ISSN: 1305-9319 e-ISSN: 1305-9327 Number: 4 Volume: 21 Year: 2025

MEDICAL JOURNAL OF BAKIRKÖY

www.bakirkoytip.org

ISSN: 1305-9319 e-ISSN: 1305-9327 Number: 4 Volume: 21 Year: 2025

MEDICAL JOURNAL OF BAKIRKÖY

Medical Journal of Bakırköy of is an official scientific Journal of University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital (BMJ)

It is published quarterly as 4 issues every year (March, June, September, December).

Please refer to the journal's webpage **(https://bakirkoymedj.org/)** for "Ethical Policy" and "Instructions to Authors".

The editorial and publication processes of the journal are shaped in accordance with the guidelines of the ICMJE, WAME, CSE, COPE, EASE, and NISO. The journal conforms with the Principles of Transparency and Best Practice in Scholarly Publishing (doaj.org/bestpractice). The Medical Journal of Bakırköy is indexed in **Web of Science; Emerging Sources Citation Index (ESCI), TUBITAK ULAKBIM TR Index, Scopus, EBSCO – Academic Search Complete, J-Gate, Embase, CINAHL Complete, ScopeMed, WorldCat, Infobase, Ulrich's Database, Ideal Online, Türkiye Atıf Dizini, Türk Medline** and **CNKI**.

The journal is published online.

Owner: University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital

Responsible Manager: Musa Çırak

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

Printing Date: December 2025
ISSN: 1305-9319 **E-ISSN:** 1305-9327
International scientific journal published quarterly.

Advisory Board

Advisory Board

Advisory Board

Advisory Board

Advisory Board

Contents

Researches

Contents

Case Report

Index

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):343-349
DOI: 10.4274/BMJ.galenos.2025.2024.7-8

Research

The Effect of Serum Uric Acid Level on Stroke Severity

Serum Ürik Asit Düzeyinin İnme Şiddetine Etkisi

Nurhan Kaya Tutar[1], Hayriye Küçükoğlu[2], Ayhan Köksal[3]

[1]University of Health Sciences Türkiye, Bağcılar Training and Research Hospital, Clinic of Neurology, İstanbul, Türkiye
[2]University of Health Sciences Türkiye, Bakırköy Prof. Dr. Mazhar Osman Training and Research Hospital, Clinic of Neurology, İstanbul, Türkiye
[3]University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital, Clinic of Neurology, İstanbul, Türkiye

ABSTRACT

Objective: Stroke remains a leading cause of disability and mortality globally, affecting millions annually. Recognizing and managing stroke risk factors is essential for reducing its prevalence and improving patient outcomes. Among the various risk factors, uric acid (UA) has garnered attention for its potential role in stroke pathophysiology. This study investigates the effect of serum UA level on stroke severity, aiming to enhance understanding and management of stroke.

Methods: A retrospective analysis was performed involving 110 patients diagnosed with acute ischemic stroke (Group 1) and 83 individuals in the healthy control group (Group 2) between August 2016 and August 2017. The average serum UA level was measured in two different groups and was analyzed together with other contributors to stroke risk. Stroke severity was evaluated in relation to serum UA levels, utilizing the National Institutes of Health Stroke Scale (NIHSS).

Results: In group 1, the average serum UA level was 5.5 mg/dL, while in Group 2, it was 4.8 mg/dL, with a statistically significant difference identified between the groups ($p<0.001$). A serum UA threshold of 5.6 mg/dL was determined ($p<0.001$), and exceeding this threshold alone was recognized as a predisposing factor for ischemic stroke, irrespective of other possible risk factors (odds ratio: 3.107; 95% confidence interval: 1.424, 6.781; $p=0.004$). There was no significant correlation between serum UA levels and NIHSS scores ($p=0.527$)

Conclusion: The findings of this study indicate that while elevated UA levels might contribute to the development of ischemic stroke, their influence on stroke severity remains uncertain. Further research is necessary to determine whether therapeutic lowering of serum UA could improve stroke outcomes.

Keywords: Ischemic stroke, uric acid, NIHSS score, stroke severity

ÖZ

Amaç: İnme, dünya çapında milyonlarca insanı etkileyen, engellilik ve ölüm oranlarında önde gelen bir neden olmaya devam etmektedir. İnme risk faktörlerinin tanınması ve yönetilmesi, inme prevalansını azaltmak ve hasta sonuçlarını iyileştirmek için esastır. Çeşitli risk faktörleri arasında ürik asit (ÜA), inme patofizyolojisindeki potansiyel rolü nedeniyle dikkat çekmiştir. Bu çalışma, serum ÜA düzeyinin inme şiddeti üzerindeki etkisini araştırarak inmenin anlaşılmasını ve yönetimini geliştirmeyi amaçlamaktadır.

Gereç ve Yöntem: Retrospektif olarak, Ağustos 2016 ve Ağustos 2017 tarihleri arasındaki 110 akut iskemik inme hastası (Grup 1) ve 83 sağlıklı kontrol (Grup 2) üzerinde değerlendirme yapıldı. İki farklı grupta ortalama serum ÜA düzeyi ölçüldü ve inme için diğer risk faktörleriyle birlikte analiz edildi. Serum ÜA düzeylerinin inme şiddetine etkisi Ulusal Sağlık Enstitüleri İnme Ölçeği (USEİÖ) skoru üzerinden değerlendirildi.

Bulgular: Grup 1'de ortalama serum ÜA seviyesi 5,5 mg/dL iken Grup 2'de 4,8 mg/dL idi ve gruplar arasında istatistiksel olarak anlamlı fark bulundu ($p<0,0019$). Serum ÜA için eşik değer 5,6 mg/dL olarak belirlendi ($p<0,001$) ve bu eşiğin aşılması tek başına, diğer potansiyel risk faktörlerinden bağımsız olarak iskemik inme gelişimi için bir risk faktörü olarak kabul edildi (olasılık oranı: 3,107; %95 güven aralığı: 1,424, 6,781; $p=0,004$). Serum ÜA düzeyleri ile USEİÖ skoru arasında herhangi bir korelasyon gözlenmedi ($p=0,527$).

Sonuç: Bu çalışma, yüksek ÜA seviyelerinin iskemik inme riskinin artmasıyla ilişkili olduğunu, ancak inme şiddetini tahmin etmede rollerinin belirsiz olduğunu göstermektedir. Serum ÜA seviyelerinin terapötik olarak düşürülmesinin inme sonuçlarını iyileştirip iyileştiremeyeceğini belirlemek için daha fazla araştırma gereklidir.

Anahtar Kelimeler: İskemik inme, ürik asit, USEİÖ skoru, inme şiddeti

Address for Correspondence: Nurhan Kaya Tutar, MD, University of Health Sciences Türkiye, Bağcılar Training and Research Hospital, Clinic of Neurology, İstanbul, Türkiye
E-mail: nurhantutar@gmail.com **ORCID ID:** orcid.org/0000-0002-2713-4282

Cite as: Kaya Tutar N, Küçükoğlu H, Köksal A. The effect of serum uric acid level on stroke severity. Med J Bakirkoy. 2025;21(4):343-349

Received: 01.08.2024
Accepted: 20.12.2024
Publication Date: 18.12.2025

INTRODUCTION

Stroke is a significant medical condition that impacts public health worldwide, ranking as one of the top causes of disability and mortality, affecting millions of people each year (1). In such an important public health issue, identifying and managing stroke risk factors is crucial to reduce the prevalence and severity of stroke. Along with non-modifiable risk factors such as age and gender, there are numerous modifiable contributors to stroke such as hypertension (HT), diabetes mellitus (DM), hyperlipidemia (HL), and abdominal obesity (2,3). The presence of many modifiable risk factors makes stroke a largely preventable neurological disease. Ongoing research is investigating emerging areas of stroke risk factors, focusing on identifying biomarkers that can predict stroke severity to improve patient outcomes.

Uric acid (UA) results from the metabolic breakdown of purines which originate from internal metabolic processes and dietary intake. Typically, UA dissolves in the blood and is excreted via the kidneys in urine (4). However, when UA levels surpass the blood's solubility threshold, urate crystals can form, leading to hyperuricemia. Hyperuricemia is primarily caused by increased production or reduced excretion of UA. Increased production can be caused by a purine-rich diet (e.g., red meat, seafood, and alcohol), increased cell turnover (e.g., cancers and chemotherapy) or genetic factors. The more common cause, reduced excretion, is usually due to renal dysfunction or certain medications such as diuretics that reduce the kidneys' ability to excrete UA.

Under physiological conditions, UA has antioxidant properties that can minimize oxidative stress by reducing the number of free radicals, thus protecting cells from damage (5). However, endogenous high serum UA levels traditionally lead to health problems such as gout and kidney stones (6). In more recent approaches, it is assumed that there is a connection between the serum UA level and vascular risk factors such as HT, DM, and coronary heart disease (CHD) (7,8). In this emerging approach, the next step requires a comprehensive investigation of the relationship between UA and stroke.

The interest in UA's role in stroke has been growing, with research exploring its potential as a biomarker for stroke risk and prognosis. Some studies indicate that high UA levels correlate with an increased risk of stroke and worse post-stroke outcomes (9-11). In contrast, others suggest that UA may have neuroprotective effects due to its antioxidant properties (12,13). Understanding the dual role of UA in the pathophysiology of ischemic cerebrovascular events is crucial for developing targeted therapies, enhancing preventive strategies, and ultimately improving clinical outcomes for patients. The purpose of this study is to investigate the influence of serum UA concentration on the severity of clinical outcomes in patients diagnosed with acute ischemic stroke.

METHODS

Study Design

A retrospective evaluation was carried out on 110 ischemic stroke patients (69 males and 41 females) who presented with acute stroke symptoms between August 2016 and August 2017 at the stroke clinic of our hospital (Group 1). The demographic data, clinical information, neuroimaging, and blood results of the subjects were obtained from our hospital's medical records. Patients were diagnosed with ischemic stroke based on neurological examination, computed tomography scans, and diffusion-weighted magnetic resonance imaging. Based on neurological examinations conducted upon patient admission, National Institutes of Health Stroke Scale (NIHSS) scores were calculated (14). The control group included 83 subjects who had no chronic neurological disease and were comparable to the patient group in both age and sex. Following their presentation to the neurology clinic with headaches, their clinical exams and neuroimaging results were normal (Group 2).

Predisposing factors for stroke were assessed in both the patient and control groups. HT was defined as systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg and/or receiving antihypertensive medications (15). DM was specified as a fasting serum glucose level ≥126 mg/dL or the use of anti-diabetic therapy (16). HL was diagnosed by fasting serum total cholesterol ≥240 mg/dL, low-density lipoprotein cholesterol ≥130 mg/dL, triglycerides ≥200 mg/dL, and/or receiving lipid-lowering medication (17). Atrial fibrillation (AF) was diagnosed using a standard 12-lead electrocardiogram. Additional relevant factors such as CHD, valvular heart disease, history of recurrent stroke, and smoking (at least five tobacco products per day) must be considered. Total number of vascular risk factors was recorded for each group. Blood samples were taken from peripheral veins after overnight fasting within 24 hours of hospitalization. Serum UA levels exceeding local laboratory reference values of 7 mg/dL in men and 5.7 mg/dL in women were considered hyperuricemia.

Exclusion criteria included conditions that could impair UA production or excretion, such as active infections, cancer, gastrointestinal disease, renal or hepatic insufficiency, chronic alcoholism, and the use of medications such as

immunosuppressants and chemotherapeutic agents. In accordance with the ethical principles of the Declaration of Helsinki, all participants gave written informed consent before the study. The study was approved by the University of Health Sciences Türkiye, Bakırköy Prof. Dr. Mazhar Osman Mental Health and Neurological Diseases Training and Research Hospital Clinical Research Ethics Committee (approval no: 569, date: 06.09.2016).

Statistical Analysis

Continuous quantitative variables are presented as means and standard deviations (SDs), while categorical variables are shown as counts (n) and percentages (%). The Shapiro-Wilk test was employed to assess the normality of the data distribution. The independent samples t-test and the Mann-Whitney U test were used to compare two independent groups based on quantitative data. The correlation between variables was evaluated using Spearman's rho test. Categorical variables were compared using Pearson chi-square or Fisher's exact tests where appropriate. Odds ratio (OR) and 95% confidence interval (CI) were calculated. The logistic backward regression method was used to investigate the cause-effect relationships. Receiver operating characteristic analysis was performed to evaluate the relationship between the two conditions. A p-value of <0.05 was considered indicative of statistical significance. All statistical analyses were conducted using SPSS version 22.0 (IBM Corp., Armonk, NY, USA) and MedCalc version 14 (MedCalc Software, Ostend, Belgium).

RESULTS

In Group 1 (n=110), 37.3% (n=69) of the patients were female, with a mean age of 63.30 (SD=14.21) years. Similarly, in Group 2 (n=83), 37.3% (n=52) were female, and the mean age was 61.87 (SD=11.75) years (Table 1). In terms of comorbidities, HT (p<0.001), HL (p=0.007), AF (p<0.001), and CHD (p=0.047) were statistically significant in Group 1, higher than in Group 2 (Table 1).

Hyperuricemia was present in 29.09% of Group 1 and 12.04% of Group 2 (Table 1). The mean serum UA levels for each group were 5.5 mg/dL in Group 1 and 4.8 mg/dL in Group 2, with Group 1 exhibiting significantly higher levels (p<0.0019) (Table 1). The UA threshold of 5.6 mg/dL was identified with an OR of 3.74, sensitivity of 49%, and specificity of 79% (p<0.001) (Table 2 and Figure 1). Multiple logistic regression analysis (95% CI) showed that HL (OR: 4.079, 95% CI: 1.806, 9.214; p=0.001), AF (OR: 12.140, 95% CI: 3.710, 39.725; p<0.001), and serum UA levels exceeding 5.6 mg/dL (OR: 3.107, 95% CI: 1.424, 6.781; p=0.004) were significantly associated with an increased risk of ischemic stroke (Table 3).

Table 1. Demographic characteristics and risk factors for stroke of each group

	Group 1 (n=110), n (%)	Group 2 (n=83), n (%)	p-value
Age (y) (mean±SD)	63.30±14.21/25-95	61.87±11.75/30-78	0.312*
Gender (female/male) (n, %)	41/69 37.3%	31/52 37.3%	1**
Hypertension	57 (51.8)	19 (22.9)	**<0.001** **3.6 (1.9-6.8)****
Diabetes mellitus	40 (36.4)	23 (27.7)	0.219**
Hyperlipidemia	49 (44.5)	21 (25.3)	**0.007** **2.4 (1.3-4.4)****
Atrial fibrillation	31 (28.2)	4 (4.8)	**<0.001** **7.6 (2.6-22.9)****
Coronary heart disease	23 (20.9)	8 (9.6)	**0.047** **2.5 (1.05-5.9)****
Valvular heart disease	1 (0.9)	0 (0.0)	1**
Smoking tobacco	9 (8.2)	0 (0.0)	**0.011** 15.6 (0.9-272.5)***
History of recurrent stroke	5 (4.5)	0 (0.0)	0.071
	Median (min.- max.)	**Median (min.- max.)**	
Total number of vascular risk factors	2 (0-6)	1 (0-3)	**<0.001***
Hyperuricemia	10 (12.04)	32 (29.09)	-
Serum UA level (mg/dL)	5.5 (2.8-11)	4.8 (2.3-9.4)	**<0.001***

*: Mann-Whitney U test, **: Chi-square test (or Fisher's exact test), ***: Odds ratio with 95% confidence interval, SD: Standard deviation, min.: Minimum, max.: Maximum, UA: Uric acid

Table 2. Thresholds and diagnostic accuracy for uric acid levels in stroke: sensitivity, specificity, and predictive values

			Patient			Control				
	Cut-off	n	PPV (%)	Sensitivity (%)	n	NPV (%)	Spesivity (%)	AUC±Se	OR (95% CI)	p-value
Serum UA level (mg/dL)	≤5.6	56	45.9%	50.9%	66	54.1%	79.5%	0.637±0.039	3.01 (1.64-5.55)*	0.001*
	>5.6	54	76.1%	49.1%	17	23.9%	20.5%	0.662±0.038	3.74 (1.954-7.2)*	<0.001*

*: ROC analysis, AUC: Area under the curve, Se: Standard error, CI: Confidence interval, PPV: Positive predictive value, NPV: Negative predictive value, UA: uric acid, ROC: Receiver operating curve, OR: Odds ratio

Table 3. Independent risk factors for ischemic stroke determined by regression analysis

	B	SD	p-value	OR (95% CI) (min.-max.)
HL	1.406	0.416	0.001	4.079 (1.806-9.214)
AF	2.497	0.605	<0.001	12.140 (3.710-39.725)
Serum UA level (>5.6 mg/dL)	1.134	0.398	0.004	3.107 (1.424-6.781)
Constant	-2.397	0.440	<0.001	0.091

HL: Hyperlipidema, AF: Atrial fibrillation, UA: Uric acid, OR: Odds ratio, CI: Confidence interval, SD: Standard deviation, min.: Minimum, max.: Maximum

Correlation analysis revealed a relationship between serum UA levels and the total number of vascular risk factors (p=0.001) (Table 4). A correlation was also found between serum UA levels and increasing age (p=0.002) (Table 4). Similarly, a correlation was found between age and the total number of vascular risk factors (p=0.005) (Table 4). No correlation was found between serum UA levels and the NIHSS scores in terms of the severity of the stroke (p=0.527) (Table 4).

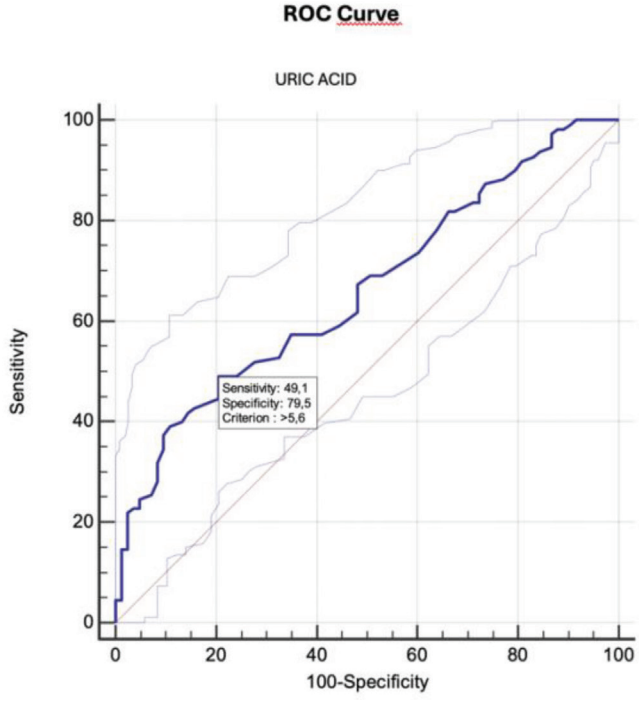


Figure 1. ROC analysis for diagnostic performance of serum uric acid levels in determining the risk of ischemic stroke

ROC: Receiver operating curve

DISCUSSION

Among the numerous biomarkers studied for their potential role in stroke, UA has garnered interest due to its dual roles as an antioxidant and a pro-oxidant. It is not yet clear whether it plays a protective role in cerebrovascular events or whether it is among the risk factors. Extensive research conducted in animal models has shown that administering UA reduces infarct size and improves neurological outcomes, supporting the idea that serum UA may serve a protective function in acute ischemic stroke (18). Clinical trials further support these findings, showing that high-dose external UA application, in addition to alteplase treatment, prevents clinical deterioration in acute ischemic stroke, as measured by the NIHSS, confirming this protective effect (19).

On the other hand, it is known that an endogenous increase in serum UA level is consistently associated with various diseases, such as HT, DM, and cardiovascular diseases (6-8). In vitro studies have also shown that hyperuricemia causes endothelial damage, thus, leading to thrombus formation, suggesting a pro-oxidant effect (20-22). This suggests that UA could be involved in the pathogenic mechanisms of stroke. These contradictory results suggest that one should be cautious before considering UA as benign in ischemic stroke.

Our study showed that serum UA concentration was significantly higher in patients with acute ischemic stroke compared with individuals without a history of stroke, suggesting that UA is a contributor to ischemic cerebrovascular events. When evaluating a novel risk factor, reported measurements need to be adjusted for known confounding factors. Although patients with acute stroke also have many vascular risk factors, our results showed that

Table 4. Correlation between serum uric acid levels, vascular risk factors, age and NIHSS scores

		p-value	p-value
Serum UA level	**Total number of vascular risk factor**	0.227	**0.001***
	Age	0.218	**0.002***
	NIH	0.061	0.527*
Total number of vascular risk factor	**Age**	0.202	**0.005***

*: Spearman's ρ test, UA: Uric acid, NIH: National Institutes of Health

high UA levels contribute to stroke susceptibility even after adjusting for these variables. This may lead us to conclude that UA plays a direct role in the pathophysiology of stroke in addition to the traditional vascular risk factors. Several meta-analyses have confirmed our findings and shown a consistent association between high UA levels and an increased incidence of ischemic cerebrovascular disease. Zhong et al. (23) demonstrated that hyperuricaemia is a predisposing factor for stroke in both sexes, focusing. Similarly, another meta-analysis showed that hyperuricemia may increase the incidence of stroke as well as the risk of mortality (24). Furthermore, research has shown that hyperuricemia not only contributes to the development of ischemic stroke, but also plays an important role in increasing the likelihood of recurrent stroke (25). This suggests that elevated UA levels may be a critical factor in both the occurrence and recurrence of cerebrovascular events.

On the other hand, there are some global epidemiologic studies suggesting that UA may not be a definitive risk factor for ischemic cerebrovascular events. In a prospective study of a large population of patients with HT but no previous stroke, hyperuricemia was not found to be a cause of ischemic cerebrovascular disease (26). Another result that has been suggested is that hyperuricemia is not associated with ischemic stroke, but it can lead to hemorrhagic stroke (27). Contradictory results in relation to hyperuricemia and cerebrovascular disease suggest that new perspectives are needed. A recent study has offered an innovative approach by highlighting the timing of UA accumulation (28). This suggests that early serum UA accumulation may lead to a higher risk of stroke compared to later accumulation, and emphasizes the need for optimal early control of serum UA levels.

The effect of UA levels on the severity of stroke is another question that needs to be clarified. Several studies suggest that higher serum UA levels are associated with more severe ischemic strokes, as determined by the scores on the NIHSS. For instance, high serum UA levels have been associated with poorer long-term functional outcomes and higher mortality rates post-stroke (29,30). This could be due to the pro-oxidant effects of serum UA under certain physiological conditions, which exacerbate neuronal injury. However, some studies do not confirm this correlation (31,32). Some studies even suggest that high baseline serum UA levels may be a biomarker for improved outcomes in patients with acute ischemic stroke (33,34). Our study indicated that there is no association between serum UA levels and the severity of stroke during the acute phase.

When discussing the effects of elevated serum UA levels, it is important to consider the factors that may influence these levels. It has been shown in the literature that vascular risk factors for stroke, such as age, HT, DM, and HL, contribute to higher serum UA levels due to impaired renal function (35). This complicates the cause-effect relationship between stroke, vascular risk factors, and serum UA levels. Our results further highlight this complexity. Consistent with the literature, we observed a corresponding increase in the number of vascular risk factors as serum UA levels increased. Furthermore, in our study, both serum UA levels and the number of vascular risk factors increased with age. This suggests that the relationship between UA and stroke risk is anything but straightforward; instead, it is shaped by age and the accumulation of vascular risks over time. These findings emphasise the importance of considering the broader clinical context when interpreting UA levels in relation to stroke risk.

In recent years, an increasing number of epidemiological and experimental studies have shown that UA can trigger pathological processes even at physiological limits, at which it is still soluble. Virdis et al. (36) showed that serum UA levels increase all-cause mortality at a cut-off value of 4.7 mg/dL, which is well below the standard thresholds used in clinical practice to define hyperuricaemia. Desideri et al. (37) even suggested redefining the "optimal health range" for UA levels as <6 mg/dL for both genders. At the outset of the study, we stratified UA levels based on gender-specific thresholds to account for the well-established physiological differences in UA metabolism between sexes. This approach ensured that we accurately captured the prevalence of hyperuricemia in both male and female populations according to their respective normal ranges. However, as

the study progressed and we analyzed serum UA levels as an independent risk factor for stroke, our findings indicated that the threshold value associated with an increased risk of stroke was consistently above 5.6 mg/dL, regardless of sex. By specifying a uniform threshold value of 5.6 mg/dL for both genders, we aimed to emphasize an important clinical finding: above this value, the risk of stroke increases regardless of the patient's gender. This finding could simplify clinical assessments and interventions aimed at stroke prevention, providing a more straightforward and universally applicable threshold.

Study Limitations

A limitation of our study is that it is retrospective and the number of patients is relatively small. Another limitation of the study is that although the groups were matched in terms of age and gender, there were some differences in vascular risk factors. We acknowledge that a study design with matched controls free of vascular risk factors might provide further insight into the independent role of UA. However, we believe that the current analysis offers valuable insights into the relationship between UA and stroke in a more clinically representative population. Given the ongoing debate in the literature regarding the role of UA in stroke, further studies are needed to provide greater clarity on this issue.

CONCLUSION

Understanding the effect of serum UA levels on stroke severity has significant clinical implications. Monitoring serum UA levels in acute stroke patients could aid risk stratification and individualized treatment planning. However, given the conflicting data, clinicians should interpret serum UA levels cautiously, considering the broader clinical context and comorbidities. Further research is needed to clarify whether lowering serum UA levels therapeutically could mitigate stroke severity or improve outcomes.

ETHICS

Ethics Committee Approval: The study was approved by the University of Health Sciences Türkiye, Bakırköy Prof. Dr. Mazhar Osman Mental Health and Neurological Diseases Training and Research Hospital Clinical Research Ethics Committee (approval no: 569, date: 06.09.2016).

Informed Consent: All participants gave written informed consent before the study.

FOOTNOTES

Authorship Contributions

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

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

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

REFERENCES

1. Guzik A, Bushnell C. Stroke Epidemiology and risk factor management. Continuum (Minneap Minn). 2017;23(1, Cerebrovascular Disease):15-39.
2. Alloubani A, Saleh A, Abdelhafiz I. Hypertension and diabetes mellitus as a predictive risk factors for stroke. Diabetes Metab Syndr. 2018;12:577-84.
3. Virani SS, Alonso A, Aparicio HJ, Benjamin EJ, Bittencourt MS, Callaway CW, et al. Heart disease and stroke statistics-2021 update: a report from the American Heart Association. Circulation. 2021;143:e254-e743.
4. Maiuolo J, Oppedisano F, Gratteri S, Muscoli C, Mollace V. Regulation of uric acid metabolism and excretion. Int J Cardiol. 2016;213:8-14. Erratum in: Int J Cardiol. 2023;387:131126.
5. El Ridi R, Tallima H. Physiological functions and pathogenic potential of uric acid: a review. J Adv Res. 2017;8:487-93.
6. Tsai CW, Lin SY, Kuo CC, Huang CC. Serum uric acid and progression of kidney disease: a longitudinal analysis and mini-review. PLoS One. 2017;12:e0170393.
7. Gan TM, Ye YY, Mo GL, Li JY. Progress of uric acid in cardiovascular disease. Cardiovasc Endocrinol Metab. 2024;13:e0300.
8. Yu W, Cheng JD. Uric acid and cardiovascular disease: an update from molecular mechanism to clinical perspective. Front Pharmacol. 2020;11:582680.
9. Ae R, Kanbay M, Kuwabara M. The causality between the serum uric acid level and stroke. Hypertens Res. 2020;43:354-6.
10. Zhang S, Liu L, Huang YQ, Lo K, Tang S, Feng YQ. The association between serum uric acid levels and ischemic stroke in essential hypertension patients. Postgrad Med. 2020;132:551-8.
11. Yang XL, Kim Y, Kim TJ, Jung S, Kim CK, Lee SH. Association of serum uric acid and cardioembolic stroke in patients with acute ischemic stroke. J Neurol Sci. 2016;370:57-62.
12. Vila E, Solé M, Masip N, Puertas-Umbert L, Amaro S, Dantas AP, et al. Uric acid treatment after stroke modulates the Krüppel-like factor 2-VEGF-A axis to protect brain endothelial cell functions: impact of hypertension. Biochem Pharmacol. 2019;164:115-28.
13. Aliena-Valero A, López-Morales MA, Burguete MC, Castelló-Ruiz M, Jover-Mengual T, Hervás D, et al. Emergent uric acid treatment is synergistic with mechanical recanalization in improving stroke outcomes in male and female rats. Neuroscience. 2018;388:263-73.
14. Lyden P. Using the national institutes of health stroke scale: a cautionary tale. Stroke. 2017;48:513-9.
15. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, et al. 2018 ESC/ESH guidelines for the management of arterial hypertension: the task force for the management of arterial hypertension of the European Society of Cardiology and the European Society of Hypertension: the task force for the management of arterial hypertension of the European Society of Cardiology and the European Society of Hypertension. J Hypertens. 2018;36:1953-2041. Erratum in: J Hypertens. 2019;37:226.
16. American Diabetes Association Professional Practice Committee. 2. classification and diagnosis of diabetes: standards of medical care in diabetes-2022. Diabetes Care. 2022;45(Suppl 1):S17-38.

17. Grundy SM, Stone NJ. 2018 American Heart Association/American College of Cardiology Multisociety Guideline on the management of blood cholesterol: primary prevention. JAMA Cardiol. 2019;4:488-9.
18. Aliena-Valero A, Rius-Pérez S, Baixauli-Martín J, Torregrosa G, Chamorro Á, Pérez S, et al. Uric acid neuroprotection associated to IL-6/STAT3 signaling pathway activation in rat ischemic stroke. Mol Neurobiol. 2021;58:408-23.
19. Amaro S, Laredo C, Renú A, Llull L, Rudilosso S, Obach V, et al. Uric acid therapy prevents early ischemic stroke progression: a tertiary analysis of the URICO-ICTUS trial (efficacy study of combined treatment with uric acid and r-tPA in acute ischemic stroke). Stroke. 2016;47:2874-6.
20. Kang DH, Ha SK. Uric acid puzzle: dual role as anti-oxidantand pro-oxidant. Electrolyte Blood Press. 2014;12:1-6.
21. Xie D, Zhao H, Lu J, He F, Liu W, Yu W, et al. High uric acid induces liver fat accumulation via ROS/JNK/AP-1 signaling. Am J Physiol Endocrinol Metab. 2021;320:E1032-43.
22. Yang L, Chang B, Guo Y, Wu X, Liu L. The role of oxidative stress-mediated apoptosis in the pathogenesis of uric acid nephropathy. Ren Fail. 2019;41:616-22.
23. Zhong C, Zhong X, Xu T, Xu T, Zhang Y. Sex-specific relationship between serum uric acid and risk of stroke: a dose-response meta-analysis of prospective studies. J Am Heart Assoc. 2017;6:e005042.
24. Qiao T, Wu H, Peng W. The relationship between elevated serum uric acid and risk of stroke in adult: an updated and dose-response meta-analysis. Front Neurol. 2021;12:674398.
25. Li M, Wang H, Gao Y. Serum uric acid levels and recurrence rate of ischemic stroke: a meta-analysis. Horm Metab Res. 2023;55:493-7.
26. Shi X, Yang J, Wang L, Zhao M, Zhang C, He M, et al. Prospective study of serum uric acid levels and stroke in a Chinese hypertensive cohort. Clin Exp Hypertens. 2017;39:527-31.
27. Wang A, Tian X, Zuo Y, Chen S, Mo D, Zhang L, et al. Effect of changes in serum uric acid on the risk of stroke and its subtypes. Nutr Metab Cardiovasc Dis. 2022;32:167-75.
28. Tian X, Chen S, Xu Q, Wang P, Zhang Y, Zhang X, et al. Cumulative serum uric acid exposure and its time course with the risk of incident stroke. Stroke. 2023;54:2077-86.
29. Zhang W, Cheng Z, Fu F, Zhan Z. Serum uric acid and prognosis in acute ischemic stroke: a dose-response meta-analysis of cohort studies. Front Aging Neurosci. 2023;15:1223015.
30. Storhaug HM, Norvik JV, Toft I, Eriksen BO, Løchen ML, Zykova S, et al. Uric acid is a risk factor for ischemic stroke and all-cause mortality in the general population: a gender specific analysis from The Tromsø Study. BMC Cardiovasc Disord. 2013;13:115.
31. Zhong J, Cai H, Zhang Z, Wang J, Xiao L, Zhang P, et al. Serum uric acid and prognosis of ischemic stroke: cohort study, meta-analysis and Mendelian randomization study. Eur Stroke J. 2024;9:235-43.
32. Zhang M, Wang Y, Wang K, Yin R, Pan X, Ma A. Association between uric acid and the prognosis of acute ischemic stroke: a systematic review and meta-analysis. Nutr Metab Cardiovasc Dis. 2021;31:3016-23.
33. Wang Z, Lin Y, Liu Y, Chen Y, Wang B, Li C, et al. Serum uric acid levels and outcomes after acute ischemic stroke. Mol Neurobiol. 2016;53:1753-9.
34. Lei Z, Cai J, Hong H, Wang Y. Serum uric acid level and outcome of patients with ischemic stroke: a systematic review and meta-analysis. Neurologist. 2019;24:121-31.
35. Xu Q, Fan X, Chen G, Ma J, Ye W, Ai S, et al. New-onset metabolic syndrome is associated with accelerated renal function decline partially through elevated uric acid: an epidemiological cohort study. Front Endocrinol (Lausanne). 2024;15:1328404.
36. Virdis A, Masi S, Casiglia E, Tikhonoff V, Cicero AFG, Ungar A, et al. Identification of the uric acid thresholds predicting an increased total and cardiovascular mortality over 20 years. Hypertension. 2020;75:302-8.
37. Desideri G, Castaldo G, Lombardi A, Mussap M, Testa A, Pontremoli R, et al. Is it time to revise the normal range of serum uric acid levels? Eur Rev Med Pharmacol Sci. 2014;18:1295-306.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):350-356
DOI: 10.4274/BMJ.galenos.2025.2025.4-12

Research

Evaluating the Efficacy and Safety of Carotid Artery Stenting: A Retrospective Analysis

Karotis Arter Stentlemenin Etkinliği ve Güvenliğinin Değerlendirilmesi: Retrospektif Bir Analiz

Serkan Civlan[1], Nevzat Doğukan Erbek[1], Işık Tekin[2], Muhammet Arslan[3], Emrah Keskin[4], Rasim Asar[1], Hakkı Peker[3], Fatih Yakar[1]

[1]Pamukkale University Faculty of Medicine, Department of Neurosurgery, Denizli, Türkiye
[2]Pamukkale University Faculty of Medicine, Department of Cardiology, Denizli, Türkiye
[3]Pamukkale University Faculty of Medicine, Department of Radiology, Denizli, Türkiye
[4]Zonguldak Bülent Ecevit University Faculty of Medicine, Department of Neurosurgery, Zonguldak, Türkiye

ABSTRACT

Objective: Stroke is the second leading cause of death and the third leading cause of disability globally, with carotid artery stenosis contributing to approximately 20% of cases. Carotid artery stenting (CAS) and carotid endarterectomy (CEA) are effective treatments to reduce stroke risk. While some studies report higher complication rates for CAS compared to CEA, others demonstrate comparable outcomes, emphasizing the importance of patient selection and procedural optimization. This study evaluates the real-world safety and effectiveness of CAS in 52 patients at a single institution.

Methods: This retrospective analysis included 52 patients who underwent CAS between 2020 and 2024. Inclusion criteria were ≥50% stenosis for symptomatic patients and >60% stenosis for asymptomatic ones. Dual antiplatelet therapy was initiated preoperatively, and distal filter embolic protection was used in all procedures. Neurological assessments and radiological imaging were performed pre- and post-procedure. Complications were categorized as periprocedural or post-procedural, and follow-ups were conducted at three, six, and twelve months.

Results: The cohort included 52 patients (78.8% symptomatic, mean stenosis rate: 80.3%±12.5%). Periprocedural ischemic stroke occurred in 5.8% of patients, and asymptomatic diffusion-restricted areas were detected in 34.6% of patients. One patient (2.2%) experienced symptomatic intracerebral hemorrhage. The overall periprocedural stroke and death rate was 7.7%.

Conclusion: CAS is a minimally invasive, effective option for treating carotid artery stenosis when patient selection and procedural protocols are optimized. Ongoing advancements in techniques and devices are anticipated to reduce complications further, supporting CAS as a safe alternative to CEA in selected patients.

Keywords: Carotid stenosis, carotid endarterectomy, stents, stroke

ÖZ

Amaç: İnme, dünya genelinde ölüm nedenleri arasında ikinci, sakatlık nedenleri arasında üçüncü sırada yer almakta olup, inme olgularının yaklaşık %20'sinden karotis arter darlığı sorumludur. Karotis arter stentleme (KAS) ve karotis endarterektomi (KEA), inme riskini azaltmada etkili tedavi yöntemleridir. Bazı çalışmalar, KAS'nin KEA'ya kıyasla daha yüksek komplikasyon oranlarına sahip olduğunu belirtirken, diğerleri benzer sonuçlar sunmakta ve hasta seçimi ile işlem optimizasyonunun önemini vurgulamaktadır. Bu çalışma, tek bir merkezde 52 hastada KAS'nin gerçek yaşam koşullarındaki güvenliği ve etkinliğini değerlendirmektedir.

Gereç ve Yöntem: Bu retrospektif analizde, 2020-2024 yılları arasında KAS uygulanan 52 hasta incelenmiştir. Dahil edilme kriterleri, semptomatik hastalar için ≥%50, asemptomatik hastalar için >%60 darlık olarak belirlenmiştir. İşlem öncesinde çift antiplatelet tedavi başlatılmış ve tüm işlemlerde distal filtre emboli koruma cihazı kullanılmıştır. İşlem öncesi ve sonrası nörolojik değerlendirmeler ve radyolojik görüntülemeler gerçekleştirilmiştir. Komplikasyonlar işlem sırasında ve erken dönemde ortaya çıkan ve işlem sonrası geç dönemde ortaya çıkan olarak sınıflandırılmış, takipler üçüncü, altıncı ve on ikinci aylarda yapılmıştır.

Address for Correspondence: Serkan Civlan, Asst. Prof., Pamukkale University Faculty of Medicine, Department of Neurosurgery, Denizli, Türkiye
E-mail: serkancivlan@hotmail.com **ORCID ID:** orcid.org/0000-0001-8915-8186

Cite as: Civlan S, Erbek ND, Tekin I, Arslan M, Keskin E, Asar R, et al. Evaluating the efficacy and safety of carotid artery stenting: a retrospective analysis. Med J Bakirkoy. 2025;21(4):350-356

Received: 26.04.2025
Accepted: 31.05.2025
Publication Date: 18.12.2025

ÖZ

Bulgular: Çalışmaya 52 hasta (%78,8 semptomatik, ortalama darlık oranı: %80,3±%12,5) dahil edilmiştir. İşlem sırasında ve erken dönemde iskemik inme %5,8 hastada görülmüş, asemptomatik difüzyon kısıtlı alanlar %34,6 hastada tespit edilmiştir. Bir hastada (%2,2) semptomatik intrakraniyal hemoraji meydana gelmiştir. Toplam işlem sırasında ve erken dönemde inme ve ölüm oranı %7,7 olarak bulunmuştur.

Sonuç: KAS, hasta seçimi ve protokoller optimize edildiğinde karotis arter darlığı tedavisinde minimal invaziv ve etkili bir seçenektir. Teknikler ve cihazlarda devam eden ilerlemelerle komplikasyon oranlarının daha da azalması beklenmekte, KAS'ın seçilmiş hastalarda KEA'ya güvenli bir alternatif olarak desteklenmesi gerektiği düşünülmektedir.

Anahtar Kelimeler: Karotis darlığı, karotis endarterektomi, stentler, inme

INTRODUCTION

Stroke is the second leading cause of death and the third leading cause of disability worldwide (1). Carotid artery stenosis is responsible for approximately 20% of all strokes (2). Carotid artery stenting (CAS) and carotid endarterectomy (CEA) are treatment modalities that have been shown to reduce the risk of stroke in both asymptomatic and symptomatic patients with carotid artery stenosis. Although several prospective studies, such as SPACE, EVA-3S, and ICSS (3-5), have reported higher complication rates for CAS compared with CEA, other studies, including SAPPHIRE, CREST, ACT I, and ACST-II (6-9), have demonstrated the non-inferiority of CAS with respect to composite primary outcomes, including stroke, myocardial infarction, and death. Case series from various regions of the world continue to be published, demonstrating the efficacy and low-risk profile of CAS (10-14).

There is still no clear consensus on the optimal treatment, primarily due to the variability in results among these studies. Differing outcomes are likely due to a lack of standardization in patient selection, procedural methods, and physician experience. Given these discrepancies, we believe that CAS is a valuable treatment option when patient selection is carefully individualized. To further assess the real-world outcomes of CAS, we performed CAS on 52 carefully selected patients at a single institution. This was done to evaluate the effectiveness and safety of CAS in a clinical setting.

METHODS

Study Design

We retrospectively analyzed the outcomes of CAS performed on 52 patients between 2020 and 2024 at a single institution. Written informed consent was obtained from all patients. The study was conducted in accordance with the rules of the Declaration of Helsinki and was approved by the Pamukkale University Non-Interventional Clinical Research Ethics Committee (approval no: E-60116787-020-617909, date: 02.12.2024).

Patient Selection

The study included both symptomatic and asymptomatic patients. Inclusion criteria were set at ≥50% stenosis for symptomatic cases and >60% stenosis for asymptomatic ones. Additionally, two patients who had prior unsuccessful CEA attempts were included. Symptomatic carotid artery stenosis was defined as the occurrence of a cerebrovascular event—such as stroke, transient ischemic attack, or amaurosis fugax—within 180 days prior to the procedure.

Procedural Details

Three neurosurgeons, each with at least five years' experience in carotid stenting and angioplasty, performed the procedures. For lesion evaluation, we used combinations of Doppler ultrasonography, magnetic resonance angiography (MRA), computed tomography (CT) angiography, and digital subtraction angiography (DSA).

Pre-procedural Preparation

Dual antiplatelet therapy with aspirin plus clopidogrel or ticagrelor was initiated at least five days before the procedure in all patients. The CAS procedures were performed using a transfemoral approach under local anesthesia.

Intra-procedural Management

After arterial sheath insertion, intravenous heparin was administered, and the activated clotting time was monitored. A minimum activated clotting time of 300 seconds was required to proceed with the intervention. Distal filter protection devices were used as the standard embolic protection method in all cases. Open-cell stents were deployed in 36 cases, and closed-cell stents were deployed in 16 cases.

Complications and Outcomes

Documented complications included retinal stroke, ischemic stroke, asymptomatic hyperintense lesions on diffusion-weighted imaging (DWI), in-stent plaque protrusion, hyperperfusion syndrome, symptomatic intracerebral hemorrhage, morbidity, and mortality. Periprocedural outcomes were defined as complications occurring during the procedure or within 30 days post-intervention, while

post-procedural outcomes were defined as complications occurring beyond 30 days.

Assessment Protocols

A comprehensive neurological examination was performed before the procedure, immediately after the intervention, and daily until discharge. The modified Rankin scale (mRS) was used to assess functional prognosis. Additionally, magnetic resonance imaging (including DWI, apparent diffusion coefficient, and fluid-attenuated inversion recovery sequences) and CT scans were conducted within 12-24 hours post-procedure to evaluate outcomes. All patients underwent MRA at three months and DSA at six months, with annual outpatient MRA follow-ups thereafter.

Definitions

Stroke was defined as a neurological deficit persisting for more than 24 hours, accompanied by corresponding hyperintensity on DWI. Minor morbidity was classified as a 1-point increase in the mRS score, while major morbidity was classified as a 2-point or greater increase. High-risk criteria, as defined in the SAPPHIRE study, include clinically significant cardiac disease, severe pulmonary disease, contralateral carotid occlusion, contralateral laryngeal nerve palsy, previous radical neck surgery or radiation therapy to the neck, recurrent stenosis after endarterectomy, and age over 80 years.

Statistical Analysis

All data were processed with SPSS Statistics, version 26 (IBM Corp., Armonk, NY, USA). Continuous variables are presented as mean±standard deviation when normally distributed (Shapiro-Wilk $p>0.05$) and as median with interquartile range when non-parametric. Categorical variables are summarised as counts and percentages.

The primary end-point was any stroke or death occurring during the procedure or within 30 days after the procedure. Secondary end-points were clinically silent DWI lesions, symptomatic intracerebral haemorrhage, hyperperfusion syndrome, and restenosis of greater than 50% on follow-up angiography.

Exploratory univariable logistic regression was applied to screen potential predictors of the primary end-point. Variables with $p<0.10$ were considered for inclusion in a forward stepwise multivariable model; however, only four primary events occurred and none reached statistical significance, so multivariable analysis was not feasible. Restenosis-free survival was illustrated with Kaplan-Meier curves. Two-sided $p<0.05$ was considered statistically significant for all planned analyses.

Owing to the modest sample size and low event rate, the statistical evaluation remains chiefly descriptive; effect estimates are reported with 95% confidence intervals when calculable, and results should be interpreted with caution.

RESULTS

The characteristics of the study populations and lesion details are summarized in Table 1. A total of 52 patients were included in the study, with 40 males and 12 females. Among these patients, 41 (78.8%) had symptomatic lesions. According to the SAPPHIRE study (6) criteria, 48% of the patients had at least one high-risk factor for CEA and the majority of these factors were related to clinically significant cardiac disease. The mean stenosis rate was 80.3%±12.5%. Five (9.6%) patients had stenosis in the contralateral internal carotid artery (ICA), and five (9.6%) had occlusion of the contralateral ICA.

Asymptomatic high-intensity areas on DWI were detected in 18 patients (34.6%). Symptomatic ischemic stroke occurred in three patients (5.8%). All three patients had symptomatic carotid stenosis, and one had a high-risk factor for CEA. In the first patient, plaque protrusion into the distal filter was observed during the procedure (Figure 1A). Following filter retrieval, the patient developed dysphasia, left-sided facial paralysis, and weakness of the left extremities. Occlusion of the superior trunk of the M2 segment of the right middle cerebral artery was identified (Figure 1B). Thrombectomy was performed using a stent retriever, resulting in successful vessel recanalization (Figure 1C). The patient achieved full recovery and was discharged asymptomatic.

The second patient presented with right-sided paresis affecting the upper and lower extremities prior to the procedure. Dysphasia occurred during the procedure. No significant vascular occlusion was observed on DSA. Post-procedural DWI revealed small, multifocal diffusion-restricted areas. The patient returned to their baseline neurological status within a two-week follow-up period.

The third patient developed dysphasia during the procedure. No significant vascular occlusion was observed on DSA. DWI demonstrated a diffusion-restricted area in the perfusion territory supplied by the distal branches of the callosomarginal artery. The patient was managed medically; the symptoms resolved completely within one week of follow-up.

One patient (2.2%) developed post-procedural confusion and left-sided hemiparesis. Brain CT revealed an intraparenchymal hematoma in the frontoparietal region. The patient was treated medically and discharged with an

mRS score of 2 after two weeks' follow-up. This patient had 99% carotid stenosis and multiple high-risk factors for CEA. Periprocedural complications are summarized in Table 2.

DISCUSSION

Carotid artery stenosis is one of the major risk factors for stroke. CEA is a well-established treatment method for carotid artery stenosis and has been performed since the 1950s (15). The surgical techniques for CEA are highly standardized, and their effectiveness and complication rates are well documented. In contrast, CAS is a relatively recent treatment that was first performed in the early 1990s. Its popularity has grown since the CAVATAS trial, the first large prospective, randomized trial published in 2001, which reported similar major complication rates but lower minor complication rates for CAS than for CEA (16). The results of CAVATAS were further supported by the SAPPHIRE trial (6,16) in 2004.

Since then, numerous studies have been conducted to compare these two treatment modalities and assess the effectiveness and complication rates of CAS. However, the results have been conflicting due to non-standardised techniques, evolving devices, variations in patient populations, and differences in physician's experience.

Historically, CEA was associated with high periprocedural complication rates, with early series reporting stroke and death rates approaching 20% (17). In contrast, contemporary studies have demonstrated significant improvements, with recent trials reporting 30-day stroke and death rates between 3% and 5.7% (18-20). By comparison, recent literature reports periprocedural stroke and death rates for CAS ranging from 2.1% to 8.2% (10-14).

In our study, we carefully selected patients based on their vascular-anatomical suitability for endovascular treatment, surgical risk profile, and patient preference. The periprocedural rate of stroke and death was 7.7%, which is at the higher end of the range reported in larger CAS studies. Although this rate exceeds the average reported for contemporary CEA, the limited sample size in our study may have contributed to reduced statistical reliability.

Table 1. Demographic data of the patients in the series

Age (years)	70.1±9
Gender (n, %)	
Male	40 (77)
Female	12 (23)
Symptomatic lesion (%, in 180 days)	41 (78.8)
Risk factors (n, %)	
Hypertension	36 (69.2)
Diabetes	27 (51.9)
Dyslipidemia	7 (13.5)
Smoker	3 (5.8)
Previous ischemic heart disease (n, %)	23 (44.2)
Previous peripheral artery disease (n, %)	2 (3.8)
High-risk factor for CEA-related complications (described in SAPPHIRE study) (n, %)	25 (48)
Clinically significant cardiac disease	23 (44.2)
Severe pulmonary disease	0 (0)
Contralateral carotid occlusion	5 (9.6)
Contralateral laryngeal-nerve palsy	0 (0)
Previous radical neck surgery or radiation therapy	1 (1.9)
Recurrent stenosis after endarterectomy	3 (5.8)
Age >80 years old	6 (11.5)
Previous CEA (n, %)	3 (5.8)
Contralateral stenosis (n, %)	5 (9.6)
Contralateral occlusion (n, %)	5 (9.6)
Stenosis rate in DSA (n, %) (NASCET)	80.3 (50-99)
Antiplatelet regimen (n, %)	
ASA+clopidogrel	39 (75)
ASA+ticagrelor	13 (25)
Pre-stent PTA (n, %)	24 (46.2)
Post-stent PTA (n, %)	33 (63.5)
Stent brand (n, %)	
Boston Scientific Wallstent™ (closed-cell)	16 (30.8)
Medtronic Protégé™ RX Stent (open-cell)	36 (69.2)

CEA: Carotid endarterectomy, PTA: Percutaneous transluminal angioplasty, ASA: Acetylsalicylic acid (aspirin), DSA: Digital subtraction angiography, NASCET: North American Symptomatic Carotid Endarterectomy Trial

Table 2. Periprocedural complications of patients

Periprocedural complications	**(n, %)**
Asymptomatic DWI high spot	18 (34.6)
Stroke symptomatic group	3 (5.8)
Retinal stroke	0
In-stent plaque protrusion	1
Hyperperfusion syndrome	0
Symptomatic ICH	1 (1.9)
Minor morbidity (mRS score of 1)	0
Major morbidity (mRS score of ≥2)	1 (1.9)
Restenosis or occlusion	0
Reintervention PTA	0
Reintervention restenting	0

The patient with symptomatic ICH was managed conservatively. Among the three patients with stroke, two were treated conservatively, while one underwent thrombectomy.
DWI: Diffusion-weighted imaging, ICH: Intracerebral hemorrhage, mRS: Modified Rankin scale, PTA: Percutaneous transluminal angioplasty

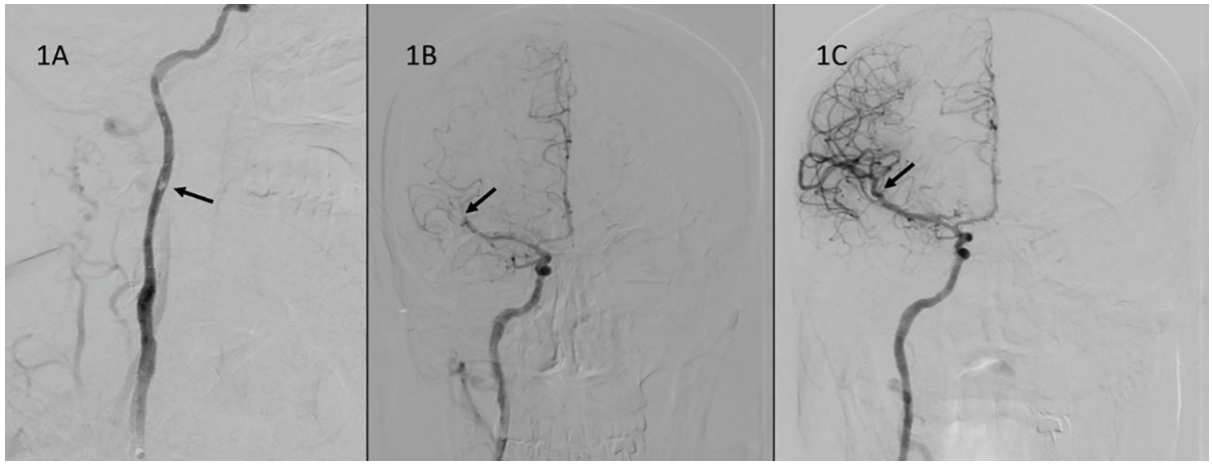


Figure 1. **1A)** Plaque protrusion is seen in distal filter after stent placement. **1B)** An occlusion of the superior trunk of the M2 segment. **1C)** Complete recanalization after successful thrombectomy. The patient achieved full recovery and was discharged asymptomatic

Notably, asymptomatic high-intensity areas on DWI were detected in 18 patients (34.6%). Although such diffusion-restricted lesions have been associated in the literature with an increased risk of future cerebrovascular events or cognitive impairment (21,22), we did not observe any related clinical manifestations in our cohort. This may be attributed to the relatively short follow-up period, which could have been insufficient to capture delayed or subtle effects.

Two major variables related to CAS devices are stent design and cerebral protection device design. Several studies have compared open-cell and closed-cell stents, but the findings remain controversial. In 2007, Bosiers et al. (23) reported that open-cell stents were associated with higher complication rates, particularly among symptomatic patients. However, a year later, Schillinger et al. (24) reported no significant difference in complication rates between open- and closed-cell stents. More recently, Faateh et al. (25) reported that closed-cell stents were associated with an increased risk of in-hospital stroke and death, especially for lesions located at the carotid bifurcation. They suggested that this might be due to the relatively low conformability of closed-cell stents in tortuous, diameter-mismatched bifurcation anatomy (25). In this study, we used open-cell stents in 36 procedures and closed-cell stents in 16 procedures.

The use of cerebral protection devices during CAS is strongly recommended, as CAS without these devices has been shown to be significantly associated with higher rates of periprocedural stroke and death (26,27). Although some studies suggest that proximal balloon occlusion is associated with lower complication rates than distal filter protection (28,29), larger studies have found that neither method is superior for periprocedural stroke and death rates (30,31). We did not perform balloon occlusion, but used distal filter protection in all cases.

Compared with CEA, CAS is less invasive and is typically performed under local anesthesia. Although concerns and debates remain about the complication rates of CAS compared with CEA, its minimally invasive nature makes it a preferable option, particularly for patients at high surgical risk. We believe that with careful patient selection and the appropriate choice of devices, CAS represents an effective and safe treatment option for stroke prevention.

It is important to recognize that CAS is still evolving, and as newer devices are developed, anatomical and technical challenges will likely be addressed, resulting in greater efficacy and reduced complication rates.

Study Limitations

This study is limited by a small patient cohort, which reduces statistical power and the precision of effect estimates. The limited number of events (n=4) precluded multivariable analysis, which should be considered a limitation in interpreting the results. The six- to twelve-month follow-up period is relatively brief, precluding evaluation of late restenosis and longer-term clinical outcomes. Its retrospective nature imposes inherent selection bias and reduces control over data completeness and uniformity. Outcomes cannot be directly compared with CEA because no contemporaneous surgical control group was included. Finally, all procedures were performed by a small, highly experienced team of physicians; therefore, the results may not be generalisable to centres with differing levels of expertise.

CONCLUSION

Our study highlights the viability of CAS as a safe and effective treatment for carotid artery stenosis, provided appropriate patient selection and procedural methods are employed. While ongoing advances and standardization are necessary to further enhance outcomes, CAS demonstrates promise comparable to that of CEA, with a favorable balance between efficacy and complication rates. As techniques and devices continue evolving, CAS is poised to become an increasingly integral component of stroke prevention, warranting its consideration alongside traditional methods in clinical practice.

ETHICS

Ethics Committee Approval: The study was conducted in accordance with the rules of the Declaration of Helsinki and was approved by the Pamukkale University Non-Interventional Clinical Research Ethics Committee (approval no: E-60116787-020-617909, date: 02.12.2024).

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

FOOTNOTES

Authorship Contributions

Surgical and Medical Practices: S.C., N.D.E., I.T., M.A., E.K., H.P., F.Y., Concept: S.C., I.T., Design: S.C., N.D.E., Data Collection or Processing: S.C., N.D.E., I.T., M.A., E.K., R.A., H.P., F.Y., Analysis or Interpretation: S.C., N.D.E., I.T., Literature Search: S.C., N.D.E., I.T., Writing: S.C., N.D.E., I.T.

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

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

REFERENCES

1. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol. 2021;20:795-820.
2. Sardar P, Chatterjee S, Aronow HD, Kundu A, Ramchand P, Mukherjee D, et al. Carotid artery stenting versus endarterectomy for stroke prevention: a meta-analysis of clinical trials. J Am Coll Cardiol. 2017;69:2266-75.
3. SPACE Collaborative Group; Ringleb PA, Allenberg J, Brückmann H, Eckstein HH, Fraedrich G, et al. 30 day results from the SPACE trial of stent-protected angioplasty versus carotid endarterectomy in symptomatic patients: a randomised non-inferiority trial. Lancet. 2006;368:1239-47. Erratum in: Lancet. 2006;368:1238.
4. Mas JL, Arquizan C, Calvet D, Viguier A, Albucher JF, Piquet P, et al. Long-term follow-up study of endarterectomy versus angioplasty in patients with symptomatic severe carotid stenosis trial. Stroke. 2014;45:2750-6.
5. International Carotid Stenting Study investigators; Ederle J, Dobson J, Featherstone RL, Bonati LH, van der Worp HB, et al. Carotid artery stenting compared with endarterectomy in patients with symptomatic carotid stenosis (International Carotid Stenting Study): an interim analysis of a randomised controlled trial. Lancet. 2010;375:985-97. Erratum in: Lancet. 2010;376:90.
6. Yadav JS, Wholey MH, Kuntz RE, Fayad P, Katzen BT, Mishkel GJ, et al. Protected carotid-artery stenting versus endarterectomy in high-risk patients. N Engl J Med. 2004;351:1493-501.
7. Brott TG, Hobson RW 2nd, Howard G, Roubin GS, Clark WM, Brooks W, et al. Stenting versus endarterectomy for treatment of carotid-artery stenosis. N Engl J Med. 2010;363:11-23. Erratum in: N Engl J Med. 2010;363:498. Erratum in: N Engl J Med. 2010;363:198.
8. Rosenfield K, Matsumura JS, Chaturvedi S, Riles T, Ansel GM, Metzger DC, et al. Randomized trial of stent versus surgery for asymptomatic carotid stenosis. N Engl J Med. 2016;374:1011-20.
9. Halliday A, Bulbulia R, Bonati LH, Chester J, Cradduck-Bamford A, Peto R, et al. Second asymptomatic carotid surgery trial (ACST-2): a randomised comparison of carotid artery stenting versus carotid endarterectomy. Lancet. 2021;398:1065-73.
10. Mudra H, Staubach S, Hein-Rothweiler R, Segerer M, Strohm H, Weber H, et al. Long-term outcomes of carotid artery stenting in clinical practice. Circ Cardiovasc Interv. 2016;9:e003940.
11. Ronchey S, Praquin B, Orrico M, Serrao E, Ciceroni C, Alberti V, et al. Outcomes of 1000 carotid wallstent implantations: single-center experience. J Endovasc Ther. 2016;23:267-74.
12. Omura Y, Imamura H, Tani S, Adachi H, Sakaguchi M, Todo K, et al. Treatment results of carotid artery stenting with an open-cell stent: analysis of 734 consecutive cases at a single center. World Neurosurg. 2024;187:e453-9.
13. Fuse Y, Kojima T, Shintai K, Seki Y. Long-term outcomes of carotid artery stenting: a single-center experience. Neurol Med Chir (Tokyo). 2020;60:121-5.
14. Karaduman BD, Ayhan H, Keleş T, Bozkurt E. Efficacy and safety of carotid artery stenting: experience of a single center. Turk Kardiyol Dern Ars. 2020;48:646-55.
15. Easton JD. History of carotid endarterectomy then and now: personal perspective. Stroke. 2014;45:e101-3.
16. Endovascular versus surgical treatment in patients with carotid stenosis in the Carotid and Vertebral Artery Transluminal Angioplasty Study (CAVATAS): a randomised trial. Lancet. 2001;357:1729-37.
17. Easton JD, Sherman DG. Stroke and mortality rate in carotid endarterectomy: 228 consecutive operations. Stroke. 1977;8:565-8.
18. Bennett KM, Scarborough JE, Shortell CK. Predictors of 30-day postoperative stroke or death after carotid endarterectomy using the 2012 carotid endarterectomy-targeted American College of Surgeons National Surgical Quality Improvement Program database. J Vasc Surg. 2015;61:103-11.
19. AlSheikh S, Aljabri B, Alanezi T, Al-Salman M, Aldossary MY, Almashat AH, et al. Outcomes of carotid endarterectomy: insights from a single-center retrospective cohort study. Saudi Med J. 2024;45:405-13.
20. Suphathamwit A, Leewatchararoongjaroen C, Rujirachun P, Poopong K, Leesakul A, Junyavoraluk A, et al. Incidence of postoperative, major, adverse cardiac events in patients undergoing carotid endarterectomy: a single-center, retrospective study. SAGE Open Med. 2022;10:20503121211070367.
21. Kühne M, Krisai P, Coslovsky M, Rodondi N, Müller A, Beer JH, et al. Silent brain infarcts impact on cognitive function in atrial fibrillation. Eur Heart J. 2022;43:2127-35.
22. Gensicke H, van der Worp HB, Nederkoorn PJ, Macdonald S, Gaines PA, van der Lugt A, et al. Ischemic brain lesions after

carotid artery stenting increase future cerebrovascular risk. J Am Coll Cardiol. 2015;65:521-9.

23. Bosiers M, de Donato G, Deloose K, Verbist J, Peeters P, Castriota F, et al. Does free cell area influence the outcome in carotid artery stenting? Eur J Vasc Endovasc Surg. 2007;33:135-41; discussion 142-3.
24. Schillinger M, Gschwendtner M, Reimers B, Trenkler J, Stockx L, Mair J, et al. Does carotid stent cell design matter? Stroke. 2008;39:905-9.
25. Faateh M, Dakour-Aridi H, Mathlouthi A, Locham S, Naazie I, Malas M. Comparison of open- and closed-cell stent design outcomes after carotid artery stenting in the vascular quality initiative. J Vasc Surg. 2021;73:1639-48.
26. Wang SX, Marcaccio CL, Patel PB, Giles KA, Soden PA, Schermerhorn ML, et al. Distal embolic protection use during transfemoral carotid artery stenting is associated with improved in-hospital outcomes. J Vasc Surg. 2023;77:1710-9.e6.
27. Mas JL, Chatellier G, Beyssen B; EVA-3S Investigators. Carotid angioplasty and stenting with and without cerebral protection: clinical alert from the endarterectomy versus angioplasty in patients with symptomatic severe carotid stenosis (EVA-3S) trial. Stroke. 2004;35:e18-20.
28. Montorsi P, Caputi L, Galli S, Ciceri E, Ballerini G, Agrifoglio M, et al. Microembolization during carotid artery stenting in patients with high-risk, lipid-rich plaque. A randomized trial of proximal versus distal cerebral protection. J Am Coll Cardiol. 2011;58:1656-63.
29. Stabile E, Sannino A, Schiattarella GG, Gargiulo G, Toscano E, Brevetti L, et al. Cerebral embolic lesions detected with diffusion-weighted magnetic resonance imaging following carotid artery stenting: a meta-analysis of 8 studies comparing filter cerebral protection and proximal balloon occlusion. JACC Cardiovasc Interv. 2014;7:1177-83.
30. Giri J, Parikh SA, Kennedy KF, Weinberg I, Donaldson C, Hawkins BM, et al. Proximal versus distal embolic protection for carotid artery stenting: a national cardiovascular data registry analysis. JACC Cardiovasc Interv. 2015;8:609-15.
31. Omran J, Mahmud E, White CJ, Aronow HD, Drachman DE, Gray W, et al. Proximal balloon occlusion versus distal filter protection in carotid artery stenting: a meta-analysis and review of the literature. Catheter Cardiovasc Interv. 2017;89:923-31.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):357-367
DOI: 10.4274/BMJ.galenos.2025.2024.4-10

Research

A Machine Learning Model to Predict HbA1c Values of Patients

Hastaların HbA1c Değerlerinin Tahminlenmesi için Bir Makine Öğrenmesi Modeli

Gökhan Silahtaroğlu[1], Özge Doğuç[1], Yavuz Furuncuoğlu[2]

[1]İstanbul Medipol University Faculty of Business and Management Sciences, Department of Management Information Systems, İstanbul, Türkiye

[2]Bahçeşehir University, Medical Park Göztepe Hospital, Department of Internal Medicine, İstanbul, Türkiye

ABSTRACT

Objective: To develop an automated method that accurately predicts patients' hemoglobin A1c (HbA1c) values.

Methods: Data from 118,265 unique patients were collected from 30 hospitals in İstanbul. Although the study began with a large dataset, extensive data cleaning reduced the final modeling dataset to 180 complete records. While this limits generalizability, it allows for a controlled evaluation of prediction models under realistic constraints of missing clinical data. For the analysis, the data was parted 70%-30% ratio as training and testing. A gradient boosting algorithm was used to train a machine learning model to predict HbA1c values.

Results: In this study, a machine learning model was developed to predict HbA1c using patients' readily obtainable vital data. For testing, the overall accuracy of the prediction is 80%, Cohen's kappa of 0.549, which is acceptable. Validation results are also promising, with an accuracy of 82%.

Conclusion: Although predicting HbA1c using different parameters can be advantageous, these predictions may be inaccurate and should be used and interpreted with caution in patients with anemia, polycythemia, or hemoglobinopathies. Our values, with specificity of 86.72%, sensitivity of 76.71%, and accuracy of 81.76%, can help eliminate this problem. In addition to satisfactory testing and validation results, the study not only explains the gradient-boosting machine learning model but also provides detailed information on cleaning noisy data and imputing missing health data.

Keywords: Data management, diabetes complications, internal medicine, machine learning

ÖZ

Amaç: Bu çalışmanın amacı hastaların hemoglobin A1c (HbA1c) değerlerini doğru bir şekilde tahmin edebilecek otomatik bir yöntem geliştirmektir.

Gereç ve Yöntem: İstanbul'daki 30 farklı hastaneden 118.265 benzersiz hasta verisi toplanmıştır. Başlangıçta büyük bir veri setiyle çalışılmış olsa da, kapsamlı veri temizleme işlemleri sonucunda modelleme için kullanılan nihai veri seti 180 tam doldurulmuş kayıttan oluşmuştur. Bu durum, sonuçların genellenebilirliğini sınırlasa da, eksik klinik verilerin oluşturduğu gerçekçi kısıtlamalar altında tahmin modellerinin kontrollü bir şekilde değerlendirilmesine olanak sağlamaktadır. Analiz için veriler %70-%30 oranında eğitim ve test olarak ayrılmıştır. Makineyi HbA1c değerlerini tahmin edecek şekilde eğitmek için gradient boost algoritması kullanıldı.

Bulgular: Bu çalışmada, hastaların kolayca elde edilebilen hayati verilerini kullanarak, HbA1c biyobelirteçlerini tahmin etmek için bir makine öğrenme modeli geliştirilmiştir. Test verileri için, tahminin genel doğruluğu, kabul edilebilir olan Cohen'in kappa katsayısı 0,549 değeriyle %80'dir. Doğrulama sonuçları ise %82 doğrulukla umut vericidir.

Sonuç: HbA1c'yi farklı parametreler kullanarak tahmin etmek bir avantaj olsa da anemi, polisitemi veya hemoglobinopatisi olan hastalarda kesin sonuçlar alınamadığından kullanılması ve yorumlanması yanlış olabilmektedir. %86,72-%76,71 özgüllük ve %81,76 duyarlılık ve doğruluk değerlerimiz, bu sorunun ortadan kaldırılmasına yardımcı olabileceği düşünülmektedir. Çalışma, gradient-boost makine öğrenmesi modelini açıklamanın yanı sıra yüksek doğruluk değerleri sağlamakta ve aynı zamanda sağlıkta gürültülü verinin nasıl temizleneceği ve eksik sağlık verilerinin nasıl yeniden oluşturulacağı hakkında ayrıntılı bilgi vermektedir.

Anahtar Kelimeler: Veri yönetimi, diyabet komplikasyonları, iç hastalıkları, makine öğrenimi

Address for Correspondence: Özge Doğuç, Assoc. Prof., İstanbul Medipol University Faculty of Business and Management Sciences, Department of Management Information Systems, İstanbul, Türkiye
E-mail: odoguc@medipol.edu.tr **ORCID ID:** orcid.org/0000-0002-5971-9218

Cite as: Silahtaroğlu G, Doğuç Ö, Furuncuoğlu Y. A machine learning model to predict HbA1c values of patients. Med J Bakirkoy. 2025;21(4):357-367

Received: 01.10.2024
Accepted: 13.06.2025
Publication Date: 18.12.2025

INTRODUCTION

Artificial intelligence (AI) and machine learning (ML) have rapidly become essential tools in healthcare for predicting disease outcomes, supporting diagnosis, and personalizing treatment plans. Among chronic conditions, diabetes mellitus stands out due to its prevalence and long-term complications. Accurate prediction and early intervention are critical to improving patient outcomes. One biomarker central to diabetes management is hemoglobin A1c (HbA1c), which reflects average blood glucose levels over a two- to three-month period and is a key metric for both diagnosis and monitoring of the disease (1).

While HbA1c is typically measured through laboratory tests, several real-world challenges hinder its timely availability (2), such as cost, limited access to laboratories in rural areas, and delays resulting from the required three-month measurement period. Moreover, in cases where patients have hemoglobin-related disorders (e.g., anemia, polycythemia, hemoglobinopathies), direct measurement may yield unreliable results. In such scenarios, a reliable prediction of HbA1c based on alternative, easily obtainable physiological parameters (e.g., age, glucose level, blood pressure) becomes clinically valuable, potentially enabling faster interventions and improved continuity of care.

Although various ML approaches have been applied to predict glycemic trends, most existing models fall short in predicting HbA1c. For instance, Plis et al. (3) used support vector regression and autoregressive integrated moving average models informed by glucose dynamics to predict short-term glucose levels, but these models rely heavily on patient-specific time-series glucose data, limiting their generalizability. Zou et al. (4) compared decision trees, random forests, and neural networks for diabetes prediction using physical exam data, but did not focus on HbA1c as a predictive target, nor did they address the implications of missing or noisy data—a common issue in real-world health datasets.

This study addresses these gaps by proposing a gradient boosting model specifically tailored to predict HbA1c levels using routinely collected patient data. Unlike prior approaches, the proposed model incorporates robust data-preprocessing strategies—including advanced outlier detection, synthetic data augmentation, and ML-based imputation of missing values—resulting in a cleaner, more usable dataset (5). While the study began with a large dataset, strict data quality protocols and completeness criteria narrowed the final sample used for modeling. Gradient boosting is chosen because of its strengths in handling imbalanced datasets, reducing bias, and offering higher accuracy with less overfitting compared with traditional ensemble methods such as random forests (6).

In summary, this research contributes to the field by demonstrating that HbA1c can be predicted with high accuracy from non-invasive, widely available clinical parameters, using a refined gradient-boosting model, offering a practical solution in settings where HbA1c cannot be directly measured or trusted.

METHODS

A two-way analysis was performed in this study. The main purpose of the study is to estimate the HbA1c value, which is effective for predicting the development of diabetes. However, due to missing data, this estimation model could not directly adopted, and the data deficiencies were addressed beforehand, as explained above.

In the second stage, the HbA1c value, the main purpose of the study, was estimated. For this purpose, the HbA1c value was converted to discrete data. At this stage, HbA1c values were categorized into clinically relevant risk groups in alignment with American Diabetes Association guidelines and were further stratified as follows: values <5.7% were labeled as low risk (normal), 5.7-6.4% as moderate (prediabetes), 6.5-7.9% as high (controlled diabetes), and ≥8.0% as very high (uncontrolled diabetes). These thresholds were chosen to reflect increasing risk of diabetes-related complications and to align with clinical treatment targets. A total of 180 patients with complete HbA1c data were included in the study. Ethical approval was obtained from the Non-Invasive Clinical Research Ethics Committee of İstanbul Medipol University (approval no: 685, date: 03.09.2020). As this was a retrospective study using anonymized data, no informed consent was required from the patients. The data were divided into 70% training and 30% test sets using stratified sampling before ML model training. Synthetic minority over-sampling technique was applied to the 70% data slice used for learning, increasing the number of records to 375, and testing was performed with the remaining 30% slice (7,8). No synthetic data generation was carried out for the testing data.

ML algorithms are highly configurable by their hyperparameters. These parameters substantially influence the behavior, complexity, speed, and other aspects of the algorithm, and their values must be selected carefully to achieve optimal performance (9). In the estimation of the HbA1c variable, the best result was obtained when the gradient boosting algorithm was customized with the following hyperparameters: tree depth=8, number of models=200, and learning rate=0.05. A learning rate of

0.1 is usually a good starting point for gradient boosting algorithms, but the optimal learning rate depends on the number of models. For this reason, the model achieved its best performance when the number of models was high and the learning rate was low. Also, for the k-nearest neighbor algorithm, the Euclidean distance metric was used, and k was set to 5. The optimum values for the support vector machine (SVM) algorithm were kernel type radial basis function and sigma=1.6.

Selection and Identification of Cases

Data

In this study, data from 118,265 unique patients were collected at 30 hospitals in İstanbul. This patient dataset was anonymized before they were included in the study. By stripping off the identifying features. The dataset contains 12 patient attributes (variables). These variables are continuous or categorical and include variables such as age, gender, height, weight, patient diabetes status, family history of diabetes, and average glucose level. All variables are given in Table 1.

Preparation of Data

The main objective of the study is to estimate HbA1c levels, an important parameter in individuals with diabetes. However, before reaching this stage, some data preprocessing needed to be carried out to prepare the data for analysis. It is well known that one problem with health data is that they may be incomplete, inaccurate, missing, or noisy. The performance of the AI models may be poor, or the models' estimation accuracy may be low due to noisy data. Data preprocessing is probably the most important stage of data analysis and data mining, and therefore of ML.

Firstly, it was spotted that there were missing data in the raw dataset. Since the number of missing mean glucose values in the dataset is high (i.e., postprandial blood sugar; 107,012 missing values), these were completely excluded from the dataset. After removal of records with missing data, 11,253 unique patient records remained in the relevant dataset. After that, outlier cleaning was performed. To detect the outliers for each variable, the first and third quartiles (Q1 and Q3) were computed. An observation was flagged as an outlier if it lies outside the range R=[Q1–k (IQR), Q3+k (IQR)] with IQR=Q3-Q1, where k=1.5, which corresponds to the smallest value in R (10,11). After this process, 6,754 records remained in the dataset.

Table 1. Variables used

Variable name
Gender
Age
Weight
Height
Pulse/heartbeat
Average blood pressure systolic
Average blood pressure diastolic
Average glucose
Has diabetes

Among these records, 6,686 mean fasting glucose values were missing and 68 were present. The aim is to fill records with missing mean fasting glucose valuesusing existing values. As a first step, missing mean fasting glucose values were predicted and imputed into the dataset. In the original dataset, the fasting glucose value was a continuous variable. To handle missing values for this variable, the fasting glucose value was first converted into discrete categories. Accordingly, fasting glucose values between 0 and 70 were labeled as hypoglycemia; values between 70 and 100 were labeled as normal; those between 100 and 125 were labeled as prediabetes; and values higher than 125 were labeled as diabetes. After this transformation, the dataset included 16 patients with diabetes, 25 with hypoglycemia, and 27 with prediabetes.

Because the aim was to fill in the missing mean fasting glucose records using the non-missing values, 68 existing values were synthetically reproduced using the k-nearest neighbor technique, increasing the number of non-missing values to 204. This was done to increase the capability of the AI ML model.

An AI model was built, using these 204 records, to replace missing fasting glucose values with predicted values. The model was trained using a gradient boosting algorithm to impute the missing values for 6,686 patients. However, not all the predictions made by the AI model were added to the dataset. Only those predicted with a confidence of 95% or higher were added to the dataset. Subsequently, 4,331 records met this criterion. Together with the original non-missing dataset, the total number of records reached 4,399 (4,331+68).

Nevertheless, the dataset included missing values for the target variable HbA1c. All records with missing HbA1c were removed from the dataset, and, finally, 180 clean, complete records were ready for analysis. Figure 1 shows the workflow diagram of the data cleaning and preprocessing steps, including (1) exclusion of records with missing glucose values, (2) outlier detection using the IQR method, (3) imputation of missing fasting glucose values using a ML model, and (4) removal of incomplete HbA1c records.

The final dataset used for training consisted of 180 fully populated records.

A Close Look at the Cleaned Data

In the final case, the data set includes 80 female and 100 male patients; 78 patients with prediabetes, 72 with hypoglycemia, and 30 with diabetes. Table 2 summarizes descriptive statistics for all numeric variables in the dataset.

Technical Information

In this study, the KNIME platform (version 4.4.2; Konstanz Information Miner, Germany) (12) was used to create learning models, develop predictive models from data, and visualize data. The learning algorithm used throughout the study is gradient boosting trees (13). Boosting is an ensemble learning method that sequentially generates base models. Boosting algorithms aim to construct a strong learner by sequentially combining weak learners generated at each iteration according to predefined rules. Gradient boosting, a boosting algorithm, is used to reduce model bias (6).

Gradient-based decision-tree algorithms are also weak learners. A weak learner is a ML model that performs slightly better than chance. In the case of gradient-boosting trees, the weak learners are shallow decision trees. Each new tree added to the ensemble model (i.e., the combination of all previous trees) minimizes the loss function associated with the ensemble (14). Decision trees rely on iteratively asking questions to split the data; however, they are prone to overfitting, and this is true for gradient-based trees as well. To reduce this risk, the gradient boosting algorithm employs a model that combines multiple decision trees.

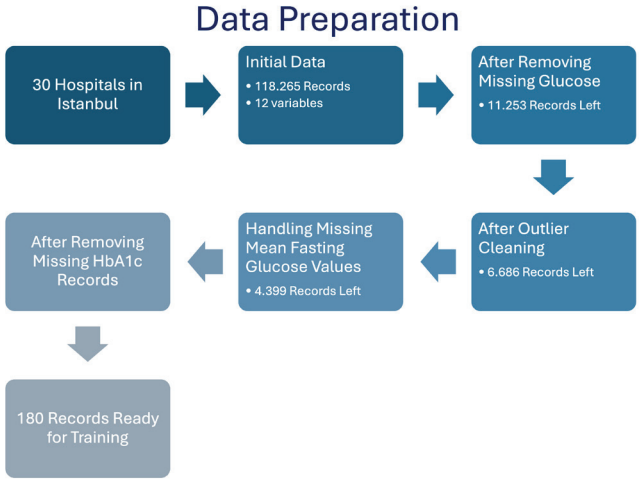


Figure 1. Data cleaning and preparation workflow
Hb1Ac: Hemoglobin A1c

The steps of the boosting method are listed below:

1. The weights of each sample are the same at the beginning,
2. It is used to learn basic learner 1 training examples,
3. When the learning is complete, the weight of the wrong samples is increased and the weight of the correct samples is decreased,
4. Basic learner 2 is used for learning,
5. Steps 2 and 4 are repeated to obtain M core-learners,
6. The results of M core learners are combined to produce the final learning outcome.

The weight of each base learner in the boosting method differs (15). Hyperparameters of the gradient boosting algorithm are set to tree depth=8, number of models=200, and learning rate=0.05. As a result, the average glucose hunger value was imputed from incomplete records using estimates with 95% confidence. Figure 2 depicts the overall model for training, testing, and validation.

Statistical Analysis

Testing

As mentioned above, 30% of the dataset, which consists of 55 records, was used for testing. As shown in Figure 3, the model achieved an accuracy of 80% and Cohen's kappa of 0.549 during testing. While accuracy alone can be misleading in imbalanced datasets, Cohen's kappa provides a more balanced measure of classification performance by accounting for agreement occurring by chance. A kappa value of 0.549 indicates moderate agreement, suggesting that the model produces reliable predictions beyond random chance but still has room for improvement.

Table 2. Sample data with minimum, maximum and mean point

Variables	Min	Max	Mean	Standard deviation
Age	16.00	94.00	56.64	16.53
Average blood pressure systolic	102.00	149.00	121.13	10.97
Average blood pressure diastolic	56.50	91.00	71.78	6.51
Weight	40.00	122.10	77.21	17.80
Height	144.00	190.00	165.85	9.21
Pulse/heartbeat	54.00	112.00	78.76	12.03
Average glucose	68.00	178.00	108.38	22.52

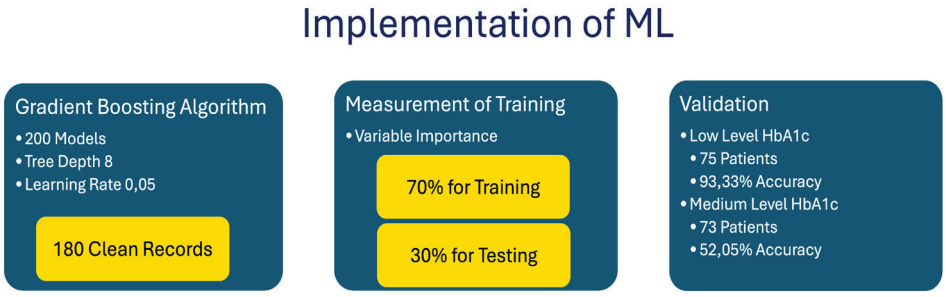


Figure 2. Implementation of gradient boosting algorithm
ML: Machine learning, HbA1c: Hemoglobin A1c

low (Actual)	0	34	2	0	94.44%
medium (Actual)	0	6	7	0	53.85%
veryhigh (Actual)	0	0	0	0	undefined
	100.00%	79.07%	77.78%	undefined	

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
80.00%	20.00%	0.549	44	11

	high (Predicted)	low (Predicted)	medium (Predicted)	veryhigh (Predict...	
high (Actual)	4	0	0	0	100.00%
low (Actual)	1	24	3	0	85.71%
medium (Actual)	2	4	11	0	64.71%
veryhigh (Actual)	0	0	0	1	100.00%
	57.14%	85.71%	78.57%	100.00%	

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
80.00%	20.00%	0.655	40	10

	high (Predicted)	low (Predicted)	medium (Predicted)	veryhigh (Predict...	
high (Actual)	2	2	0	0	50.00%
low (Actual)	0	28	0	0	100.00%
medium (Actual)	0	8	9	0	52.94%
veryhigh (Actual)	0	0	0	1	100.00%
	100.00%	73.68%	100.00%	100.00%	

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
80.00%	20.00%	0.608	40	10

Figure 3. Gradient boost, KNN and SVM evaluation of training with test data
KNN: K-nearest neighbors, SVM: Support vector machine

Examination of Figure 3 shows that the machine predicts a high level of HbA1c with 100% accuracy. On the other hand, there were no data to test for very high HbA1c values. Finally, the accuracies for the low and medium levels are 79.07% and 77.78%, respectively.

Table 3 shows a detailed evaluation of Gradient Boosting, k-nearest neighbors, and SVM algorithms. In the first table, all F-measures are acceptable and both sensitivity and specificity values are adequate for the gradient boost algorithm.

While the algorithm made the above predictions, it ranked the variables according to their importance. It should be kept in mind that this grading is relative and is performed by the gradient boosting algorithm itself. It means that the algorithm has obtained information from the features shown in Table 4. As can be seen, the patient's average glucose, age, height, and blood pressure are more important to the algorithm than other features.

RESULTS

Since the testing data were drawn from the training data, were preprocessed, and had some missing values imputed using predictions from other variables, the results could have been biased. Due to this possibility, data collection

Table 3. Detailed evaluation of training

Gradient boost								
HbA1c	**TP**	**FP**	**TN**	**FN**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
High	3	0	86	1	1.0	0.75	1.0	0.86
Low	33	12	31	14	0.73	0.70	0.72	0.72
Medium	26	14	38	12	0.65	0.68	0.73	0.67
Very high	1	1	88	0	0.5	1.0	0.99	0.67
KNN								
HbA1c	**TP**	**FP**	**TN**	**FN**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
High	4	3	43	0	0.57	1.00	0.93	0.73
Low	24	4	18	4	0.86	0.86	0.82	0.86
Medium	11	3	30	6	0.79	0.65	0.91	0.71
Very high	1	0	49	0	1.00	1.00	1.00	1.00
SVM								
HbA1c	**TP**	**FP**	**TN**	**FN**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
High	2	0	46	2	1.00	0.50	1.00	0.67
Low	28	10	12	0	0.74	1.00	0.55	0.85
Medium	9	0	33	8	1.00	0.53	1.00	0.69
Very high	1	0	49	0	1.00	1.00	1.00	1.00

HbA1c: Hemoglobin A1c, TP: True positive, FP: False positive, TN: True negative, FN: False negative, KNN: K-nearest neighbors, SVM: Support vector machine

Table 4. Variable importance gradient boosting

Variables	**Relative importance**	**Scaled importance**	**Percentage**
Average glucose	297.9	1.0	0.3
Age	215.1	0.7	0.2
Average blood pressure diastolic	127.4	0.4	0.1
Height	125.1	0.4	0.1
Average blood pressure systolic	117.4	0.4	0.1
Pulse/heartbeat	87.1	0.3	0.1
Weight	86.9	0.3	0.1
Gender (male)	9.2	0.0	0.0
Gender (famale)	0.0	0.0	0.0
Is diabetes	0.0	0.0	0.0
Diabetes in family	0.0	0.0	0.0

continued during the study, and these data were used to validate the learning. The test dataset used to evaluate the model's performance consisted of only 55 records, and the validation set included 73 medium and 75 low HbA1c values. No records representing the "high" or "very high" categories were present in the validation data. This limited sample size, particularly the absence of two of the four target classes during validation, reduces generalizability of the findings and undermines multi-class classification performance.

As it is seen in Table 5 and Figure 4 are examined it is seen that area under curve is 0.81 which close to 1 and sensitivity and specificity statistics are above 76% suggesting that machine predictions are accurate enough to be used. On the other hand, the area under the confidence curve of the machine appears to be approximately 50%. This value can be interpreted as: although the machine predicts with an accuracy of 81.76%, its confidence is low (Table 6, Figure 5). In this case, when the model is applied to real data, any prediction exceeding 55% may be sufficiently reliable.

DISCUSSION

HbA1c value is a criterion used in the diagnosis and treatment of diabetes. It is an important biomarker because, as the HbA1c level increases or decreases, the risk of diabetes-related complications changes accordingly. For this reason, it plays a critical role in assessing both the possibility of disease progression and the effectiveness of treatment. This value reflects the average of fasting and postprandial blood glucose levels over the preceding 2-3 months and is a robust indicator of long-term glycemic control. It is not only useful for monitoring diabetes but also predictive of cardiovascular complications.

While traditional glucose tests, such as fasting blood glucose, provide only a snapshot of a patient's current state, they do not reflect long-term trends. Continuous glucose monitoring addresses this gap partially, but remains costly and less accessible. In contrast, HbA1c provides a time-averaged measure; predicting it from easily collected variables may be useful in contexts where laboratory testing is unavailable or unreliable.

Study Limitations

One potential limitation of the study is the significant reduction in usable data due to missing values. While the initial dataset was large, our strict criteria for data quality and the need for complete records led to a smaller final sample size. This decision was intended to avoid bias and to ensure model integrity. However, it does mean that the findings should be interpreted as a proof-of-concept, and further research with larger, more complete datasets is warranted to validate the model's generalizability.

In our study, the gradient boosting model achieved a test accuracy of 80% and a Cohen's kappa of 0.549. This kappa score indicates moderate agreement between the predicted and actual HbA1c categories, suggesting the model performs better than random classification. In the validation phase, the model demonstrated 81.76% accuracy, with sensitivity and specificity values exceeding 76% for low and medium HbA1c categories. These metrics support the model's robustness and generalizability within the observed data range. However, due to the absence of high and very high HbA1c values in the validation dataset, further validation with a balanced sample is necessary to confirm performance across all classes.

Table 5. Evaluation of validation data for low and medium level HbA1c values

Gradient boost								
HbA1c	**TP**	**FP**	**TN**	**FN**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low	65	17	56	10	0.79	0.87	0.77	0.83
Medium	56	10	65	17	0.85	0.77	0.87	0.81
KNN								
HbA1c	**TP**	**FP**	**TN**	**FN**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low	10	1	4	1	0.91	0.91	0.80	0.91
Medium	3	1	11	1	0.75	0.75	0.92	0.75
SVM								
HbA1c	**TP**	**FP**	**TN**	**FN**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low	11	3	2	0	0.79	1	0.40	0.88
Medium	2	0	12	2	1	0.5	1	0.67

HbA1c: Hemoglobin A1c, TP: True positive, FP: False positive, TN: True negative, FN: False negative, KNN: K-nearest neighbors, SVM: Support vector machine

Confusion Matrix

	high (Predicted)	low (Predicted)	medium (Predicted)	veryhigh (Predict...	
high (Actual)	0	0	0	0	undefined
low (Actual)	0	65	10	0	86.67%
medium (Actual)	0	17	56	0	76.71%
veryhigh (Actual)	0	0	0	0	undefined
	undefined	79.27%	84.85%	undefined	

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
81.76%	18.24%	0.635	121	27

	high (Predicted)	low (Predicted)	medium (Predicted)	veryhigh (Predict...	
high (Actual)	1	0	0	0	100.00%
low (Actual)	0	10	1	0	90.91%
medium (Actual)	0	1	3	0	75.00%
veryhigh (Actual)	0	0	0	0	undefined
	100.00%	90.91%	75.00%	undefined	

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
87.50%	12.50%	0.729	14	2

	high (Predicted)	low (Predicted)	medium (Predicted)	veryhigh (Predict...	
high (Actual)	0	1	0	0	0.00%
low (Actual)	0	11	0	0	100.00%
medium (Actual)	0	2	2	0	50.00%
veryhigh (Actual)	0	0	0	0	undefined
	undefined	78.57%	100.00%	undefined	

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
81.25%	18.75%	0.489	13	3

Figure 4. Gradient boost, KNN and SVM validation results of medium level HbA1c patients
KNN: K-nearest neighbors, SVM: Support vector machine, HbA1c: Hemoglobin A1c

Another critical consideration in clinical ML is model interpretability. Gradient boosting models are often perceived as black boxes, which can undermine trust in their outputs. To address this, we conducted a variable importance analysis. Features such as average glucose, age, height, and blood pressure were identified as the most influential predictors of HbA1c. This aligns with clinical knowledge and increases the model's transparency. Future work should incorporate advanced interpretability techniques, such as SHapley Additive exPlanations or local interpretable model-agnostic explanations, to further explain individual predictions and enhance physicians' confidence in the system.

Table 6. Validation statistics

Gradient boost				
Results	**Low HbA1c validation**		**Medium HbA1c validation**	
Statistic	**Value**	**95% CI**	**Value**	**95% CI**
Sensitivity	86.67%	76.84% to 93.42%	76.71%	65.35% to 85.81%
Specificity	76.71%	65.35% to 85.81%	86.67%	76.84% to 93.42%
Positive likelihood ratio	3.72	2.43 to 5.70	5.75	3.19 to 10.39
Negative likelihood ratio	0.17	0.10 to 0.31	0.27	0.18 to 0.41
Disease prevalence	50.68%	42.34% to 58.98%	49.32%	41.02% to 57.66%
Positive predictive value	79.27%	71.41% to 85.41%	84.85%	75.62% to 91.00%
Negative predictive value	84.85%	75.62% to 91.00%	79.27%	71.41% to 85.41%
Area under curve (ROC)	0.817			
Accuracy	81.76%	74.58% to 87.62%	81.76%	74.58% to 87.62%
KNN				
Results	**Low HbA1c validation**		**Medium HbA1c validation**	
Statistic	**Value**	**95% CI**	**Value**	**95% CI**
Sensitivity	90.91%	58.72% to 99.77%	75.00%	19.41% to 99.37%
Specificity	80.00%	28.36% to 99.49%	91.67%	61.52% to 99.79%
Positive likelihood ratio	4.55	0.78 to 26.50	9	1.27 to 63.89
Negative likelihood ratio	0.11	0.02 to 0.77	0.27	0.05 to 1.50
Disease prevalence	68.75%	41.34% to 88.98%	25.00%	7.27% to 52.38%
Positive predictive value	90.91%	63.17% to 98.31%	75.00%	29.70% to 95.52%
Negative predictive value	80.00%	36.98% to 96.46%	91.67%	66.64% to 98.38%
Area under curve (ROC)	87.50%	61.65% to 98.45%	87.50%	61.65% to 98.45%
Accuracy	90.91%	58.72% to 99.77%	75.00%	19.41% to 99.37%
SVM				
Results	**Low HbA1c validation**		**Medium HbA1c validation**	
Statistic	**Value**	**95% CI**	**Value**	**95% CI**
Sensitivity	100.00%	71.51% to 100.00%	50.00%	6.76% to 93.24%
Specificity	40.00%	5.27% to 85.34%	100.00%	73.54% to 100.00%
Positive likelihood ratio	1.67	0.81 to 3.41		
Negative likelihood ratio	0		0.5	0.19 to 1.33
Disease prevalence	68.75%	41.34% to 88.98%	25.00%	7.27% to 52.38%
Positive predictive value	78.57%	64.19% to 88.24%	100.00%	
Negative predictive value	100.00%		85.71%	69.25% to 94.11%
Area under curve (ROC)	81.25%	54.35% to 95.95%	87.50%	61.65% to 98.45%
Accuracy	100.00%	71.51% to 100.00%	50.00%	6.76% to 93.24%

HbA1c: Hemoglobin A1c, KNN: K-nearest neighbors, SVM: Support vector machine, ROC: Receiver operating characteristic, CI: Confidence interval

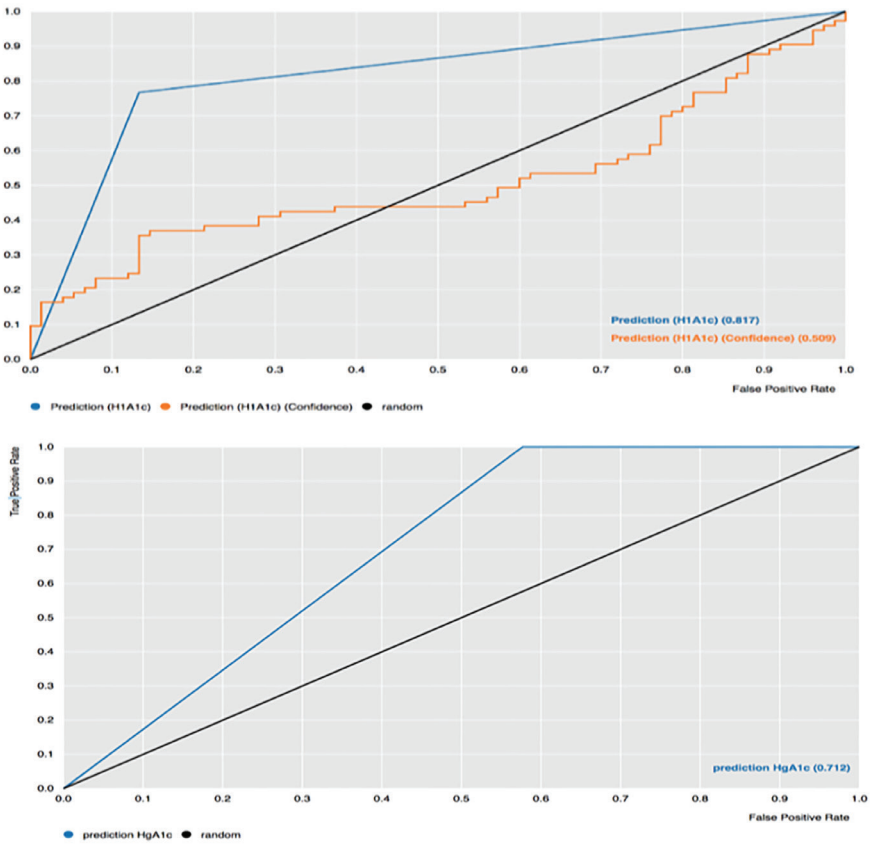


Figure 5. ROC of validation gradient boost and KNN
KNN: K-nearest neighbors, ROC: Receiver operating characteristic, HbA1c: Hemoglobin A1c

CONCLUSION

Although it is an advantage to predict HbA1c using different parameters, it can be wrong to use and interpret because accurate results cannot be obtained for the patients with anemia, polycythemia or hemoglobinopathy. Our values with 86.72-76.71% specificity and sensitivity and accuracy of 81.76% can help to eliminate this problem. Positive and negative predictivity values are also above acceptable levels. Based on these results, it is a useful estimation tool in situations where HbA1c cannot be measured, when testing is requested without requiring a 3-month waiting period, or when Hb-related diseases render HbA1c measurements unreliable.

ETHICS

Ethics Committee Approval: Ethical approval was obtained from the Non-Invasive Clinical Research Ethics Committee of İstanbul Medipol University (approval no: 685, date: 03.09.2020).

Informed Consent: As this was a retrospective study using anonymized data, no informed consent was required from the patients.

FOOTNOTES

Authorship Contributions

Surgical and Medical Practices: G.S., Y.F., Concept: G.S., Y.F., Design: Ö.D., Data Collection or Processing: Ö.D., Y.F., Analysis or Interpretation: G.S., Ö.D., Literature Search: Ö.D., Writing: G.S., Ö.D.

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

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

REFERENCES

1. Canbolat ZN, Silahtaroğlu G, Doğuç Ö, Yılmaztürk N. A machine learning approach to predict creatine kinase test results. Emerg Sci J. 2020;4:283-96.

2. Sarwar MA, Kamal N, Hamid W, Shah MA. Prediction of diabetes using machine learning algorithms in healthcare. 2018 24th International Conference on Automation and Computing (ICAC); Newcastle Upon Tyne, UK; 2018. p. 1-6.
3. Plis K, Bunescu R, Marling C, Shubrook J, Schwartz F. A machine learning approach to predicting blood glucose levels for diabetes management. Proceedings of the AAAI 14 Workshop on Modern Artificial Intelligence for Health Analytics; 2014; Toronto, Canada.
4. Zou L, Lam NSN, Cai H, Qiang Y. Mining twitter data for improved understanding of disaster resilience. Ann Am Assoc Geogr. 2018;108:1422-41.
5. Sherwani SI, Khan HA, Ekhzaimy A, Masood A, Sakharkar MK. Significance of HbA1c test in diagnosis and prognosis of diabetic patients. Biomark Insights. 2016;11:95-104.
6. Jeatrakul P, Wong KW, Fung CC. Classification of imbalanced data by combining the complementary neural network and SMOTE algorithm. In: Wong K W, Mendis B S U, Bouzerdoum A (eds). Neural Information Processing. Models and Applications. ICONIP 2010. Lecture Notes in Computer Science, vol 6444. Berlin, Heidelberg: Springer-Verlag; 2010:152-9.
7. Basgall MJ, Hasperué W, Naiouf M, Fernández A, Herrera F. An analysis of local and global solutions to address big data imbalanced classification: a case study with SMOTE preprocessing. In: Rucci E, Naiouf M, Chichizola F, De Giusti L, De Giusti A (eds). Cloud Computing and Big Data: Revised Selected Papers from JCC&BD 2019. Communications in Computer and Information Science, vol 1050. Cham: Springer; 2019. p. 75-85.
8. Bischl B, Binder M, Lang M, Pielok T, Richter J, Coors S, et al. Hyperparameter optimization: foundations, algorithms, best practices, and open challenges. WIREs Data Mining Knowl Discov. 2023;13:e1484.
9. Tukey JW. Some thoughts on clinical trials, especially problems of multiplicity. Science. 1977;198:679-84.
10. Zimek A, Filzmoser P. There and back again: outlier detection between statistical reasoning and data mining algorithms. WIREs Data Mining Knowl Discov. 2018;8:e1280.
11. Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, et al. KNIME – the Konstanz information miner: version 2.0 and beyond. SIGKDD Explorations Newsletter. 2009;11:26-31.
12. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat. 2001;29:1189-232.
13. Cao DS, Xu Q, Liang YZ, Zhang LX. The boosting: a new idea of building models. Chemometr Intell Lab Syst. 2010;100:1-11.
14. Callens A, Morichon D, Abadie S, Delpey M, Liquet B. Using random forest and gradient boosting trees to improve wave forecast at a specific location. Applied Ocean Research. 2020;104:102339.
15. Cheng J, Li G, Chen X. Research on travel time prediction model of freeway based on gradient boosting decision tree. IEEE Access. 2018;7:7466-80.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):368-373
DOI: 10.4274/BMJ.galenos.2025.2024.3-5

Research

The Prognostic Impact of Acute Kidney Injury in Patients with Moderate and Low-Risk Pulmonary Embolism

Orta ve Düşük Riskli Pulmoner Emboli Hastalarında Akut Böbrek Hasarının Prognostik Etkisi

Berrin Zinnet Eraslan[1], Meriç Oruç[2], Sevda Şener Cömert[1]

[1]University of Health Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital, Clinic of Chest Diseases, İstanbul, Türkiye

[2]University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Nephrology, İstanbul, Türkiye

ABSTRACT

Objective: This study aimed to evaluate the incidence of acute kidney injury (AKI) and its impact on prognosis in patients diagnosed with acute pulmonary embolism (PE).

Methods: Patients hospitalized with a diagnosis of acute PE between January 2018 and March 2023 were retrospectively reviewed. Demographic data; comorbidities; risk factors such as atrial fibrillation, malignancy, immobilization, history of surgery, fracture, and deep vein thrombosis (DVT); biochemical parameters; and 30-day mortality were recorded. AKI was diagnosed using the kidney disease: improving global outcomes criteria.

Results: AKI was identified in 11.8% of the 186 patients. Compared with those without AKI, patients with AKI were older (p=0.024) and had a higher prevalence of diabetes, chronic kidney disease, atrial fibrillation, and immobilization (p=0.035, 0.001, 0.027, and <0.001, respectively). In the AKI group, serum creatinine and troponin levels were higher, while estimated glomerular filtration rate and serum albumin levels were lower. In the mortality group, comorbidities such as diabetes mellitus, malignancy, immobilization, concomitant DVT, and a simplified pulmonary embolism severity index score ≥1 were observed more frequently. However, there was no significant difference in mortality between patients with and without AKI at admission or during hospitalization (p=0.079).

Conclusion: Among patients with acute PE, AKI present at admission does not independently affect 30-day mortality. However, the occurrence of AKI at admission or during hospitalization was associated with a non-significant trend toward increased mortality (p=0.079).

Keywords: Acute kidney injury, pulmonary embolism, mortality

ÖZ

Amaç: Bu çalışmanın amacı akut pulmoner emboli (PE) tanısı almış hastalarda akut böbrek hasarı (ABH) insidansını ve prognoz üzerindeki etkisini değerlendirmektir.

Gereç ve Yöntem: Ocak 2018 ile Mart 2023 arasında akut PE tanısıyla hastaneye yatırılan hastalar retrospektif olarak incelendi. Demografik veriler, eşlik eden hastalıklar, atriyal fibrilasyon, malignite, immobilizasyon, cerrahi öyküsü, kırık ve derin ven trombozu (DVT) gibi risk faktörleri, biyokimyasal parametreler ve 30 günlük mortalite durumları kaydedildi. ABH, böbrek hastalığı: küresel sonuçları iyileştirme kriterleri kullanılarak teşhis edildi.

Bulgular: Yüz seksen altı hastanın %11,8'inde ABH tespit edildi. ABH olmayan hastalarla karşılaştırıldığında, ABH olan hastalar daha yaşlıydı (p=0,024) ve diyabet, kronik böbrek hastalığı, atriyal fibrilasyon ve immobilizasyon prevalansı daha yüksekti (sırasıyla p=0,035, 0,001, 0,027 ve <0,001). ABH grubunda, serum kreatinin ve troponin seviyeleri daha yüksekken, tahmini glomerüler filtrasyon hızı ve serum albümin seviyeleri daha düşüktü. Mortalite görülen grupta diyabet, malignite, immobilizasyon, eş zamanlı DVT ve basitleştirilmiş pulmoner emboli şiddet indeksi ≥1 daha fazla görüldü. Ancak, ABH olan ve olmayan hastalar arasında hastaneye yatışta ve/veya yatış sırasında mortalite açısından anlamlı bir fark yoktu (p=0,079).

Sonuç: Akut PE hastalarında başvuru anında ABH'nin 30 günlük mortalite üzerine doğrudan bir etkisi görülmemiştir. Ancak başvuru anında ve/veya hastanede yatış sırasında ABH gelişimi, mortaliteyi artırma eğiliminde olmasına rağmen bu artış istatistiksel olarak anlamlı değildir (p=0,079).

Anahtar Kelimeler: Akut böbrek hasarı, pulmoner emboli, mortalite

Address for Correspondence: Berrin Zinnet Eraslan, MD, University of Health Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital, Clinic of Chest Diseases, İstanbul, Türkiye
E-mail: berzinbalta@yahoo.com **ORCID ID:** orcid.org/0000-0002-2425-6903

Cite as: Eraslan BZ, Oruç M, Şener Cömert S. The prognostic impact of acute kidney injury in patients with moderate and low-risk pulmonary embolism. Med J Bakirkoy. 2025;21(4):368-373

Received: 14.03.2024
Accepted: 13.06.2025
Publication Date: 18.12.2025

INTRODUCTION

Acute kidney injury (AKI), defined as a rapid and often reversible decline in kidney function, is characterized by decreased urine output or elevated serum creatinine (SCr) levels. It is a prevalent condition, particularly among hospitalized patients (1). The incidence of AKI in hospitalized patients is approximately 10-15%, increasing to as high as 50% in those admitted to intensive care units (2).

Pulmonary embolism (PE), a clinical condition caused by partial or complete obstruction of the pulmonary artery or its branches by a thrombus, is the third most common cardiovascular cause of death worldwide (3). The clinical spectrum of PE ranges from asymptomatic cases to fatal outcomes due to hemodynamic instability. In acute PE cases, risk assessment is the most crucial step for appropriate treatment planning, where "risk" refers to mortality. Stratifying patients diagnosed with acute PE into high, intermediate-, or low-risk categories for early mortality directly influences treatment options (anticoagulation or reperfusion) and prognosis (4). The main clinical manifestation of cardiovascular collapse due to massive PE and acute right ventricular failure is hypotension. Patients presenting with persistent hypotension and cardiogenic shock are considered high-risk because these conditions are directly associated with early mortality (5). In patients without detected hypotension or shock, risk assessment using the simplified pulmonary embolism severity index (sPESI) after diagnosis helps identify low- and intermediate-risk patients. Patients with PESI class I-II or sPESI <1 are deemed low-risk, while those with PESI class III-IV or sPESI ≥1 are considered medium risk (6).

AKI can develop in patients with acute PE due to renal congestion, hemodynamic instability, or both. A study by Murgier et al. (7) has shown that the presence of AKI in PE patients is associated with increased mortality across all risk groups (high, medium, and low). Additionally, when AKI status is added to the sPESI score, it enhances mortality prediction and can be useful in deciding whether low-risk patients can be treated at home (7).

In our study, we aimed to determine the frequency of AKI at admission and during hospitalization among medium- and low-risk patients with acute PE, and to investigate the impact of AKI on their prognosis.

METHODS

Patients admitted to the pulmonology department with a diagnosis of hemodynamically stable acute PE between January 2018 and March 2023 were retrospectively reviewed. Patients aged 18 years or older with a confirmed diagnosis of PE by computed tomography pulmonary angiography or ventilation/perfusion scintigraphy were included in the study. Patients with PE who were hemodynamically unstable were excluded from the study because they required monitoring in the intensive care unit. Additionally, patients with missing data, patients with an uncertain diagnosis of embolism, patients with duplicate records, and patients undergoing hemodialysis were excluded from the study (Table 1). Patients' demographic characteristics, comorbidities, risk factors, presence of deep vein thrombosis (DVT), sPESI scores, vital signs, creatinine, estimated glomerular filtration rate (eGFR), albumin, D-dimer, troponin values, and 30-day mortality status were recorded. The eGFR was calculated using the formula provided by the Chronic Kidney Disease Epidemiology Collaboration (8). The diagnosis of AKI was established according to the kidney disease: improving global outcomes (KDIGO) guidelines (9). The presence of AKI at hospital admission and its development during hospitalization were documented.

The sPESI is calculated based on six clinical parameters: age >80 years, history of cancer, chronic cardiopulmonary disease, heart rate ≥110 bpm, systolic blood pressure <100 mmHg, and oxygen saturation <90%. Patients with an sPESI score of 0 are considered low risk, whereas those with a score ≥1 are considered high risk.

In accordance with the Turkish Thoracic Society Pulmonary Embolism Consensus Report, patients were further categorized into moderate- and low-risk groups. Specifically, those with a PESI class I-II or sPESI <1 were categorized as low-risk, while patients with a PESI class III-IV or sPESI ≥1 were classified as moderate-risk. It is important to emphasize that these classification systems are exclusively applicable to hemodynamically stable patients (10).

Early-period mortality was defined as death due to any cause occurring within 30 days. Patients were divided into two groups: those with AKI at the time of admission and those without, Additionally, patients were categorized as survivors or non-survivors.

This study was conducted in accordance with the Declaration of Helsinki and was approved by our University of Health

Table 1. Patient selection and exclusion criteria

Patients initially assessed for eligibility	n=343
Patients with missing data excluded)	n=32
Uncertain diagnosis of embolism (excluded)	n=72
Patients with duplicate records (excluded)	n=46
Undergoing hemodialysis (excluded)	n=7
Total number of included patients	n=186

Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital Clinical Research Ethics Committee (approval no: 2022/514/238/20, date: 29.11.2022).

Statistical Analysis

Continuous data are presented as median and interquartile range (IQR). Categorical data are presented as percentages. For multiple-group comparisons of categorical variables, the chi-square test was used. Continuous variables were compared using the Mann-Whitney U test. Multivariate logistic regression analysis was performed to identify risk factors associated with mortality. Variables were selected for regression analysis based on the Mann-Whitney U test ($p<0.1$). All tests were performed using SPSS for Windows, version 17.0 (SPSS Inc., Chicago, IL, USA). P-values of less than 0.05 were considered statistically significant.

RESULTS

The study included 186 patients, of whom 82 (44.1%) were male. The average age of the patients was 63 years, and at the time of admission, 22 patients (11.8%) had AKI. The most common comorbidities were hypertension (44.1%), diabetes (23.1%), and coronary artery disease (16.7%), in that order. Active malignancy was present in 62 patients (33.3%). Atrial fibrillation was detected in 18 (9.7%) patients and DVT in 70 (39.5%) patients at the time of admission (Table 2).

The group with AKI at admission had a significantly higher mean age (p=0.024). Additionally, diabetes, chronic kidney disease, chronic atrial fibrillation, and immobilization were more frequent (p=0.035, p<0.001, p=0.027, p<0.001, respectively). Vital signs indicated that the group with AKI had lower oxygen saturation and pulse rate (p=0.025 and p=0.022, respectively) (Table 2).

In the group with AKI at admission, SCr and troponin levels were significantly higher, and eGFR and serum albumin levels were lower (p<0.000, p=0.002, p<0.000, p=0.003, respectively) (Table 3).

Among patients who died within 30 days, diabetes, active malignancy, immobilization, concurrent DVT, and a sPESI score ≥1 were significantly more common than in the group that did not die (p=0.037, p=0.007, p=0.039, p=0.048, and p=0.005, respectively). Mortality was higher in the group with AKI at admission or during hospitalization, although the difference was not statistically significant (p=0.079). Vital signs indicated that the group who died had a higher pulse

Table 2. Baseline characteristics of pulmonary thromboembolism patients according to AKI presence at admission

	All patients (n=186) median (IQR)	No AKI (n=164) median (IQR)	AKI (n=22) median (IQR)	p-value
Age (years)	63 (50-76)	62 (49-75)	73 (60.5-80.5)	**0.024**
Gender (male), n (%)	82 (44.1)	76 (46.3)	6 (27.3)	0.091
Comorbidities				
Hypertension, n (%)	82 (44.1)	70 (42.7)	12 (54.5)	0.293
Diabetes mellitus, n (%)	43 (23.1)	34 (20.7)	9 (40.9)	**0.035**
Coronary artery disease, n (%)	31 (16.7)	28 (17.1)	3 (13.6)	1.000
Chronic kidney disease, n (%)	21 (11.3)	14 (8.5)	7 (31.8)	**0.001**
Previous stroke, n (%)	15 (8.1)	12 (7.3)	3 (13.6)	0.271
Atrial fibrillation, n (%)	18 (9.7)	13 (7.9)	5 (22.7)	**0.027**
Malignancy, n (%)	62 (33.3)	51 (31.1)	11 (50)	0.077
Immobilization, n (%)	72 (38.7)	56 (34.1)	16 (72.7)	**0.000**
History of surgery n (%)	24 (12.9)	22 (13.4)	2 (9.1)	0.744
Fracture of lower limb n (%)	11 (5.9)	11 (6.7)	0 (0)	0.367
Coexistence with DVT, n (%)	70[a] (39.5)	61 (38.9)	9 (45)	0.596
Physical examination findings				
sBP, mmHg	115 (100-130)	119 (100-130)[b]	112 (100-122.5)	0.318
dBP, mmHg	70 (60-80)	70 (60-80)	70 (60-70)	0.189
Pulse, /min	85.5 (84-97)	85.5 (77.8-99)	77 (70-87.3)	**0.022**
SpO_2	94 (90-96)	95 (90-96)	90.5 (87.5-95)	**0.025**

Values are expressed as median (IQR), count (percentage).
[a]: Nine values were missing, [b]: Three values were missing, AKI: Acute kidney injury, DVT: Deep venous thrombosis, sBP: Systolic blood pressure, dBP: Diastolic Blood pressure, SpO_2: Pulse oximeter oxygen saturation, IQR: Interquartile range

rate (p=0.004), and biochemical parameters showed lower albumin and higher troponin T levels (p<0.000, p=0.009, respectively) (Table 4).

Logistic regression analysis of variables associated with mortality found that only concurrent malignancy was an independent risk factor [p=0.044; 95% confidence interval (CI), 1.04-25.76] (Table 5).

Table 3. Laboratory values of patients admitted with pulmonary embolism according to AKI presence at admission

	All patients (n=186) median (IQR)	No AKI (n=164) median (IQR)	AKI (n=22) median (IQR)	p-value
Serum creatinine, mg/dL	0.83 (0.66-1.04)	0.78 (0.65-0.97)	1.34 (1.07-1.67)	0.000
eGFR, mL/min/1.73 m^2	90.2 (69.8-105.4)	91.6 (77.6-106.4)	40.9 (32.6-61.1)	0.000
Albumin, g/dL	3.6 (3-4)	3.6 (3.1-4.03)	3.3 (2.8-3.5)	0.003
Troponin T, ng/L	19.5 (8-40)	18 (7-35.5)	40 (17-52)	0.002
D-dimer, µg/L	3875 (2722.5-7652.5)	3640 (2267.5-7487.5)	6295 (3200-13057.5)	0.108

IQR: Interquartile range, AKI: Acute kidney injury, eGFR: Estimated glomerular filtration rate

Table 4. Comparison of the demographic, clinical and laboratory characteristics on admission between patients who survived and who died

Variables	n	Patients who survived median (IQR)	n	Patients who died median (IQR)	p-value
Age (years)	**163**	62 (50-75)	**22**	73 (50.3-79)	0.179
Gender (male)	**163**	69 (42.3)	**22**	13 (59.1)	0.137
Comorbidities					
Hypertension	**163**	72 (44.2)	**22**	9 (40.9)	0.772
Diabetes mellitus	**163**	34 (20.9)	**22**	9 (40.9)	**0.037**
Coronary artery disease	**163**	28 (17.2)	**22**	3 (13.6)	1.000
Chronic kidney disease	**163**	17 (10.4)	**22**	4 (18.2)	0.285
Atrial fibrillation	**163**	15 (9.2)	**22**	3 (13.6)	0.454
Malignancy	**163**	49 (30.1)	**22**	13 (59.1)	**0.007**
Immobilization	**163**	59 (36.2)	**22**	13 (59.1)	**0.039**
Coexistence with DVT	**154**	57 (37.01)	**22**	13 (59.1)	**0.048**
Physical examination findings					
sBP, mmHg	**160**	114.5 (100-130)	**22**	115 (100-140)	0.665
dBP, mmHg	**160**	70 (60-80)	**22**	70 (60-71.3)	0.215
Pulse, min	**157**	82 (72-95)	**22**	94.5 (80-111.8)	**0.004**
SpO_2 <90%	**163**	34 (20.9)	**22**	4 (18.2)	1.000
Laboratory findings					
Serum creatinine, mg/dL	**163**	0.84 (0.68-1.04)	**22**	0.82 (0.63-1.21)	0.973
eGFR, mL/min/1.73 m^2	**163**	90.5 (69.5-105.2)	**22**	85.4 (70.9-108.4)	0.771
Albumin, g/dL	**159**	3.6 (3.2-4)	**20**	2.9 (2.6-3.2)	**0.000**
D-dimer, µg/L	**101**	3770 (2785-7420)	**13**	5680 (2055-12410)	0.533
Troponin T, ng/L	**119**	18 (7-38)	**16**	34 (27-49)	**0.009**
AKI at admission	**163**	19 (11.7)	**22**	3 (13.6)	0.730
AKI (admission and/or during hospitalization)	**137**	21 (15.3)	**19**	6 (31.6)	**0.079**
sPESI ≥1	**162**	87 (53.7)	**21**	18 (85.7)	**0.005**
PAP >25 mmHg	**122**	37 (30.3)	**17**	7 (41.2)	0.368

DVT: Deep venous thrombosis, sBP: Systolic blood pressure, dBP: Diastolic blood pressure, SpO_2: Pulse oximeter oxygen saturation, eGFR: Estimated glomerular filtration rate, AKI: Acute kidney injury, sPESI: Simplified pulmonary embolism severity index score, PAP: Pulmonary artery pressure, IQR: Interquartile range

Table 5. Multivariate logistic regression analyses of risk factors associated with mortality in patients admitted with pulmonary embolism

Variables	Multivariate			
	β	OR	95% CI	p-value
Diabetes mellitus	0.050	1.051	0.269-4.116	0.943
Malignancy	1.645	5.182	1.042-25.759	**0.044**
Immobilization	0.370	1.447	0.337-6.214	0.619
Coexistence with DVT	0.473	1.605	0.423-6.092	0.487
Pulse, min	0.036	1.036	0.994-1.080	0.091
Hypoalbuminemia (alb <3.5 g/dL)	1.863	6.443	0.699-59.403	0.100
Troponin T (ng/L)	-0.002	0.998	0.981-1.016	0.824
AKI (admission or during hospitalization)	1.015	2.761	0.567-13.433	0.208
sPESI ≥1	-0.114	0.892	0.114-6.973	0.914

alb: Albumin, DVT: Deep venous thrombosis, AKI: Acute kidney injury, sPESI: Simplified pulmonary embolism severity ındex, OR: Odds ratio, CI: Confidence interval

DISCUSSION

In our study, AKI present at hospital admission in patients with moderate- and low-risk PE did not affect mortality. The mortality rate was higher in patients who developed AKI at the time of initial admission and/or during hospitalization compared with those without AKI; however, this difference was not statistically significant.

AKI is a significant complication in patients with acute PE and has been associated with adverse clinical outcomes. In our study, AKI was identified in 11.8% of patients. In contrast, Murgier et al. (11) conducted a large-scale study involving 21,131 PE patients and reported that 29.5% developed AKI. Furthermore, their analysis demonstrated that AKI was an independent predictor of 30-day all-cause mortality in PE patients [odds ratio (OR)=1.25, 95% CI: 1.02-1.54] (11). Similarly, a recent meta-analysis by Xing et al. (12) reported that AKI is an independent predictor of poor prognosis in acute PE patients, doubling the overall mortality rate (pooled OR=2.75, 95% CI: 2.45-3.08, p<0.001). However, our study, which included 186 patients with moderate- and low-risk PE, found no significant association between AKI and 30-day mortality (p=0.079). This discrepancy may be attributed to differences in patient populations, PE severity, and sample size between the two studies. While Murgier et al. (11) included high-risk PE patients within a much larger cohort, our study focused solely on hemodynamically stable patients; this difference may explain the lack of statistical significance.

In another recent meta-analysis, Wang et al. (13) reported that renal insufficiency, particularly AKI and severe renal impairment (eGFR <30 mL/min/1.73 m^2), was significantly associated with an increased risk of early mortality in PE patients (pooled OR=2.69, 95% CI: 1.24-5.84, p<0.001). Unlike this meta-analysis, our study, which focused exclusively on AKI as defined by KDIGO criteria, found no significant association between AKI and 30-day mortality (p=0.079). This discrepancy may be attributed to differences in how renal impairment is defined among studies; for example, some studies in the meta-analysis included patients with severe chronic kidney disease (eGFR <30 mL/min/1.73 m^2).

In our study, serum albumin levels were significantly lower in patients with AKI than in those without AKI (p=0.003). Similarly, among patients who died within 30-days, serum albumin levels were significantly lower than in those who survived (p<0.000). Wiedermann et al. (14) published a meta-analysis demonstrating that hypoalbuminemia is an independent risk factor for both the development of AKI (OR=2.34, 95% CI: 1.74-3.14) and AKI-related mortality (OR=2.47, 95% CI: 1.51-4.05).

In the management of acute PE, the most commonly used risk stratification tool to guide treatment decisions and predict short-term mortality is the sPESI. In our study, the 30-day mortality rate among patients with an sPESI score ≥1 was significantly higher than among those with an sPESI score <1 (p=0.005).

The sPESI is a widely used and validated tool for risk stratification in acute PE patients. In our study, patients with an sPESI score ≥1 had significantly higher 30-day mortality than those with an sPESI score of 0 (p=0.005). While our study focused on evaluating the prognostic role of sPESI, a large-scale registry study by Murgier et al. (7), which included 30,532 PE patients from the RIETE database, further investigated the impact of AKI on sPESI-based risk stratification. Their findings demonstrated that AKI significantly increased 30-day mortality in all sPESI-defined risk groups: from 0.46% to 3% in low-risk patients, from 5.4% to 10% in intermediate-risk patients, and from 9.4% to 18% in high-risk patients (7). These results suggest that renal function may play a crucial role in PE prognosis. Although our study did not assess integrating AKI into sPESI scoring, our findings align with existing literature by confirming the prognostic power of sPESI.

Concurrent malignancy in patients hospitalized for acute PE has been associated with a 90% increase in all-cause mortality and a longer hospital stay (15). Malignancy is a well-established risk factor for poor prognosis in patients with PE. In our study, malignancy was identified as the only independent predictor of 30-day mortality in the multivariate logistic regression analysis (OR=5.18, 95% CI: 1.04-25.76, p=0.044). This finding highlights the significant impact of malignancy on mortality risk, even after adjustment for other well-known prognostic factors such as immobilization, DVT, troponin elevation, AKI, and sPESI ≥1; none of these factors were independently associated with mortality in our analysis. Similarly, a large-scale retrospective study by Shalaby et al. (15) found that cancer was associated with a significantly higher in-hospital mortality rate (11.8% vs. 6.6%, OR=1.79, p<0.0001).

Study Limitations

Our study has some limitations. It was a single-center retrospective study with a smaller patient group than reported in the literature. High-risk PE patient groups were not included in our study.

CONCLUSION

Our study demonstrated that in hemodynamically stable patients hospitalized with acute PE, the presence of AKI at admission was not an independent predictor of mortality. Although mortality rates were higher in patients with AKI at admission and/or during hospitalization, this difference did not reach statistical significance Further large-scale prospective studies are needed to better elucidate the prognostic role of AKI in patients with hemodynamically stable PE.

ETHICS

Ethics Committee Approval: This study was conducted in accordance with the Declaration of Helsinki and was approved by our University of Health Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital Clinical Research Ethics Committee (approval no: 2022/514/238/20, date: 29.11.2022).

Informed Consent: Retrospective study.

FOOTNOTES

Authorship Contributions

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

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

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

REFERENCES

1. Palevsky PM. Endpoints for clinical trials of acute kidney injury. Nephron. 2018;140:111-5.
2. Ronco C, Bellomo R, Kellum JA. Acute kidney injury. Lancet. 2019;394:1949-64.
3. 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;34:2363-71.
4. Konstantinides SV, Meyer G, Becattini C, Bueno H, Geersing GJ, Harjola VP, et al. The task force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC): pulmonary embolism: risk assessment and management. Eur Heart J 2020;41:543-603.
5. Laporte S, Mismetti P, Décousus H, Uresandi F, Otero R, Lobo JL, et al. Clinical predictors for fatal pulmonary embolism in 15,520 patients with venous thromboembolism: findings from the Registro Informatizado de la Enfermedad TromboEmbolica venosa (RIETE) Registry. Circulation. 2008;117:1711-6.
6. Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, et al. Derivation and validation of a prognostic model for pulmonary embolism. Am J Respir Crit Care Med 2005;172:1041-6.
7. Murgier M, Bertoletti L, Bikdeli B, Jimenez D, Trujillo-Santos J, Merah A, et al. Prognostic impact of acute kidney injury in patients with acute pulmonary embolism data from the RIETE registry. J Thromb Thrombolysis. 2022;54:58-66.
8. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150:604-12. Erratum in: Ann Intern Med. 2011;155:408.
9. Kellum JA. KDIGO clinical practice guideline for acute kidney injury: section 2: AKI definition. Kidney Int Suppl. 2012;2:19-36.
10. Turkish Thoracic Society. Consensus report on the diagnosis and treatment of pulmonary thromboembolism 2021 [Internet]. Turkish Thoracic Society; 2021. Available from: https://toraks.org.tr/site/community/library/gBkQrJRReWp5OPm4
11. Murgier M, Bertoletti L, Darmon M, Zeni F, Valle R, Del Toro J, et al. Frequency and prognostic impact of acute kidney injury in patients with acute pulmonary embolism. Data from the RIETE registry. Int J Cardiol. 2019;291:121-6.
12. Xing X, Liu J, Deng Y, Xu S, Wei L, Yang M, et al. Impact of renal function on the prognosis of acute pulmonary embolism patients: a systematic review and meta-analysis. Expert Rev Respir Med. 2022;16:91-8.
13. Wang D, Fan G, Liu X, Wu S, Zhai Z. Renal insufficiency and short-term outcomes of acute pulmonary embolism: a systematic review and meta-analysis. Thromb Haemost. 2020;120:1025-34.
14. Wiedermann CJ, Wiedermann W, Joannidis M. Hypoalbuminemia and acute kidney injury: a meta-analysis of observational clinical studies. Intensive Care Med. 2010;36:1657-65.
15. Shalaby K, Kahn A, Silver ES, Kim MJ, Balakumaran K, Kim AS. Outcomes of acute pulmonary embolism in hospitalized patients with cancer. BMC Pulm Med. 2022;22:11.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):374-379
DOI: 10.4274/BMJ.galenos.2025.2025.7-15

Research

Determinants of Surgeon and Center Selection in Recurrent Lumbar Disc Herniation: A Survey-Based Analysis of Patient Decision-Making

Nüks Lomber Disk Hernisinde Merkez ve Cerrah Değişimi: Hasta Tercihlerini Belirleyen Etkenler Üzerine Anket Temelli Bir Araştırma

İlhan Aydın[1], Ahmed Rasim Bayramoğlu[1], Berkay Kef[2], Melih Üçer[2], Orhun Mete Çevik[1], Uzay Erdoğan[1], Hasan Burak Gündüz[1], Erhan Emel[1]

[1]University of Health Sciences Türkiye, Bakırköy Prof. Dr. Mazhar Osman Mental Health and Neurological Diseases Training and Research Hospital, Clinic of Neurosurgery, İstanbul, Türkiye

[2]Biruni University Faculty of Medicine, Department of Neurosurgery, İstanbul, Türkiye

ABSTRACT

Objective: To investigate factors influencing patients' selection of surgeon and treatment center for revision surgery for recurrent lumbar disc herniation in Türkiye. Understanding the motivations behind changing providers after primary spinal surgery is crucial for enhancing continuity of care and informing health policy within mixed healthcare systems.

Methods: This cross-sectional observational study was conducted at a tertiary neurosurgery center in İstanbul. Forty patients who had previously undergone lumbar discectomy and subsequently required revision surgery were interviewed via structured telephone questionnaires. Patients were asked about their reasons for changing surgeons or institutions, their satisfaction with their initial surgery, and whether they had contacted their index surgeon following recurrence. Descriptive statistics were used to analyze response frequencies. Additionally, two-proportion z-tests were used to compare key findings against data from the landmark international literature.

Results: The most common reasons for changing providers were seeking a second opinion (45%), dissatisfaction with prior care (30%), and financial barriers (35%). All 14 patients who had initially undergone surgery in private hospitals opted for revision in public institutions, citing cost concerns. Despite 70% of patients reporting satisfaction with their index surgery, only 35% reconsulted their original surgeon after recurrence of symptoms. A minority (10%) deliberately avoided their initial provider despite having access. Furthermore, the proportion of patients who did not reconsult their index surgeon (65%) was significantly higher than that reported in benchmark United States of America studies (p=0.017).

Conclusion: Although patient satisfaction with initial surgery was generally high, provider change was frequently driven by economic pressures and informational needs, rather than geographic relocation or outright dissatisfaction. These findings underscore the importance of transparent preoperative counseling and cost awareness, particularly in mixed public-private systems like Türkiye's. Future research should evaluate the impact of provider continuity on revision outcomes and explore strategies to support equitable access to complex spinal care.

Keywords: Choice behavior, continuity of patient care, disc herniation, health services accessibility, patient satisfaction, recurrence, reoperation, second opinion, Türkiye

ÖZ

Amaç: Bu çalışma, Türkiye'de nüks lomber disk hernisi nedeniyle revizyon cerrahisi geçiren hastalarda cerrah ve sağlık merkezi tercihini etkileyen faktörleri araştırmayı amaçlamaktadır. Primer spinal cerrahi sonrası sağlık hizmeti sağlayıcısı değişiminin altında yatan motivasyonların anlaşılması, bakım sürekliliğinin güçlendirilmesi ve karma sağlık sistemlerinde sağlık politikalarının geliştirilmesi açısından önem taşımaktadır.

Gereç ve Yöntem: Bu kesitsel gözlemsel çalışma, İstanbul'daki üçüncü basamak bir beyin cerrahisi merkezinde yürütülmüştür. Daha önce lomber diskektomi uygulanmış ve sonrasında revizyon cerrahisine ihtiyaç duyan 40 hasta ile yapılandırılmış telefon anketleri aracılığıyla görüşülmüştür. Hastalara cerrah ya da kurum değiştirme nedenleri, ilk cerrahilerine ilişkin memnuniyet düzeyleri ve nüks sonrası bu cerrahla yeniden iletişime geçip geçmedikleri sorulmuştur. Yanıt frekansları tanımlayıcı istatistiklerle analiz edilmiş; ayrıca bazı temel bulgular, uluslararası literatürdeki referans verilerle iki oran z-testleri kullanılarak karşılaştırılmıştır.

Address for Correspondence: İlhan Aydın, MD, University of Health Sciences Türkiye, Bakırköy Prof. Dr. Mazhar Osman Mental Health and Neurological Diseases Training and Research Hospital, Clinic of Neurosurgery, İstanbul, Türkiye
E-mail: ilhanaydinmd@gmail.com **ORCID ID:** orcid.org/0000-0001-7681-8528

Cite as: Aydın İ, Bayramoğlu AR, Kef B, Üçer M, Çevik OM, Erdoğan U, et al. Determinants of surgeon and center selection in recurrent lumbar disc herniation: a survey-based analysis of patient decision-making. Med J Bakirkoy. 2025;21(4):374-379

Received: 29.07.2025
Accepted: 02.09.2025
Publication Date: 18.12.2025

ÖZ

Bulgular: Sağlık hizmeti sağlayıcısı değişikliğinin en sık gerekçeleri ikinci görüş alma isteği (%45), önceki bakımdan memnuniyetsizlik (%30) ve finansal engeller (%35) olarak belirlenmiştir. Başlangıçta özel hastanelerde ameliyat olan 14 hastanın tamamı, revizyon cerrahisini kamu hastanelerinde yaptırmayı tercih etmiş ve bu kararı maliyet gerekçesiyle aldıklarını belirtmiştir. Hastaların %70'i ilk cerrahilerinden memnun olduklarını ifade etmesine rağmen, yalnızca %35'i semptomların yinelenmesinden sonra ilk cerrahlarına yeniden başvurmuştur. Erişim imkanı olmasına rağmen %10'luk bir kesim, önceki cerrahlarını bilinçli olarak tercih etmemiştir. İlk cerrahına yeniden başvurmayan hastaların oranı (%65), benzer Amerika Birleşik Devletleri kaynaklı çalışmalarda bildirilen oranlara kıyasla anlamlı ölçüde yüksektir (p=0,017).

Sonuç: Hastaların ilk cerrahiye yönelik memnuniyet düzeyi genel olarak yüksek olmakla birlikte, sağlık hizmeti sağlayıcısı değişimi çoğunlukla ekonomik nedenler ve bilgi edinme ihtiyacı ile ilişkilidir; coğrafi nedenler veya doğrudan memnuniyetsizlik daha az etkili görünmektedir. Bu bulgular, kamu-özel dengesine dayalı sağlık sistemlerinde, özellikle maliyet konusunda şeffaf bilgilendirme ve etkili preoperatif danışmanlığın önemine işaret etmektedir. Gelecek çalışmalar, sağlık hizmeti sağlayıcısı sürekliliğinin revizyon cerrahisi sonuçlarına etkisini değerlendirmeli ve kompleks spinal bakım hizmetlerine adil erişimi destekleyici stratejiler geliştirmelidir.

Anahtar Kelimeler: Tercih davranışı, hasta bakımının sürekliliği, disk hernisi, sağlık hizmetlerine erişim, hasta memnuniyeti, nüks, yeniden cerrahi, ikinci görüş, Türkiye

INTRODUCTION

Lumbar disc herniation is among the most frequently encountered degenerative spinal disorders, and surgical discectomy remains a well-established treatment for patients with persistent radiculopathy unresponsive to conservative measures. Despite the generally favorable outcomes of lumbar discectomy, recurrence rates have been reported between 5% and 15%, often requiring reoperation. In such cases, patients face complex decisions about whether to return to their original surgeon or to seek care from a different provider or institution (1,2).

Patient choice in recurrent spinal surgery is multifactorial and may be influenced by prior satisfaction with the initial surgery, trust in the surgeon, accessibility, institutional reputation, and even social or geographic mobility. In the setting of recurrent lumbar disc herniation, where surgical outcomes are less predictable and patient expectations are shaped by previous experiences, these decisions gain particular importance. Understanding the determinants of such preferences can offer valuable insights into the dynamics of surgeon-patient relationships and the broader delivery of spinal care (3-5).

While patient attrition during long-term follow-up is often assumed to reflect dissatisfaction with prior care, emerging evidence suggests that logistical and systemic factors—such as relocation, insurance compatibility, or referral patterns—may play a more prominent role. However, there is a paucity of data on these preferences in the context of revision lumbar surgery, particularly in countries with centralized health systems and limited availability of multidisciplinary spine units (6,7).

The objective of this study is to investigate the factors influencing selection of surgeon and center among patients undergoing surgery for recurrent lumbar disc herniation. Using a structured, telephone-administered questionnaire, we aim to assess the relative impact of satisfaction, communication, accessibility, and institutional variables on patient decision-making. Through this inquiry, we seek to contribute to a better understanding of patient behavior in the context of revision spine surgery and to inform strategies for optimizing continuity of care.

METHODS

Study Design and Setting

This observational, cross-sectional study was conducted at the Department of Neurosurgery of a tertiary care referral center in İstanbul, Türkiye. The study was approved by the University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital Non-Interventional Scientific Research Ethics Committee (approval no: 2025-10-03, date: 21.05.2025), and all procedures were carried out in accordance with the Declaration of Helsinki and national regulations on human subject research.

Patient Selection

Patients who underwent lumbar spine surgery between September 1, 2015, and September 1, 2024, and were subsequently diagnosed with recurrent lumbar disc herniation were retrospectively identified from hospital records. Eligibility criteria included age ≥18 years, a prior lumbar discectomy, and either reoperation for recurrence or consultation for recurrent symptoms. Patients were contacted via telephone, and those who provided verbal informed consent were included in the study. Exclusion criteria were: incomplete medical history, inability to establish communication, refusal to participate, or severe neurological deficits impairing communication.

Data Collection

A structured questionnaire was administered via telephone interviews between May and June 2025. The questionnaire included items regarding demographic data, initial and

revision surgery details, satisfaction with the index surgeon, reasons for seeking a different surgeon or institution, communication history with the initial provider, and preferences related to healthcare institution type (public, university, private).

The questionnaire was developed by adapting items from instruments used in similar studies (6-8). The final draft was then reviewed by two independent spine surgeons for content validity, clarity, and relevance.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics version 22.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics, including frequencies (n) and percentages (%), were used to summarize all patient-related data and questionnaire responses. To compare key proportions in our cohort with those reported in previous studies, a two-proportion z-test was used (7). A p-value of <0.05 was considered statistically significant.

RESULTS

Forty patients who underwent revision surgery for recurrent lumbar disc herniation completed the telephone-based questionnaire. Since all patients had previously undergone lumbar discectomy, spinal level and surgical type were not included as variables. Baseline characteristics of the index surgeons and institutions are summarized in Table 1.

Patients were asked about their reasons for not consulting the same surgeon again. The most commonly reported factors were seeking a second opinion (45.0%) and dissatisfaction with prior care (30.0%). In addition, 14 patients (35.0%) reported financial constraints, citing high costs in private hospitals as the reason they chose to undergo surgery in a public hospital. Multiple responses were allowed (Table 2).

Contact patterns and satisfaction outcomes are summarized in Table 3. Only 14 patients (35.0%) discussed their new symptoms with their original surgeon. Of the 26 who did not, four (10.0% of the total sample) had the opportunity to return but consciously chose not to return.

Table 1. Index-surgery characteristics (n=40)

Variable	Category	n	%
Specialty of index surgeon	Neurosurgery	35	87.5
	Orthopaedics	5	12.5
Institution where surgery was performed	Private hospital	20	50.0
	Tertiary training/research hospital	12	30.0
	Secondary-level public hospital	6	15.0
	University hospital	2	5.0

Despite half of the cohort initially received surgery in private hospitals, financial concerns emerged as a major determinant of subsequent surgeon preference. While dissatisfaction and the desire for a second opinion were common, 70% of patients reported being satisfied with their original surgical care and expressed a willingness to undergo the same operation again (Table 3). These findings suggest that changes of surgeon or treatment center in recurrent disc surgery are not solely due to negative experiences, but are shaped by multifactorial considerations, including cost, access, and evolving trust dynamics.

DISCUSSION

This study investigated why patients undergoing revision lumbar discectomy did not return to their original surgeons for follow-up care or reoperation. Although the majority expressed satisfaction with their initial surgical outcomes, more than half of the cohort sought consultation with a different surgeon for the recurrent episode. The most commonly cited reasons were dissatisfaction with prior care, desire for a second opinion, and financial barriers associated with reoperation in the private sector. Notably, even among patients who had access to their original surgeons, some deliberately changed providers, indicating that practical access alone does not ensure continuity of care. These findings highlight a critical gap between patient experience and provider continuity in the context of recurrent spinal

Table 2. Patient-reported reasons for not returning to the index surgeon (n=40)

Reason	n	%
Seeking a second opinion	18	45.0
Dissatisfaction with previous care	12	30.0
Surgeon no longer practiced in the area	5	12.5
Patient had relocated	5	12.5
Told a more complex operation was required	4	10.0
Other*	14	35.0

*: All 14 patients underwent surgery in private hospitals and were subsequently transferred to public care for revision surgery because of high out-of-pocket costs

Table 3. Follow-up behavior and patient satisfaction (n=40)

Outcome	n	%
Contacted original surgeon regarding new symptoms	14	35.0%
Did not contact original surgeon	26	65.0%
Of whom actively chose not to return despite access	4	10.0%
Satisfied with prior surgical care	28	70.0%
Would undergo the same procedure again for same outcome	28	70.0%

surgery, emphasizing the multifactorial nature of surgical loyalty and decision-making.

Patients' decisions to change surgeons despite access to their original providers are consistent with prior literature highlighting the complexity of follow-up behavior in spine care. In the landmark study by Daffner et al. (7), approximately 46% of patients did not contact their original surgeon upon experiencing new symptoms, and only 4% consciously chose not to return despite having the opportunity. In contrast, our study found both a significantly higher rate of patients who did not reconsult their index surgeon (65% vs. 46.4%; z=2.39, p=0.017) and a greater proportion of patients who actively avoided follow-up (10% vs. 4%). This suggests a higher rate of intentional provider change in our setting, possibly influenced by contextual factors such as the healthcare structure or cultural expectations. Moreover, while Daffner et al. (7) reported that 71% of patients were satisfied with their previous care, we found a nearly identical satisfaction rate (70%), indicating that switching surgeons is not always driven by dissatisfaction alone. Rather, as seen in both cohorts, second-opinion seeking emerged as a major driver. Although the rate of second-opinion seeking was notably higher in our cohort, the difference was not statistically significant compared with Daffner's study (45.0% vs. 31.9%; z=1.37, p=0.17) (7). This finding aligns with evidence from a recent scoping review that second opinions in spine surgery are both frequent and often discordant with initial recommendations. Such discrepancies can erode the trust in the original treatment plan and influence patient loyalty (7,8).

The reasons for this discordance are multifactorial. For example, in one study, 61.3% of patients who were initially recommended surgery for lumbar disc herniation received a different treatment plan upon seeking a second opinion, and 75% of these were advised to pursue non-surgical options (8). This highlights not only the variability in treatment approaches among spine surgeons but also the potential for subjective interpretation of patient expectations and clinical presentations. In particular, when a patient initially advised to undergo surgery is later recommended for conservative management, it can lead to confusion and diminished trust in the healthcare system. These conflicting recommendations significantly shape patient preferences: some patients opt for a more invasive approach, while others choose a more cautious path to avoid potential complications. Therefore, second opinions in spine surgery should be understood as influencing not only treatment decisions but also patient satisfaction, outcome expectations, and adherence to long-term follow-up.

Beyond individual satisfaction, patients' decisions were strongly influenced by structural factors related to the healthcare system. In our cohort, 35% of patients reported switching providers, primarily due to the high cost of reoperation in the private sector, despite having undergone their initial surgery in the private sector. There was a strong association between undergoing the index surgery in a private hospital and citing "financial constraints" as the primary reason for changing the surgeon ($p<0.001$). This behavior reflects a broader tendency in Türkiye's hybrid health system, where publicly insured patients can access both public and contracted private hospitals, yet may face substantial out-of-pocket expenses for complex or repeat procedures. The Health Transformation Program implemented in Türkiye over the past two decades has increased patient mobility and choice, allowing individuals to change providers with fewer bureaucratic barriers. However, this flexibility also leads to fragmentation of follow-up care, especially when provider continuity is deprioritized in favor of logistical convenience or financial necessity. Previous research suggests that although patient choice is formally guaranteed, true informed decision-making is limited by disparities in information availability and the uneven distribution of specialized services . In our setting, patients who initially received care in the private sector often sought revision in public hospitals due to affordability and perceived comprehensiveness of state-funded care. This aligns with studies showing that Turkish patients often rely on informal sources such as social networks and physician reputation when navigating between sectors. Thus, accessibility in Türkiye is not merely a matter of geographic proximity, but also of systemic financial pathways and perceptions of institutional trustworthiness (9).

Notably, although 70% of patients in our study reported satisfaction with their initial surgery and indicated they would undergo the same procedure again under similar circumstances, a substantial proportion nevertheless chose not to return to their original surgeon. This apparent paradox highlights the nuanced interplay between technical satisfaction and relational trust. As previous reviews have emphasized, fulfillment of patient expectations is a more reliable predictor of satisfaction than objective surgical outcomes alone. Even when clinical results are acceptable, dissatisfaction may arise from unmet expectations, perceived lack of empathy, or poor communication. In fact, one systematic review concluded that patients who experienced a mismatch between their expectations and outcomes were more likely to report dissatisfaction even if their symptoms improved. Changing surgeons may not necessarily reflect a negative surgical result but rather a

rupture in the therapeutic alliance. Our findings support this view, as some patients expressed overall satisfaction with the surgical outcome yet described feeling "ignored" or "unimportant" during postoperative follow-up. This aligns with studies showing that interpersonal aspects of care—such as perceived respect, attentiveness, and clarity—strongly influence long-term trust and provider loyalty. In surgical disciplines like spine surgery, where treatment decisions are preference-sensitive and recovery is often prolonged, the strength of the patient-surgeon relationship may be as critical as technical proficiency. Failure to maintain open, trust-based communication may lead even clinically successful patients to seek future care elsewhere (5,10-15).

This study has several strengths. It addresses an understudied area in spinal surgery—patient preferences and behavior in the setting of revision lumbar discectomy—through real-world data obtained from direct patient contact. Unlike registry-based or chart-review studies, our structured telephone interviews enabled us to capture patient perspectives, motivations, and values that may not be documented in routine clinical records. Furthermore, the homogeneity of the surgical indication (recurrent lumbar disc herniation) lends specificity to the findings, reducing confounding by diagnosis type. The inclusion of patients from both private and public healthcare settings enhances the generalizability of our findings across the Turkish health system.

Study Limitations

Nonetheless, several limitations must be acknowledged. First, the relatively small sample size (n=40) reduces statistical power and may limit detection of subtle associations between demographic variables and decision patterns. Second, the study design was retrospective and survey-based, making it susceptible to recall bias, especially regarding patients' perceptions of prior care and interactions with surgeons. Third, because all data were self-reported, social desirability bias may have influenced responses some participants may have understated dissatisfaction or overemphasized rational decision-making. Additionally, we did not formally assess patient literacy, psychological status, or cultural background, all of which are known to affect healthcare preferences and trust. Lastly, the results are drawn from a single tertiary center, which may limit external validity, particularly for rural or lower-resource settings.

Despite these limitations, our findings offer valuable insights into patient behavior following recurrent disc herniation surgery and suggest actionable strategies to improve continuity of care and patient-centered decision-making.

In summary, our findings underscore that surgeon and center preferences following recurrent lumbar disc herniation are shaped not solely by clinical outcomes but by a complex interplay of patient expectations, communication quality, system-level accessibility, and cultural norms. Even in the presence of postoperative satisfaction, lapses in relational trust or logistical barriers can lead patients to seek care elsewhere. Addressing these multifactorial influences through improved expectation management, strengthened follow-up protocols, and a more transparent, integrated healthcare system may enhance continuity of care and reinforce long-term patient-surgeon relationships in spinal practice.

CONCLUSION

This study highlights the multifactorial nature of surgeon and center preferences among patients undergoing revision surgery for recurrent lumbar disc herniation in Türkiye. Despite a high satisfaction rate with their initial operation, a substantial proportion of patients opted to change providers due to seeking second opinions, a perceived need for specialized expertise, and financial constraints, particularly among those initially treated in private hospitals.

These findings suggest that patient satisfaction alone is not a reliable predictor of continuity with the index surgeon, especially in the context of recurrent disease and system-level cost barriers. Financial accessibility and transparent communication about recurrence risks should be emphasized during preoperative counseling. Additionally, the limitations of the private sector in providing long-term continuity of care for complex cases warrant further investigation within Türkiye's healthcare system.

Future research should incorporate clinical outcomes and economic evaluations to better understand how changes in provider affect revision surgery outcomes and to support more equitable spine care planning across public and private sectors.

ETHICS

Ethics Committee Approval: The study was approved by the University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital Non-Interventional Scientific Research Ethics Committee (approval no: 2025-10-03, date: 21.05.2025).

Informed Consent: Patients were contacted via telephone, and those who provided verbal informed consent were included in the study.

FOOTNOTES

Authorship Contributions

Surgical and Medical Practices: İ.A., M.Ü., O.M.Ç., U.E., H.B.G., E.E., Concept: İ.A., B.K., A.R.B., U.E., Design: İ.A., B.K., A.R.B., H.B.G., Data Collection or Processing: İ.A., M.Ü., O.M.Ç., U.E., H.B.G., E.E., Analysis or Interpretation: İ.A., A.R.B., M.Ü., E.E., Literature Search: İ.A., B.K., Writing: İ.A., A.R.B., B.K., M.Ü., O.M.Ç., U.E., H.B.G., E.E.

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

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

REFERENCES

1. Parai C, Hägg O, Willers C, Lind B, Brisby H. Characteristics and predicted outcome of patients lost to follow-up after degenerative lumbar spine surgery. Eur Spine J. 2020;29:3063-73.
2. Cômes PC, Gavotto A, Zouakia Z, Lonjon G, Amelot A, Edgard-Rosa G, et al. Repeat discectomy or instrumented surgery for recurrent lumbar disk herniation: an overview of French spine surgeons' practice. Global Spine J. 2025;15:1533-43.
3. Andersen SB, Birkelund R, Andersen MØ, Carreon LY, Coulter A, Steffensen KD. Factors affecting patient decision-making on surgery for lumbar disc herniation. Spine (Phila Pa 1976). 2019;44:143-9.
4. Alsalamah RA, Aljohani EA, Aljasser R, Alsaud JS, Alsherbi R, Albalawi IAJ, et al. Patient perceptions and preferences when choosing a surgeon: a cross-sectional study, Qassim Region, Saudi Arabia. Cureus. 2023;15:e39547.
5. Menendez JY, Omar NB, Chagoya G, Tabibian BE, Elsayed GA, Walters BC, et al. Patient satisfaction in spine surgery: a systematic review of the literature. Asian Spine J. 2019;13:1047-57.
6. Levy HA, Karamian BA, Pezzulo J, Canseco JA, Sherman MB, Kurd MF, et al. Do patient outcomes predict loss to long-term follow-up after spine surgery? World Neurosurg. 2023;170:e301-12.
7. Daffner SD, Hilibrand AS, Riew KD. Why are spine surgery patients lost to follow-up? Global Spine J. 2013;3:15-20.
8. Gattas S, Fote GM, Brown NJ, Lien BV, Choi EH, Chan AY, et al. Second opinion in spine surgery: a scoping review. Surg Neurol Int. 2021;12:436.
9. Yıldırım HH, Hughes D, Yıldırım T. Markets, modernisation and national interest: three faces of patient choice policy in Turkey. Health Soc Care Community. 2011;19:168-77.
10. Rampersaud YR, Canizares M, Perruccio AV, Abraham E, Bailey CS, Christie SD, et al. Fulfillment of patient expectations after spine surgery is critical to patient satisfaction: a cohort study of spine surgery patients. Neurosurgery. 2022;91:173-81.
11. Witiw CD, Mansouri A, Mathieu F, Nassiri F, Badhiwala JH, Fessler RG. Exploring the expectation-actuality discrepancy: a systematic review of the impact of preoperative expectations on satisfaction and patient reported outcomes in spinal surgery. Neurosurg Rev. 2018;41:19-30.
12. Godil SS, Parker SL, Zuckerman SL, Mendenhall SK, Devin CJ, Asher AL, et al. Determining the quality and effectiveness of surgical spine care: patient satisfaction is not a valid proxy. Spine J. 2013;13:1006-12.
13. Abtahi AM, Brodke DS, Lawrence BD, Zhang C, Spiker WR. Association between patient-reported measures of psychological distress and patient satisfaction scores after spine surgery. J Bone Joint Surg Am. 2015;97:824-8. Erratum in: J Bone Joint Surg Am. 2015;97:e54.
14. Rabah NM, Levin JM, Winkelman RD, Mroz TE, Steinmetz MP. The association between physicians' communication and patient-reported outcomes in spine surgery. Spine (Phila Pa 1976). 2020;45:1073-80.
15. Adogwa O, Carr K, Fatemi P, Verla T, Gazcon G, Gottfried O, et al. Psychosocial factors and surgical outcomes: are elderly depressed patients less satisfied with surgery? Spine (Phila Pa 1976). 2014;39:1701-6.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):380-385
DOI: 10.4274/BMJ.galenos.2025.2025.7-3

Research

Nutritional and Inflammatory Markers as Predictors of Major Amputation in Patients with Diabetic Foot Infections

Diyabetik Ayak Enfeksiyonu Olan Hastalarda Majör Amputasyon Öngörücüsü Nutrisyonel ve Enflamatuvar Belirteçler

Ayşegül İnci Sezen[1], Yusuf Emre Özdemir[1], Deniz Borcak[1], Mine Karzaoğlu[1], Kübranur Bali Yurdakul[1], Dilhan Günay[1], Serkan İskender[2], Ali Bayram Kasım[3], Kadriye Kart Yaşar[1]

[1]University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Infectious Diseases and Clinical Microbiology, İstanbul, Türkiye

[2]University of Health Sciences Türkiye, İzmir City Hospital, Clinic of General Surgery, İstanbul, Türkiye

[3]University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Center for Diabetic Foot and Chronic Wounds, Department of Family Medicine, İstanbul, Türkiye

ABSTRACT

Objective: Diabetic foot infections (DFIs) are serious complications of diabetes mellitus and a major cause of lower extremity amputations. Identifying predictive factors for major amputation can guide treatment and improve patient outcomes.

Methods: In this retrospective cross-sectional study, 77 patients with DFIs who were treated between December 2022 and July 2024 were evaluated. Clinical and laboratory data were collected, including inflammatory markers [C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), leukocyte count], hematologic parameters, and nutritional indices [albumin, hemoglobin, prognostic nutritional index (PNI)]. Amputations were classified as major or minor based on anatomical level. Statistical analyses were conducted to identify predictors of major amputation.

Results: Major amputations were performed in 25.9% of patients. Compared with patients without major amputation, those who underwent major amputation had significantly higher leukocyte, neutrophil, and platelet counts, CRP, and ESR levels and lower hemoglobin, albumin, and PNI scores (all $p<0.05$). No significant differences were observed in neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, or systemic immune-inflammation index values between the groups. PNI was strongly associated with the need for major amputation.

Conclusion: Nutritional status, especially PNI and serum albumin, and inflammatory markers are valuable predictors of major amputation in patients with DFIs. Early assessments of nutritional and inflammatory status may reduce the risk of amputation and improve prognosis.

Keywords: Diabetic foot infection, major amputation, prognostic nutritional index, inflammation, albumin

ÖZ

Amaç: Diyabetik ayak enfeksiyonu (DAE), diabetes mellitusun ciddi komplikasyonlarından biridir ve alt ekstremite amputasyonlarının başlıca nedenleri arasında yer alır. Majör amputasyon gerekliliğini öngörecek belirteçlerin tanımlanması, tedavi planlamasını kolaylaştırabilir ve hasta sonuçlarını iyileştirebilir.

Gereç ve Yöntem: Bu retrospektif, kesitsel çalışmada, Aralık 2022 ile Temmuz 2024 tarihleri arasında DAE nedeniyle takip edilen 77 hasta değerlendirildi. Hastalara ait klinik ve laboratuvar verileri [C-reaktif protein (CRP), eritrosit sedimentasyon hızı, lökosit sayısı, hemoglobin, albümin, prognostik beslenme indeksi (PBİ) gibi] toplandı. Amputasyonlar anatomik düzeye göre majör ve majör olmayan olarak sınıflandırıldı. İstatistiksel analizlerle majör amputasyonu öngören parametreler incelendi.

Bulgular: Hastaların %25,9'unda majör amputasyon yapıldı. Majör amputasyon uygulanan grupta lökosit, nötrofil, trombosit, CRP ve sedimentasyon düzeyleri anlamlı şekilde yüksek; hemoglobin, albümin ve PBİ değerleri ise anlamlı şekilde düşüktü ($p<0,05$). Nötrofil/lenfosit oranı, trombosit/lenfosit oranı ve sistemik immün-enflamatuvar indeks değerlerinde anlamlı fark gözlenmedi. Düşük PBİ skoru majör amputasyonla güçlü şekilde ilişkiliydi.

Address for Correspondence: Ayşegül İnci Sezen, MD, University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Infectious Diseases and Clinical Microbiology, İstanbul, Türkiye
E-mail: incinarin@hotmail.com **ORCID ID:** orcid.org/0000-0001-8920-9019

Cite as: Sezen Aİ, Özdemir YE, Borcak D, Karzaoğlu M, Bali Yurdakul K, Günay D, et al. Nutritional and inflammatory markers as predictors of major amputation in patients with diabetic foot infections. Med J Bakirkoy. 2025;21(4):380-385

Received: 17.07.2025
Accepted: 16.10.2025
Publication Date: 18.12.2025

ÖZ

Sonuç: DAE olan hastalarda beslenme durumu (özellikle PBİ ve serum albümin) ve enflamatuvar belirteçler majör amputasyon riskini öngörmede değerli olabilir. Bu parametrelerin erken değerlendirilmesi, amputasyon riskini azaltabilir ve prognozu iyileştirebilir.

Anahtar Kelimeler: Diyabetik ayak enfeksiyonu, majör amputasyon, prognostik beslenme indeksi, enflamasyon, albümin

INTRODUCTION

Diabetes mellitus is a chronic illness that affects several organ systems and is increasing in prevalence worldwide. Diabetes-related complications can cause serious harm, especially to the cardiovascular, neurological, and microvascular systems, and increase morbidity and mortality among patients. Diabetic foot infections (DFIs) represent one of the most severe complications of diabetes and are among the leading causes of hospitalizations and lower extremity amputations (1).

Reports indicate that 40-60% of non-traumatic lower-extremity amputations are due to DFIs (2). In diabetic patients, the development of peripheral neuropathy, peripheral arterial disease, and a weakened immune system impairs wound healing and predisposes patients to infection. Amputation becomes inevitable in cases where infection cannot be controlled, tissue necrosis occurs, or systemic complications develop (3).

It is well established that the extent of amputation directly affects patients' clinical prognosis. When there is extensive tissue loss or systemic spread, major amputation is the preferred option. Minor amputations, on the other hand, aim to preserve the limb's function. It is essential to understand the factors that influence the level of amputation in order to plan surgery for each patient and to prevent potential complications. The goal of this study was to identify the clinical and laboratory factors that led to the need for major amputation in patients with DFI.

METHODS

This single-center, retrospective, cross-sectional study included patients aged 18 years or older who were hospitalized for DFIs and followed-up at our hospital's wound care clinic between December 2022 and July 2024. Only patients with complete medical records were included in the study.

Demographic data (age, sex), comorbidities (hypertension, coronary artery disease, chronic kidney disease, peripheral artery disease, etc.), and laboratory parameters [leukocyte count, hemoglobin (Hb), neutrophil count, lymphocyte count, platelet count, creatinine, HbA1c, etc.] were retrospectively collected from the electronic medical records system.

Amputations performed proximal to the ankle joint were classified as "major amputations," while those performed distal to the ankle joint were classified as "minor amputations."

Inflammatory markers [C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), procalcitonin], blood-cell ratio indices [platelet-to-lymphocyte ratio (PLR), neutrophil-to-lymphocyte ratio (NLR), systemic immune-inflammation index (SII)=(platelet count×neutrophil count)/lymphocyte count], and nutritional indices such as the prognostic nutritional index (PNI)=serum albumin (g/L)+[0.005×total lymphocyte count (mm^3)] and controlling nutritional status scores were examined to determine whether significant differences between the two groups. Causative pathogens were identified through deep tissue culture samples obtained during the debridement process.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics software (version 29, SPSS Inc., Chicago, IL, USA). Normality of the distribution of continuous variables was assessed using the Shapiro-Wilk test. For non-parametric variables, the Mann-Whitney U test was used; for parametric variables, the Student's t-test was applied. The chi-square test or Fisher's exact test, as applicable, was used to analyze categorical variables. A statistically significant difference was defined as $p<0.05$.

Ethical Approval

This study was approved by the University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital Non-Interventional Scientific Research Ethics Committee (approval no: 2025-08-12; date: 24.04.2025). As the study was retrospective and personal data were anonymized, informed consent was not required.

RESULTS

Of the 77 patients included in the study, 87% (n=67) were male, with a mean age of 62.6±11.9 years. The most common comorbidities were hypertension (n=46, 59.7%), coronary artery disease (n=36, 46.7%), chronic kidney disease (n=25, 32.4%), and peripheral arterial disease (n=25, 32.4%).

Major amputations were performed in 20 patients (25.9%), and minor amputations were performed in 15 patients (19.4%). Patients who underwent major amputation had significantly higher leukocyte counts (13006±5193 vs. 10160±4763; p=0.032), neutrophil counts (10638±5443 vs. 7975±4669; p=0.034), platelet counts (368750±107035 vs. 309166±105900; p=0.041), CRP levels (109±69 vs. 76±73; p=0.027), and ESR levels (82±30 vs. 60±30; p=0.004). In contrast, Hb (9.8±1.1 vs. 10.8±1.9; p=0.010), albumin (28±4 vs. 34±5; p<0.001), and PNI scores (36.79±6.11 vs. 42.17±6.26; p=0.003) were significantly lower in the major amputation group (Table 1).

Although patients with severe diabetes (HbA1c>10%) had a higher rate of major amputations, this difference was not statistically significant (38.4% vs. 23.4%; p=0.266).

A pathogenic microorganism was identified in 51 patients (66.2%). Among these, 9 (17.6%) had Gram-positive organisms, 26 (50.9%) had Gram-negative organisms, and 16 (31.5%) had polymicrobial infections. A total of 67 isolates were obtained. The most common pathogens identified

Table 1. Comparison of diabetic foot infection patients with and without major amputation

Variables	Total (n=77) n (%)/mean±SD	Major amputation + (n=20) n (%)/mean±SD	Major amputation - (n=57) n (%)/mean±SD	p-value	OR
Age	62.6±11.9	64.9±10.7	61.8±12.2	0.323	
Gender					
Male Female	67 (87) 10 (13)	18 (90) 2 (10)	49 (86) 8 (14)	0.644	0.68
Hypertension	46 (59.7)	12 (60)	34 (59.6)	0.970	1.01
Coronary artery disease	36 (46.7)	10 (50)	26 (45.6)	0.735	1.19
Chronic kidney disease	25 (32.4)	7 (35)	18 (31.6)	0.784	1.16
Peripheral artery disease	25 (32.4)	7 (35)	27 (47.4)	0.338	0.91
Laboratory results					
HbA1c	8.1±1.7	8.4±1.7	8.0±1.7	0.368	
White blood cell count (cell/µL)	11899±5003	13006±5193	10160±4763	**0.032**	
Hemoglobin (g/dL)	10.6±1.8	9.8±1.1	10.8±1.93	**0.010**	
Neutrophil count (cell/µL)	8667±4985	10638±5443	7975±4669	**0.034**	
Lymphocyte count (cell/µL)	1551±636	1550±719	1551±612	0.70	
Platelet count (cell/µL)	324642±108716	368750±107035	309166±105900	**0.041**	
C-reactive protein (mg/dL)	85±73	109±69	76±73	**0.027**	
Procalcitonin (ng/mL)	0.34±1.17	0.30±0.35	0.35±1.30	0.076	
Glucose (mg/dL)	171±86	179±92	169±84	0.86	
Blood urea nitrogen (mg/dL)	55±30	60±33	53±29	0.291	
Creatinine (mg/dL)	1.24±1.06	1.18±0.82	1.44±1.58	0.306	
Alanine aminotransferase (U/L)	19±25	19±14	20±27	0.214	
Aspartate aminotransferase (U/L)	22±25	23±14	21±28	0.176	
Albumin (g/L)	33±5	28±4	34±5	**<0.001**	
Total cholesterol (mg/dL)	148±44	156±48	146±43	0.479	
Triglycerides (mg/dL)	166±94	173±70	163±102	0.06	
Erythrocyte sedimentation rate	65±31	82±30	60±30	**0.004**	
PLR	231±95	262±99	221±91	0.76	
NLR	6.8±6.0	9.0±7.3	6.0±5.4	0.119	
SII	79.4±59.0	79.1±67.4	79.5±56.9	0.76	
PNI	40.8±6.6	36.79±6.11	42.17±6.26	0.003	
CONUT score	4.1±2.6	3.5±2.7	4.3±2.6	0.235	

PLR: Platelet-to-lymphocyte ratio, NLR: Neutrophil-to-lymphocyte ratio, SII: Systemic immune-inflammation index=(platelet count×neutrophil count)/lymphocyte count, PNI: Prognostic nutritional index=serum albumin (g/L)+[0.005×total lymphocyte count (mm^3)], CONUT: Controlling nutritional status score, OR: Odds ratio, SD: Standard deviation, HbA1c: Hemoglobin A1c

were *Proteus mirabilis* (n=18, 26.8%), *Enterococcus* spp. (n=10, 14.9%), *Klebsiella pneumoniae* (n=7, 10.4%), *Pseudomonas aeruginosa* (n=7, 10.4%), and *Escherichia coli* (n=6, 8.9%).

DISCUSSION

In our study, the demographic and clinical characteristics were comparable between the two groups. While inflammation is known to be a poor prognostic indicator in DFIs, previous studies have reported associations of inflammatory markers such as NLR, PLR, and SII with amputation and osteomyelitis in DFI patients (4-6). However, we did not observe significant differences in NLR, PLR, or SII values in patients who underwent major amputation. In contrast, a lower PNI was associated with an increased risk of major amputation in DFIs in our study.

PNI has been utilized as a diagnostic and prognostic tool in various conditions in which inflammation plays a central role, including malignancies, infectious diseases, and cardiovascular disorders (7-9). The PNI reflects the relationship between immune function and nutritional status, as its components are serum albumin level and total lymphocyte count. In line with our results, previous studies have demonstrated that reduced albumin levels are strongly associated with major amputations and mortality (10). Our findings revealed that the mean serum albumin level was 28±4 g/L in the major amputation group, compared with 34±5 g/L in the minor/non-amputation group; this difference was statistically significant ($p<0.001$).

Albumin is widely recognized as an indicator of nutritional status; however, in patients with diabetic foot, it also reflects disease severity, prognosis, and malnutrition (11,12). PNI includes both nutritional and inflammatory indicators, such as albumin level and lymphocyte count. Lower serum albumin levels indicate both poor nutritional status and increased disease severity (13). Additionally, nutritional status significantly influences immune function, and protein-energy malnutrition has been associated with immunodeficiency (14). Lower lymphocyte counts may suggest immune suppression or dysfunction (15). As in other studies (16,17), our study found that lymphocyte counts did not differ significantly between groups with and without major amputations. The use of the PNI formula, compared with lymphocyte counts alone, allowed us to observe a significant prognostic effect in the major amputation group. This situation also highlights the importance of malnutrition in these patients.

Malnutrition is also a significant nutritional concern among patients with diabetes mellitus (18,19). Given the frequent coexistence of diabetes and hypertension, vascular impairments are almost inevitable in patients with diabetic foot (20). This contributes to the development of ischemia, poor wound healing, and increased susceptibility to infection (21). Vascular impairment is a known risk factor for amputation in individuals with diabetic DFIs (21). Malnutrition, in combination with vascular pathologies, contributes to an increased risk of major amputation (21). Inflammation is a major contributor to the development of diabetes-related complications (22).

Hb levels were significantly lower in the major amputation group (9.8±1.1 g/dL vs. 10.8±1.9 g/dL; $p=0.010$). Low Hb levels typically indicate reduced oxygen-carrying capacity and, therefore, an increased risk of tissue hypoxia. In DFIs, inadequate oxygenation can delay wound healing, leading to infection progression and necrosis. This condition can be considered a significant risk factor that increases the likelihood of amputation. Furthermore, anemia can often be an indicator of underlying systemic problems such as chronic inflammation, malnutrition, and renal dysfunction. The literature demonstrates a clear relationship between anemia severity and DFI severity (23). Studies have also reported that low Hb levels are associated with poor prognosis, increased amputation rates, and even mortality (23).

In our study, leukocyte, neutrophil, and platelet counts, as well as CRP and ESR, were significantly higher in the group undergoing major amputation. Elevated levels of these parameters indicate that the infection has progressed to a systemic level and a widespread inflammatory response has developed. In diabetic patients, immune system activation in response to infection typically manifests as increased total leukocyte and neutrophil counts. High CRP and ESR levels are indicators of an acute-phase response and ongoing tissue inflammation. An increase in platelet count may indicate reactive thrombocytosis, a condition that develops in response to the release of cytokines during infection and inflammation. These findings suggest that inflammation is more severe and its systemic effects are more pronounced in cases requiring major amputation. Similarly, previous studies have reported mean CRP, white blood cell, ESR, and PLT values were significantly higher in patients with major amputations than in other patient groups (24).

Our study adds a new perspective to clinical practice by highlighting not only inflammatory markers but also parameters such as the PNI, which reflects nutritional status, as predictors of the risk of major amputation in patients with DFIs. One novel aspects of this study is that although PNI has previously been used primarily in oncology and cardiovascular disease, this study clearly demonstrates,

for the first time, its significant association with major amputation in this patient group.

Furthermore, the combined evaluation of inflammatory markers (CRP, ESR, leukocyte count) with hematological and nutritional indicators offers a multidimensional approach to assessing amputation risk. Data obtained from Türkiye are crucial for filling a gap in the literature on this topic and for demonstrating that low-cost, accessible laboratory parameters (PNI, Hb, albumin) can be used effectively in risk assessment, particularly in developing countries.

Study Limitations

There are certain limitations to this study. First, the retrospective and single-center design may limit the generalizability of the findings. Second, the nutritional assessment might have been affected by the lack of information on the patients' body mass indices and dietary intake on admission. Third, the unequal distribution between major and minor amputations limited subgroup comparisons because major amputations were relatively few. In future studies with larger and more balanced cohorts, it may be possible to better evaluate whether PNI differs significantly between groups undergoing major versus minor amputations. Despite these limitations, one significant advantage of our research is that every patient was treated by the same clinical team, ensuring standardized treatment approaches and minimizing variation that might affect amputation outcomes.

CONCLUSION

Among patients with DFIs, both increased inflammatory markers and decreased albumin and PNI, reflecting poor nutritional status, may serve as significant predictors of major amputation. Therefore, in addition to monitoring inflammatory markers, nutritional support—particularly albumin supplementation—should not be neglected during patient follow-up.

ETHICS

Ethics Committee Approval: This study was approved by the University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital Non-Interventional Scientific Research Ethics Committee (approval no: 2025-08-12; date: 24.04.2025).

Informed Consent: Retrospective study.

FOOTNOTES

Authorship Contributions

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

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

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

REFERENCES

1. Armstrong DG, Boulton AJM, Bus SA. Diabetic foot ulcers and their recurrence. N Engl J Med. 2017;376:2367-75.
2. McDermott K, Fang M, Boulton AJ, Selvin E, Hicks CW. Etiology, epidemiology, and disparities in the burden of diabetic foot ulcers. Diabetes Care. 2023;46:209-21.
3. Maity S, Leton N, Nayak N, Jha A, Anand N, Thompson K, et al. A systematic review of diabetic foot infections: pathogenesis, diagnosis, and management strategies. Front Clin Diabetes Healthc. 2024;5:1393309.
4. Aragón-Sánchez J, Víquez-Molina G, López-Valverde ME, Rojas-Bonilla JM. Systemic immune-inflammation index in diabetic foot infections and osteomyelitis. Int J Low Extrem Wounds. 2025;24:1278-80.
5. Yapıcı O, Berk H, Öztoprak N, Seyman D, Tahmaz A, Merdin A. Can ratio of neutrophil-to-lymphocyte count and erythrocyte sedimentation rate in diabetic foot infecti on predict osteomyelitis and/or amputation? Hematol Rep. 2017;9:6981.
6. Afshar F, Daraie M, Mohammadi F, Seifouri K, Amin Afshari S, Heidari Some'eh S, et al. Neutrophil-lymphocyte ratio (NLR); an accurate inflammatory marker to predict diabetic foot ulcer amputation: a matched case-control study. BMC Endocr Disord. 2025;25:120.
7. Liu X, Duan Y, Wang Y, Zhang X, Lv H, Li Q, et al. Predictive value of prognostic nutritional index as prognostic biomarkers in patients with lymphoma: a systematic review and meta-analysis. Clin Transl Oncol. 2025;27:1274-87.
8. Huang Y, Zhang Q, Li P, Chen M, Wang R, Hu J, et al. The prognostic nutritional index predicts all-cause mortality in critically ill patients with acute myocardial infarction. BMC Cardiovasc Disord. 2023;23:339.
9. Karagöz İ, Özer B, Aktas G. The predictors of outcome in patients that require management in intensive care units: a narrative review. Hitit Med J. 2024;6:367-78.
10. Coşkun B, Ayhan M, Ulusoy S. Relationship between prognostic nutritional index and amputation in patients with diabetic foot ulcer. Diagnostics (Basel). 2024;14:738.
11. Xu S, Wang Y, Hu Z, Ma L, Zhang F, Liu P. Effects of neutrophil-to-lymphocyte ratio, serum calcium, and serum albumin on prognosis in patients with diabetic foot. Int Wound J. 2023;20:1638-46.
12. Gremese E, Bruno D, Varriano V, Perniola S, Petricca L, Ferraccioli G. Serum albumin levels: a biomarker to be repurposed in different disease settings in clinical practice. J Clin Med. 2023;12:6017.
13. Hong J, Huang QQ, Liu WY, Hu X, Jiang FF, Xu ZR, et al. Three nutritional indices are effective predictors of mortality in patients with type 2 diabetes and foot ulcers. Front Nutr. 2022;9:851274.
14. Bourke CD, Berkley JA, Prendergast AJ. Immune dysfunction as a cause and consequence of malnutrition. Trends Immunol. 2016;37:386-98.

15. Karagoz I, Peltek Ozer S, Ozer B, Aktas G. Prognostic nutritional index could serve as a reliable prognostic marker in intensive care population. Med Sci (Basel). 2025;13:59.
16. Kahraman S, Zencirkıran Aguş H, Kalkan AK, Uzun F, Ertürk M, Kalkan ME, et al. Prognostic nutritional index predicts mortality in infective endocarditis. Turk Kardiyol Dern Ars. 2020;48:392-402.
17. Ünver Ulusoy T, Hekimoğlu CH, Kayhan S, Altın N, Şencan İ. Prognostic nutritional index: is it associated with the prognosis of Crimean Congo hemorrhagic fever. J Med Virol. 2022;94:4910-7.
18. Skórka M, Bazaliński D, Więch P, Kłęk S, Kozieł D, Sierżantowicz R. Nutritional status in a group of patients with wounds due to diabetic foot disease and chronic venous insufficiency. J Clin Med. 2024;14:43.
19. Wadivkar P, Thomas N, Jebasingh F, Bacot-Davis VR, Maini R, Hawkins M. Undernutrition-associated diabetes mellitus: pathophysiology of a global problem. Physiology (Bethesda). 2025;40:0.
20. Zhang HZ, Wang YH, Ge YL, Wang SY, Sun JY, Chen LL, et al. Obesity, malnutrition, and the prevalence and outcome of hypertension: evidence from the National Health and Nutrition Examination Survey. Front Cardiovasc Med. 2023;10:1043491.
21. Jude EB, Eleftheriadou I, Tentolouris N. Peripheral arterial disease in diabetes--a review. Diabet Med. 2010;27:4-14.
22. Nguyen DV, Shaw LC, Grant MB. Inflammation in the pathogenesis of microvascular complications in diabetes. Front Endocrinol (Lausanne). 2012;3:170.
23. Yammine K, Hayek F, Assi C. Is there an association between anemia and diabetic foot ulcers? A systematic review and meta-analysis. Wound Rep Reg. 2021;29:432-42.
24. Demirkol D, Aktaş Ş, Özcan T, Tannier X, Selçukcan Erol Ç. Analysis of risk factors for amputation in patients with diabetic foot ulcers: a cohort study from a tertiary center. Acta Orthop Traumatol Turc. 2022;56:333-9.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):386-391
DOI: 10.4274/BMJ.galenos.2025.2025.9-5

Research

Vertebroplasty vs. Balloon Kyphoplasty for Vertebral Compression Fractures: A 331-Patient Experience

Vertebral Kompresyon Kırıklarında Vertebroplasti ve Balon Kifoplasti: 331 Hastadan Oluşan Tek Merkezli Bir Kohort

Aykut Gökbel[1], Atakan Emengen[2], Eren Yılmaz[1], Ayşe Uzuner[3], Savaş Ceylan[2]

[1]İstinye University Faculty of Medicine, Department of Neurosurgery, İstanbul, Türkiye
[2]Bahçeşehir Universty Faculty of Medicine, Department of Neurosurgery, İstanbul, Türkiye
[3]Afşin State Hospital, Clinic of Neurosurgery, Kahramanmaraş, Türkiye

ABSTRACT

Objective: Vertebral augmentation with percutaneous vertebroplasty (PVP) or balloon kyphoplasty (BKP) is widely used for painful vertebral compression fractures (VCFs), yet examining that examining single-center outcomes across mixed etiologies remains informative.

Methods: We retrospectively analyzed 331 consecutive patients treated with PVP (n=155) or BKP (n=176). Mean age was 57.2 years, and 202 participants (61.0%) were women. Overall, etiologies were traumatic: 166 (50.2%) and osteoporotic: 165 (49.8%); men predominantly had traumatic fractures (101/129), whereas women predominantly had osteoporotic fractures (137/202). The primary endpoint was early pain change [Visual Analog Scale (VAS)]. Safety endpoints included radiographic/symptomatic cement leakage, new/adjacent symptomatic fractures, cardiopulmonary events, infection, and reoperation for progressive kyphosis.

Results: Mean VAS improved from 8.23±0.82 to 3.03±1.60 at early follow-up (Δ-5.19, p<0.001). Radiographic cement leakage occurred in 13/331 (3.9%); with 1/331 (0.3%) of these cases being symptomatic. New/adjacent symptomatic VCFs occurred in 24/331 (7.2%). Other complications were infrequent: pulmonary embolism occurred in 3 out of 331 cases (0.9%), myocardial infarction in 1 out of 331 cases (0.3%), and infection in 1 out of 331 cases (0.3%). Reoperation for progressive kyphosis (posterior instrumentation and fusion) was required in 17/331 (5.1%) (4 men, 13 women). Outcomes were favorable across both techniques, with choice individualized by fracture morphology and alignment goals.

Conclusion: In this large, single-center experience, vertebral augmentation provided rapid, clinically meaningful pain relief with low symptomatic complication rates and modest new fracture incidence. The observed 5.1% reoperation rate underscores the need for longitudinal alignment surveillance and aggressive osteoporosis management. Findings support selection-driven use of PVP or BKP with meticulous cement management and structured follow-up.

Keywords: Vertebroplasty, kyphoplasty, vertebral compression fractures, osteoporosis, traumatic spine fractures

ÖZ

Amaç: Ağrılı vertebra kompresyon kırıklarında vertebral augmentasyon [perkütan vertebroplasti (PVP) ve balon kifoplasti (BKP)] sık uygulanmaktadır. Bu çalışma, tek merkezde ardışık olgular üzerinden PVP/BKP'nin kısa dönem klinik ve güvenlik sonuçlarını ortaya koymayı amaçlamıştır.

Gereç ve Yöntem: Retrospektif olarak 331 ardışık hasta incelendi (PVP, n=155; BKP, n=176). Ortalama yaş 57,2 yıl; 202 (%61,0) kadın idi. Etiyoloji dağılımı genel kohortta travmatik 166 (%50,2) ve osteoporotik 165 (%49,8) olup erkeklerde travma (101/129), kadınlarda osteoporoz (137/202) baskındı. Birincil sonlanım erken dönemde ağrı değişimiydi [Görsel Analog Skala (GAS)]. Güvenlik sonlanımları; radyografik/semptomatik çimento kaçışı, yeni/komşu semptomatik kırık, kardiyopulmoner olaylar, enfeksiyon ve kifoz progresyonu nedeniyle reoperasyonu içerdi.

Bulgular: Ortalama GAS 8,23±0,82'den 3,03±1,60'a geriledi (Δ-5,19; p<0,001). Radyografik çimento kaçışı 13/331 (%3,9) olup 1/331 (%0,3) olguda semptomatikti. Yeni/komşu semptomatik kırık 24/331 (%7,2) saptandı. Diğer komplikasyonlar seyrek izlendi: pulmoner emboli 3/331 (%0,9), miyokard enfarktüsü 1/331 (%0,3), enfeksiyon 1/331 (%0,3). Kifoz progresyonu nedeniyle posterior enstrümantasyon-füzyon 17/331 (%5,1) hastada (4 erkek, 13 kadın) uygulandı. Sonuçlar her iki teknikte de olumluydu; yöntem seçimi kırık morfolojisi ve deformite hedeflerine göre bireyselleştirildi.

Sonuç: Geniş tek merkez serimizde vertebral augmentasyon hızlı ve anlamlı analjezi sağlarken semptomatik komplikasyon oranları düşük, yeni kırık insidansı ılımlı düzeyde kaldı. %5,1'lik reoperasyon oranı, uzunlamasına deformite izlemi ve agresif osteoporoz tedavisinin önemini

Address for Correspondence: Savaş Ceylan, Prof, Afşin State Hospital, Clinic of Neurosurgery, Kahramanmaraş, Türkiye
E-mail: ssceylan@yahoo.com **ORCID ID:** orcid.org/0000-0002-2747-8907

Cite as: A, Emengen A, Yılmaz E, Uzuner A, Ceylan S. Vertebroplasty vs. balloon kyphoplasty for vertebral compression fractures: a 331-patient experience. Med J Bakirkoy. 2025;21(4):386-391

Received: 29.09.2025
Accepted: 08.10.2025
Publication Date: 18.12.2025

ÖZ

vurgulamaktadır. Bulgular, özenli hasta seçimi ve titiz çimento yönetimi eşliğinde PVP/BKP'nin uygun endikasyonlarda etkin ve güvenli olduğunu desteklemektedir.

Anahtar Kelimeler: Vertebroplasti, kifoplasti, vertebral kompresyon kırıkları, osteoporoz, travmatik omurga kırıkları

INTRODUCTION

Osteoporotic vertebral compression fractures (VCFs) are a leading cause of pain, disability, and excess mortality in older adults. When analgesics, bracing, and activity modification fail, image-guided vertebral augmentation—percutaneous vertebroplasty (PVP) and balloon kyphoplasty (BKP)—is considered to achieve rapid analgesia and stabilization (1,2). Randomized and systematic evidence indicates that both procedures can improve pain and function, whereas BKP offers greater vertebral height restoration and kyphotic angle correction in selected patients (3-6). Although two sham-controlled trials questioned routine use of PVP in chronic or subacute fractures (7,8), subsequent randomized trials and meta-analyses, particularly in acute fractures and oncology settings, support BKP over non-surgical management for early quality of life gains, without excess peri-procedural risk (5,6). Safety profiles are favorable overall; the most frequent event is cement leakage—typically asymptomatic—with symptomatic neurological or cardiopulmonary complications, being uncommon and mitigable through technique optimization (2,3,9). Importantly, large registry and meta-analytic data associates vertebral augmentation with reduced long-term mortality compared with non-surgical care (≈22% relative reduction) (10).

Our team's series of 331 consecutive PVP/BKP procedures reports clinical outcomes, radiographic changes, and complications, and explores technical nuances that may influence safety (e.g., approach selection, cement viscosity control, unipedicular vs. bipedicular cannulation).

METHODS

Study Design

We conducted a retrospective, single-center cohort study of consecutive patients undergoing PVP or BKP for painful thoracic and/or lumbar VCFs between 2020 and 2025 at the department of neurosurgery. Ethical approval was obtained from Bahçeşehir University Non-Interventional Studies Ethics Committee (approval no: 2025-14/08, date: 19.09.2025). All patients provided written informed consent consistent with local regulations.

Patient Selection

Inclusion criteria

(I) Focal back pain attributable to one or more VCFs; (II) magnetic resonance imaging (MRI) or computed tomography (CT) confirmation, with active edema on short-tau inversion-recovery/T2 when available; (III) refractoriness to ≥2-4 weeks of conservative therapy; (IV) osteoporotic, traumatic, or neoplastic etiology.

Exclusion criteria

Uncontrolled coagulopathy; active infection; unstable burst fracture with significant posterior wall retropulsion compressing neural elements; allergy to polymethyl methacrylate (PMMA); inability to tolerate prone positioning. Selection aligns with contemporary policy/guideline statements (6,11).

Preoperative Assessment

Baseline pain [Visual Analog Scale (VAS)] and function [oswestry disability index (ODI)] were recorded. Standing lateral radiographs, and when indicated, CT/MRI, were used to characterize fracture level(s), vertebral height loss, kyphotic angle, intravertebral cleft, and posterior wall integrity. Bone mineral density (BMD) was documented via dual-energy X-ray absorptiometry in osteoporotic cases. Anticoagulants/antiplatelets were managed per institutional protocol.

Anesthesia, Positioning, and Imaging

Procedures were performed in a dedicated operating room with biplanar or high-resolution C-arm fluoroscopy. Local anesthesia with conscious sedation was preferred. General anesthesia was used for non-cooperative patients or extensive multilevel augmentation. Patients were positioned prone on a radiolucent table with chest/pelvic bolsters to reduce epidural venous pressure.

Surgical Procedure

Level-specific approaches were selected to optimize safety: transpedicular access for lumbar levels, parapedicular/costotransverse for mid-thoracic levels with narrow pedicles. Unipedicular cannulation was attempted when central vertebral body cross-fill could be achieved; bipedicular access was used when cement spread was predicted to

be insufficient, or in biconcave deformities. These choices are consistent with comparative and review data on access safety and cement distribution (2-4) (Figure 1).

Balloon kyphoplasty: After guidewire and working cannula placement, a balloon tamp was advanced into the vertebral body and gradually inflated under continuous anteroposterior and lateral fluoroscopy until (I) cortical resistance was felt, (II) desired height restoration/kyphosis correction was achieved, or (III) manufacturer-recommended pressure limits were reached. The balloon was deflated and removed, creating a cavity. Medium-to-high viscosity PMMA was prepared and injected in 0.5-1.0 mL aliquots to fill the cavity while avoiding posterior third encroachment. BKP has shown superior restoration of vertebral height and local alignment vs. PVP and non-surgical care in trials/meta-analyses (3-6).

Table 1. Demographic and procedural characteristics

Variable	p-value
Number of patients	331
Mean age (years)	57.2
Sex (female/male)	202/129
Etiology-traumatic	166 (50.2%)
Etiology-osteoporotic	165 (49.8%)
PVP cases	182 (55%)
BKP cases	149 (45%)
Access type (PVP)-unipedicular	125/182 (68.6%)
Access type (PVP)-bipedicular	57/182 (31.4%)
Complications	
Radiographic cement leakage	13 (3.9%)
Symptomatic cement leakage	1 (0.3%)
New/adjacent symptomatic fractures	24 (7.2%)
Reoperation for progressive kyphosis	17 (5.1%)
Pulmonary embolism	3 (0.9%)
Myocardial infarction	1 (0.3%)
Infection	1 (0.3%)

PVP: Percutaneous vertebroplasty, BKP: Balloon kyphoplasty

During vertebroplasty (PVP), following cannula placement into the anterior third of the vertebral body, PMMA at optimal (toothpaste-like) viscosity was injected slowly under continuous biplanar fluoroscopy. Injection was paused whenever cement approached the posterior third, or extravasation was suspected. The endpoint was homogeneous trabecular filling with mechanical stabilization. Although early sham-controlled trials in mixed-chronicity fractures did not show superiority of PVP over simulated procedures (7,8), subsequent evidence supports careful patient selection (acute, edematous fractures; concordant pain), and meticulous technique to optimize outcomes (1,3,5).

Cement powder-to-monomer ratios followed manufacturer guidance. To mitigate leakage, we used high-viscosity materials, low-pressure incremental injection, and avoided

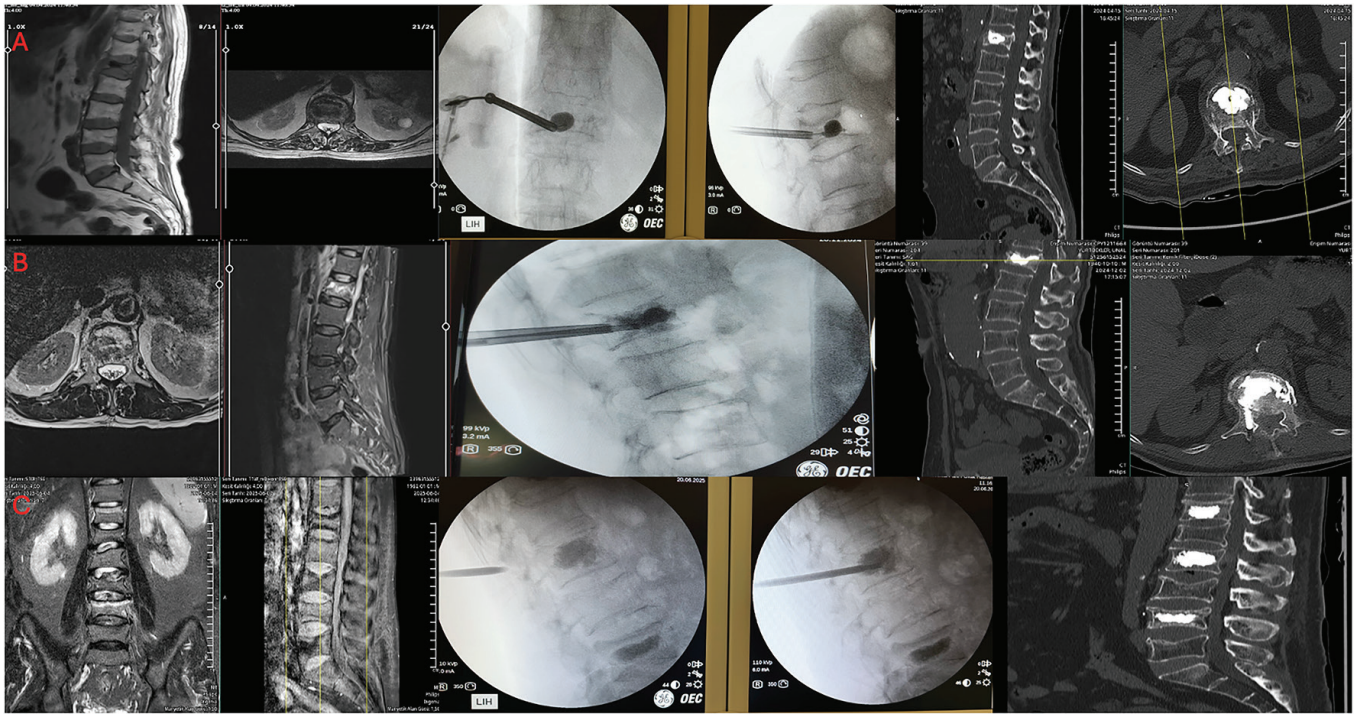


Figure 1. A-C) Preoperative, intraoperative, and postoperative imaging in patients treated with percutaneous vertebroplasty and balloon kyphoplasty

cortical defects when feasible. In clefted fractures or posterior wall compromise, BKP was favored to create a low-pressure cavity, where PVP was performed in such contexts, reduced volumes and earlier cessation thresholds were applied. Recognized leakage risk factors include intravertebral cleft, cortical disruption, low viscosity, and larger injected volumes. Symptomatic neurological/cardiopulmonary events are rare but recognized, vigilance during injection is critical (2,9). Cannulas were withdrawn, skin sites were dressed without sutures. Patients underwent serial neurological checks for 2-4 hours and were able to ambulate the same day when feasible. Bracing for 4-6 weeks was individualized for each patient.

Outcome Measures

Primary outcomes: Change in VAS and ODI from baseline to 24 hours, 1-month, and 6 months.

Secondary outcomes: Vertebral body height restoration and local kyphotic angle correction; procedure time; length of stay.

Safety outcomes: Cement leakage (radiographic/symptomatic), adjacent vertebral fractures, pulmonary embolism, new neurological deficits, infection, and reinterventions. Outcome selection reflects prior randomized and observational literature (2,5,6,10).

Statistical Analysis

Continuous data are reported as mean±SD or median (interquartile range) and compared with paired t-tests or Wilcoxon signed-rank tests, as appropriate. Categorical variables (e.g., leakage, adjacent-level fracture) were compared using chi-square or Fisher's exact tests. Multivariable logistic regression explored predictors of leakage and adjacent fractures (candidate variables: age, sex, BMD T-score, fracture age, approach, cement volume/viscosity, intravertebral cleft, posterior wall defect). Significance was set at $p<0.05$ (two-sided). Analyses were conducted in SPSS (IBM Corp., Armonk, NY, USA).

RESULTS

Cohort, Procedures, and Perioperative Metrics

The cohort comprised 331 patients (mean age 57.2 years), including 202 women (61.0%) and 129 men (39.0%). Overall distribution of etiologies was balanced: traumatic fractures in 166/331 (50.2%) and osteoporotic fractures in 165/331, (49.8%). By sex, men had predominantly traumatic fractures (101/129; 78.3%), whereas women had predominantly osteoporotic fractures (137/202; 67.8%).

A total of 331 consecutive patients were analyzed; 182 PVP (55%) and 149 BKP (45%), including 202 women (61%) and 129 men (39%). Procedures were performed under monitored anesthesia care with local infiltration in most cases; general anesthesia was reserved for a minority. Unipedicular access predominated at lower lumbar levels and bipedicular access at mid-thoracic levels (distribution to be specified). Among PVP cases, access was unipedicular in 125/182 (68.6%) and bipedicular in 57/182 (31.4%) (Table 1). Median cement volume per level was 4.2 mL overall [PVP (3.2 mL); BKP (5.4 mL)], with median fluoroscopy time of 15.9 minutes, and length of stay of 1 day.

Clinical Outcomes

Mean baseline pain for the overall cohort was 8.23±0.82, early post-procedure pain was 3.03±1.6, for an average ΔVAS≈-5.19 ($p<0.001$). Technique-weighted benchmarks suggest a slightly larger absolute VAS reduction with PVP (≈-5.68) than BKP (≈-4.60), yet post-procedure pain levels were clinically similar (PVP$\approx$2.68 vs. BKP$\approx$3.46). Functional scores (ODI) are expected to mirror pain trajectories; center specific values will be entered at 24-h, 1-month, and 6 months.

Radiographic Outcomes

Consistent with comparative literature, BKP is expected to yield greater mean vertebral height restoration and larger kyphotic angle correction than PVP. Height gain typically concentrates in acute, edematous fractures, and at levels with preserved endplate compliance. Final values for anterior/middle height (mm) and local Cobb angle (°) will be entered once radiographic measurements are finalized, between group differences are expected to favor BKP for alignment metrics, with similar clinical analgesia at short-term follow-up.

Complications

Overall radiographic cement leakage occurred in 13/331 procedures (3.9%), driven by higher per-procedure leakage in PVP (8/182) versus BKP (5/149). Most leaks were anterior/lateral and clinically silent. A symptomatic cement leak occurred in one patient during PVP. New or adjacent symptomatic VCFs occurred in 24/331 (7.2%) (PVP 16/182 vs. BKP 8/149). Posterior instrumentation and fusion for progressive kyphosis during follow-up was required in 17/331 patients (5.1%): 4/129 men (3.1%) and 13/202 women (6.4%). Other events were uncommon: pulmonary embolism 2/331 (PVP, 1; BKP, 1), myocardial infarction 1/331 (0.3%; BKP), hematoma 2/331 (0.6%), transient hemodynamic change 1/331 (0.3%), pneumonia/hypoxia 1/331 (0.3%), and infection 1/331 (0.3%).

DISCUSSION

Principal findings and clinical meaning. In this single-center, mixed PVP/BKP series of 331 patients, pain improved substantially (≈5-point early VAS decrease), radiographic cement leakage was low (3.9%) with only one symptomatic event, and new/adjacent symptomatic VCFs were infrequent (7.2%). A 5.1% reoperation rate for progressive kyphosis (posterior instrumentation/fusion) provides a meaningful, patient-centered endpoint. These results align with contemporary evidence that vertebral augmentation delivers rapid analgesia with a generally favorable safety profile. The balance between PVP and BKP often reflects anatomy, fracture chronicity, and the need for height restoration (2,8,12-17).

Context within randomized and comparative data. Our early pain trajectory is consistent with randomized controlled trials and meta-analyses showing large, early analgesic effects after augmentation, most notably the BKP FREE program versus non-surgical care, and positive PVP results when acute, edematous fractures are selected (e.g., VERTOS II, VAPOUR). By contrast, the 2009 New England Journal of Medicine sham-controlled trials suggested limited benefit for PVP in mixed-chronicity cohorts; taken together, these data emphasize patient selection and timing rather than a blanket endorsement or rejection of PVP. Head-to-head data suggest broadly similar pain/functional improvements for PVP vs BKP, while BKP tends to yield greater height/kyphosis correction (2,13,14,17).

Radiographic alignment and technique choice. Although our manuscript focuses on clinical outcomes, the access strategy and device selection imply a pragmatic approach: BKP for cases where creating a cavity aids in cement control or height restoration, and PVP when controlled trabecular fill is feasible without cavity preparation. Prior trials and reviews show greater vertebral height restoration and local kyphosis correction with BKP (FREE and subsequent analyses), whereas post-procedure pain tends to be comparable across techniques at short-term follow-up. This supports individualized modality selection, guided by fracture morphology (endplate compliance, intravertebral cleft, posterior wall integrity) (2,8,12,18).

Cement leakage: A technique-sensitive endpoint. Our overall 3.9% radiographic leakage rate (PVP, 8/182; BKP, 5/149), with only one symptomatic case, appears lower than pooled estimates (e.g., ≈20% PVP vs. ≈7% BKP, many asymptomatic) from older meta-analyses (likely reflecting viscosity-timed injection, incremental low-pressure delivery, and cavity creation when cortical defects or clefts were present). Evidence consistently links intradiscal spread, posterior cortical disruption, low viscosity, and larger injected volumes to higher leakage (and to earlier adjacent fractures), highlighting the importance of stopping rules near the posterior third and avoiding intradiscal migration (13,19,20).

New/adjacent vertebral fractures and biological background. The 7.2% rate of symptomatic new/adjacent fractures in our cohort, is at the lower end of the historical range. Observational data and biomechanical studies suggest that both disease biology (severe osteoporosis, thoracolumbar junction mechanics) and technical factors (intradiscal cement, overcorrection, abrupt stiffness gradients) may influence these events. Notably, intradiscal leakage has been associated with earlier adjacent fractures; rigorous osteoporosis management (antiresorptive or anabolic therapy), attention to cement distribution, and gradual alignment correction remain key mitigations (2,15-17,20).

Reoperation for progressive kyphosis: who is at risk? Our 5.1% reoperation rate (more frequent in women, consistent with the higher osteoporotic burden) underlines that deformity progression can still occur despite initial stabilization particularly in multilevel disease, severe baseline collapse, or when fracture chronicity blunts height recovery. Although high-quality comparative data on post-augmentation fusion are limited, this finding supports longitudinal radiographic monitoring, targeted bracing/rehabilitation, and aggressive bone health optimization. Future work might integrate morphological predictors (cleft, posterior wall breach), cement metrics (viscosity/volume/centroid), and BMD into a practical risk model to anticipate which patients may benefit from earlier deformity control strategies (16,17).

Study Limitations

This study has several limitations. It was a single-team, retrospective analysis, which may limit the generalizability of the findings. The follow-up period was relatively short, possibly underestimating late complications such as adjacent fractures or progressive kyphosis. In addition, the choice between PVP and BKP was based on surgeon preference rather than randomization, introducing potential selection bias. Future multicenter prospective studies with longer follow-up are needed to validate these results.

CONCLUSION

In a 331-patient, mixed PVP/BKP experience, augmentation provided rapid and durable pain relief with low symptomatic complication rates and a modest incidence of new fractures; reoperation for progressive kyphosis occurred in

approximately 5%. Together with randomized and meta-analytic data, these findings support selection-driven use of PVP or BKP, prioritizing cement control, avoidance of intradiscal spread, and alignment goals while pairing the procedure with aggressive osteoporosis therapy and structured follow-up to minimize downstream fractures and deformity.

ETHICS

Ethics Committee Approval: Ethical approval was obtained from Bahçeşehir University Non-Interventional Studies Ethics Committee (approval no: 2025-14/08, date: 19.09.2025).

Informed Consent: All patients provided written informed consent consistent with local regulations.

FOOTNOTES

Authorship Contributions

Surgical and Medical Practices: A.G., A.E., E.Y., S.C., Concept: A.E., E.Y., A.U., S.C., Design: A.G., E.Y., A.U., S.C., Data Collection or Processing: A.E., S.C., Analysis or Interpretation: E.Y., S.C., Literature Search: A.G., S.C., Writing: A.E., S.C.

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

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

REFERENCES

1. Clerk-Lamalice O, Beall DP, Ong K, Lorio MP. ISASS policy 2018-vertebral augmentation: coverage indications, limitations, and/or medical necessity. Int J Spine Surg. 2019;13:1-10.
2. Hulme PA, Krebs J, Ferguson SJ, Berlemann U. Vertebroplasty and kyphoplasty: a systematic review of 69 clinical studies. Spine (Phila Pa 1976). 2006;31:1983-2001.
3. McCall T, Cole C, Dailey A. Vertebroplasty and kyphoplasty: a comparative review of efficacy and adverse events. Curr Rev Musculoskelet Med. 2008;1:17-23.
4. Santiago FR. Comparative review of vertebroplasty and kyphoplasty. World J Radiol. 2014;6:329.
5. Van Meirhaeghe J, Bastian L, Boonen S, Ranstam J, Tillman JB, Wardlaw D, et al. A randomized trial of balloon kyphoplasty and nonsurgical management for treating acute vertebral compression fractures: vertebral body kyphosis correction and surgical parameters. Spine (Phila Pa 1976). 2013;38:971-83.
6. Wardlaw D, Cummings SR, Van Meirhaeghe J, Bastian L, Tillman JB, Ranstam J, et al. Efficacy and safety of balloon kyphoplasty compared with non-surgical care for vertebral compression fracture (FREE): a randomised controlled trial. Lancet. 2009;373:1016-24.
7. Buchbinder R, Osborne RH, Ebeling PR, Wark JD, Mitchell P, Wriedt C, et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. N Engl J Med. 2009;361:557-68.
8. Kallmes DF, Comstock BA, Heagerty PJ, Turner JA, Wilson DJ, Diamond TH, et al. A randomized trial of vertebroplasty for osteoporotic spinal fractures. N Engl J Med. 2009;361:569-79.
9. Zhan Y, Jiang J, Liao H, Tan H, Yang K. Risk factors for cement leakage after vertebroplasty or kyphoplasty: a meta-analysis of published evidence. World Neurosurg. 2017;101:633-42.
10. Hinde K, Maingard J, Hirsch JA, Phan K, Asadi H, Chandra RV. Mortality outcomes of vertebral augmentation (vertebroplasty and/or balloon kyphoplasty) for osteoporotic vertebral compression fractures: a systematic review and meta-analysis. Radiology. 2020;295:96-103.
11. Berenson J, Pflugmacher R, Jarzem P, Zonder J, Schechtman K, Tillman JB, et al. Balloon kyphoplasty versus non-surgical fracture management for treatment of painful vertebral body compression fractures in patients with cancer: a multicentre, randomised controlled trial. Lancet Oncol. 2011;12:225-35.
12. Clark W, Bird P, Gonski P, Diamond TH, Smerdely P, McNeil HP, et al. Safety and efficacy of vertebroplasty for acute painful osteoporotic fractures (VAPOUR): a multicentre, randomised, double-blind, placebo-controlled trial. Lancet. 2016;388:1408-16.
13. Dohm M, Black CM, Dacre A, Tillman JB, Fueredi G; KAVIAR investigators. A randomized trial comparing balloon kyphoplasty and vertebroplasty for vertebral compression fractures due to osteoporosis. AJNR Am J Neuroradiol. 2014;35:2227-36.
14. Klazen CAH, Lohle PNM, de Vries J, Jansen FH, Tielbeek AV, Blonk MC, et al. Vertebroplasty versus conservative treatment in acute osteoporotic vertebral compression fractures (Vertos II): an open-label randomised trial. Lancet. 2010;376:1085-92.
15. Trout AT, Kallmes DF, Kaufmann TJ. New fractures after vertebroplasty: adjacent fractures occur significantly sooner. AJNR Am J Neuroradiol. 2006;27:217-23.
16. Uppin AA, Hirsch JA, Centenera LV, Pfiefer BA, Pazianos AG, Choi IS. Occurrence of new vertebral body fracture after percutaneous vertebroplasty in patients with osteoporosis. Radiology. 2003;226:119-24.
17. Yuan WH, Hsu HC, Lai KL. Vertebroplasty and balloon kyphoplasty versus conservative treatment for osteoporotic vertebral compression fractures: a meta-analysis. Medicine. 2016;95:e4491.
18. Hiwatashi A, Westesson PL, Yoshiura T, Noguchi T, Togao O, Yamashita K, et al. Kyphoplasty and vertebroplasty produce the same degree of height restoration. AJNR Am J Neuroradiol. 2009;30:669-73.
19. Krueger A, Bliemel C, Zettl R, Ruchholtz S. Management of pulmonary cement embolism after percutaneous vertebroplasty and kyphoplasty: a systematic review of the literature. Eur Spine J. 2009;18:1257-65.
20. Lin EP, Ekholm S, Hiwatashi A, Westesson PL. Vertebroplasty: cement leakage into the disc increases the risk of new fracture of adjacent vertebral body. AJNR Am J Neuroradiol. 2004;25:175-80.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):392-397
DOI: 10.4274/BMJ.galenos.2025.2025.8-8

Research

Our Institutional Experience in Laparoscopic Inguinal Hernia Surgery: Is There a Difference Between Right and Left-Sided Procedures? A Retrospective Analysis of 192 Cases

Laparoskopik İnguinal Herni Cerrahisinde Kurum Deneyimimiz: Sağ ve Sol Taraf Cerrahisi Arasında Fark Var mı? 192 Hastalık Retrospektif Değerlendirme

Süleyman Büyükaşık, Yusuf Emre Altundal, Burak Kankaya, Ergün Eskioğlu, Cansu Esen

İstanbul Aydın University Faculty of Medicine, Department of General Surgery, İstanbul, Türkiye

ABSTRACT

Objective: This study aimed to compare surgical duration, mesh placement time, complication and recurrence rates among adult patients undergoing right-sided, left-sided, or bilateral laparoscopic inguinal hernia repair.

Methods: Data from 192 patients who underwent laparoscopic transabdominal preperitoneal hernia repair at a single institution were retrospectively reviewed. Patients were grouped by the side of hernia repair (right, left, or bilateral), and surgical parameters were compared statistically. We used a standard large 3D mesh (10.3x15.7 cm) for hernia repair in all patients.

Results: Surgical duration in patients undergoing bilateral procedures was significantly longer than in patients undergoing single-sided surgeries ($p<0.05$). Although left-sided surgeries lasted an average of three minutes longer than right-sided ones, the difference was not statistically significant ($p=0.512$). Similarly, the mesh placement time was longer for left-sided hernias, but the difference was not statistically significant ($p=0.239$). However, in bilateral procedures, the duration of left-sided mesh placement was significantly longer than that of the right side ($p<0.05$). Complication rates were low (6.7%), and recurrence was detected in six patients.

Conclusion: Laparoscopic inguinal hernia repair is a safe and effective procedure; bilateral repairs requiring a significantly longer operative time. Although left-sided surgeries tend to take slightly longer than right-sided ones, this difference is not statistically significant. These parameters should be considered in surgical planning and patient counseling.

Keywords: Inguinal hernia, laparoscopy, mesh, repair, complication

ÖZ

Amaç: Bu çalışmada, erişkin hastalarda laparoskopik inguinal herni onarımı uygulanan sağ, sol ve bilateral cerrahilerin ameliyat süresi, mesh yerleştirme süresi, komplikasyon oranı ve nüks açısından karşılaştırılması amaçlanmıştır.

Gereç ve Yöntem: Tek bir merkezde laparoskopik transabdominal preperitoneal fıtık onarımı uygulanan 192 hastanın verileri retrospektif olarak incelendi. Hastalar fıtık onarımının yapıldığı tarafa (sağ, sol, bilateral) göre gruplandırıldı ve cerrahi parametreler istatistiksel olarak karşılaştırıldı. Tüm hastalarda fıtık onarımı için büyük 3D mesh (10,3x15,7 cm) kullanıldı.

Bulgular: Bilateral cerrahi uygulanan hastalarda ameliyat süresi, sağ ve sol cerrahilere göre istatistiksel olarak anlamlı şekilde daha uzundu ($p<0,05$). Sağ ve sol inguinal herni onarımı karşılaştırıldığında, sol taraf cerrahisinin süresi sağdan ortalama 3 dakika daha uzun olmasına rağmen fark istatistiksel olarak anlamlı değildi ($p=0,512$). Benzer şekilde mesh yerleştirme süresi de sol tarafta daha uzun olmasına rağmen istatistiksel olarak anlamlı bulunmadı ($p=0,239$). Ancak bilateral cerrahi grubunda aynı hastada sağ ve sol mesh yerleştirme süreleri karşılaştırıldığında, sol taraf süresi anlamlı derecede daha uzundu ($p<0,05$). Komplikasyon oranları düşüktü (%6,7). Altı hastada nüks tespit edildi.

Sonuç: Laparoskopik inguinal herni onarımı güvenli ve etkili bir yöntem olup, bilateral cerrahilerde süre belirgin olarak uzamaktadır. Sağ ve sol herni onarımı karşılaştırıldığında sol taraf cerrahisinde süreler daha uzun eğilim göstermekle birlikte farklar istatistiksel olarak anlamlı değildir. Cerrahi planlama ve hasta bilgilendirmelerinde bu parametreler göz önünde bulundurulmalıdır.

Anahtar Kelimeler: İnguinal herni, laparoskopi, mesh, onarım, komplikasyon

Address for Correspondence: Süleyman Büyükaşık, Asst. Prof., İstanbul Aydın University Faculty of Medicine, Department of General Surgery, İstanbul, Türkiye
E-mail: suleymanbuyukask@gmail.com **ORCID ID:** orcid.org/0000-0001-5536-4395

Cite as: Büyükaşık S, Altundal YE, Kankaya B, Eskioğlu E, Esen C. Our institutional experience in laparoscopic inguinal hernia surgery: is there a difference between right and left-sided procedures? A retrospective analysis of 192 cases. Med J Bakirkoy. 2025;21(4):392-397

Received: 22.08.2025
Accepted: 04.10.2025
Publication Date: 18.12.2025

INTRODUCTION

Abdominal wall hernias are among the most common surgical conditions, with inguinal hernias accounting for approximately 75% of cases. The lifetime risk is estimated at 27% for men and 3% for women (1). Over recent decades, significant progress has been made in their management. Although conventional open repairs, both mesh and non-mesh, were once predominant, laparoscopic mesh repairs have become increasingly preferred. This shift has been largely driven by comparative studies demonstrating advantages over open techniques (2).

Multiple studies have confirmed that minimally invasive approaches to inguinal hernia repair are associated with less postoperative pain, fewer wound complications, faster return to normal activities, and lower rates of chronic groin pain (2,3). Although recurrence rates appear similar between laparoscopic and open repairs, the overall benefits of laparoscopy have established it as a preferred approach (3).

The two principal laparoscopic techniques are transabdominal preperitoneal repair (TAPP) and totally extraperitoneal repair (TEP). TAPP provides a wider operative field but carries a greater risk of intra-abdominal injury, whereas TEP avoids peritoneal entry and thus reduces the risk of visceral injury. However, TEP is technically more demanding due to limited working space and restricted visualization (4,5).

A wide range of synthetic meshes is available for laparoscopic repair. Recently, anatomically contoured and three-dimensional (3D) meshes have been developed to overcome limitations such as folding or migration associated with flat polypropylene meshes (6). This evolution has contributed to the growing adoption of 3D mesh designs (6).

With the increasing use of laparoscopy for inguinal hernia repair, studies comparing open, laparoscopic, and robotic-assisted approaches have become more prevalent, further highlighting the advantages of minimally invasive surgery (1,3,7-10).

In the present study, we report our single-center experience with 192 patients who underwent laparoscopic TAPP repair.

METHODS

Between June 2021 and June 2025, the medical records of 303 patients who underwent surgery for inguinal hernia at the Department of General Surgery, İstanbul Aydın University Faculty of Medicine, were retrospectively reviewed. Fifty-nine patients treated with open repair were excluded. Surgical videos were available for 244 patients; however, 52 were excluded due to incomplete or missing video data. The remaining 192 patients, all of whom underwent laparoscopic TAPP repair, constituted the study cohort.

Polypropylene mesh was used in 6 patients, whereas 3D anatomically contoured mesh was used in the remaining 186 patients. All operations were performed by the same surgical team. Operative times were measured by reviewing surgical videos and operative notes: the interval from the insertion of the first trocar to the end of the procedure was defined as the total operative time, and the interval from the insertion of the mesh into the abdominal cavity to peritoneal closure was defined as the mesh placement time.

Postoperative outcomes assessed included hospital length of stay, drain duration (if applicable), early perioperative complications, and recurrence rates.

Statistical Analysis

Statistical analyses were performed using SPSS Statistics for Windows, version 15.0 (SPSS Inc., Armonk, NY, USA). Descriptive statistics were presented as mean±standard deviation for continuous variables and as percentages for categorical variables. The normality of data distribution was assessed using the Kolmogorov-Smirnov test. For categorical variables, either the chi-square test or Fisher's exact test was applied, as appropriate. For continuous variables, comparisons between groups were made using the Student's t-test or the Mann-Whitney U test, depending on data distribution. A two-tailed p-value <0.05 was considered statistically significant.

Artificial Intelligence-Assisted Figure Generation

Visualizations, including operative times and mesh-placement durations, were generated using the artificial intelligence tool ChatGPT (GPT-5; OpenAI; August 2025 version). Aggregated and anonymized data were provided to the system without any patient-identifiable information. The prompts specified figure types (e.g., bar charts, grouped comparisons), axis labeling, and legend formatting. All generated figures were carefully reviewed and cross-checked by the authors to ensure accuracy and consistency with the study data.

Ethical Approval

This retrospective study was conducted in accordance with the ethical standards of the institutional and national research committees and with the 1964 Declaration of Helsinki and its later amendments. The study protocol was approved by the İstanbul Aydın University Non-Interventional Clinical Research Ethics Committee (approval no: 164/2025, date: 06.08.2025). No interventions or changes were made to patient care as part of this research. As the study had a retrospective design, informed consent was not required from the patients.

RESULTS

Of the 192 patients included, 180 were male (93.8%) and 12 were female (6.3%). The mean age was 53.7±15.1 years (range, 21-83). Bilateral inguinal hernia repair was performed in 84 patients (43.8%), right-sided repair in 51 patients (26.6%), and left-sided repair in 57 patients (29.7%).

Drains were not placed in 168 patients (87.5%), whereas the surgical team inserted drains in 24 patients (12.5%). The mean length of hospital stay was 1.3 days overall. Patients with drains had a significantly longer hospital stay than those without drains (mean 1.79 vs. 1.29 days; $p<0.05$).

Complications were observed in 13 patients: seven were early perioperative complications and six were recurrences. Among early complications, four patients developed postoperative hematomas, one developed a bladder hematoma, and two sustained intraoperative bowel injuries; all bowel injuries were managed laparoscopically with drain placement. No drains were placed in patients with hematomas.

The mean operative time was 33.0±11.8 minutes (median 32.0) for right-sided repairs, minutes (median 35.0) for left-sided repairs, and 54.4±24.7 minutes (median 50.0) for bilateral repairs. Bilateral repairs were significantly longer than unilateral repairs ($p<0.05$; Figure 1). Although left-sided repairs averaged approximately 3 minutes longer than right-sided repairs, this difference was not statistically significant ($p=0.512$).

Mesh placement times were 6.7±3.3 minutes (median 6.0) for right-sided repairs and minutes (median 7.0) for left-sided repairs; the difference was not statistically significant ($p=0.239$; Figure 2). In bilateral procedures, no significant difference was found between the right and left sides in operative duration (26.6±13.6 min vs. 27.3±12.9 min, respectively; $p>0.05$). However, mesh placement time was significantly longer on the left side (6.6±3.1 min vs. 6.3±3.8 min; $p<0.05$) (Figure 3).

Recurrence was detected in 6 patients (3.1%) with a minimum follow-up of 6 months. This rate approached the upper limit reported in the literature (0.5-3%). When the 52 patients who had been excluded for lack of surgical video data were included in the denominator, the overall recurrence rate was 2.4%, consistent with previously published series.

Among patients with recurrence, 2 were repaired with polypropylene mesh and 4 patients with 3D mesh ($p<0.05$). Four recurrences occurred after bilateral repairs and two after left-sided repairs. No recurrences were observed in patients with right-sided unilateral repair. In the bilateral group, two recurrences occurred on the left and two on the right. In total, 6 recurrences occurred on the left side and 2 on the right (Table 1).

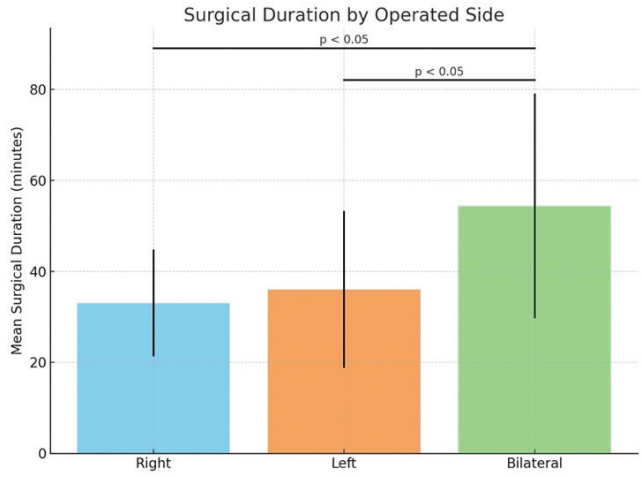


Figure 1. Mean surgical duration according to operated side (right, left, bilateral). Error bars indicate standard deviation. Significant differences are marked ($p<0.05$)

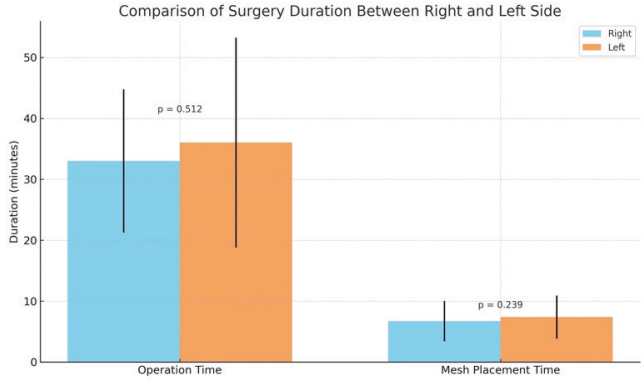


Figure 2. Comparison of operative time and mesh placement time between right- and left-sided inguinal hernia repairs. Error bars indicate standard deviation

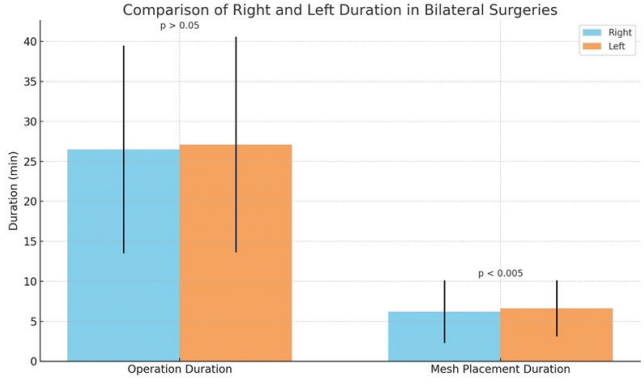


Figure 3. Comparison of operative time and mesh placement duration in bilateral surgeries (right and left sides). Error bars indicate standard deviation. Significant differences are marked ($p<0.05$)

Table 1. Clinical and demographic characteristics

Parameter	Value
Total number of patients	192
Gender (male/female)	180 (93.75%)/12 (6.25%)
Mean age (years)	**53.65±15.05 (21-83)**
Surgical side (right/left/bilateral)	51 (26.56%)/57 (29.69%)/84 (43.75%)
Patients with drain	24 (12.5%)
Mean hospital stay (days)	1.3
Hospital stay with/without drain	**1.79/1.29 (p<0.05)**
Number of complications (total)	13
Recurrence	6
Hematoma (early period)	4
Bladder hematoma	1
Intraoperative bowel injury	2
Operative time (right/left/bilateral)	33.02/36.02/54.37 min
Mesh placement time (right/left)	6.73/7.39 min
Mesh placement time in bilateral surgeries (right/left)	**6.31/6.63 min (p<0.05)**
Recurrence rate	6 patients (3.12 %)
Type of mesh in recurrences (prolene/3D)	2/4 (p<0.05)
Side of recurrence (right/left)	2/4

DISCUSSION

Laparoscopic approaches to inguinal hernia repair have increasingly replaced conventional open techniques and are now widely adopted (11). Nonetheless, utilization rates vary considerably, with reports of 38% in the United States (12) and 23% in the United Kingdom (13). Several studies have also suggested that laparoscopy is more commonly employed in bilateral hernia cases (14). In our series, 80.5% of inguinal hernia repairs over a four-year period were performed laparoscopically, and 43.8% of patients underwent bilateral repair. These rates are notably higher than global averages. Remarkably, in the past two years, open surgery was avoided altogether except in elderly patients deemed unsuitable for general anesthesia. This trend reflects an increasing preference for laparoscopic repair as experience accumulates.

Although recurrence rates after laparoscopic and open repair are generally comparable, rates of acute and chronic pain are significantly lower after laparoscopic repair (3,15). These advantages have driven wider adoption and training in laparoscopic techniques. In our cohort, the recurrence rate was 2.4%, consistent with published data and comparable to open repair (16,17).

Among laparoscopic techniques, debate continues between TAPP and TEP. While some studies have found equivalent outcomes, others favor one approach over the other. In our practice, the TAPP procedure is preferred because it provides wider exposure and facilitates more straightforward mesh placement, which is particularly useful during the learning phase (18-20).

Initially, large polypropylene meshes, typically used in open surgery, were employed to reduce costs. However, this practice was quickly abandoned due to longer operative times. In the present series, six patients received polypropylene mesh; two (33.3%) experienced recurrence. By contrast, among the 186 patients treated with 3D meshes, only four recurrences were observed (2.1%). Despite the small number of polypropylene cases, operative and mesh placement times were longer and recurrence rates were higher compared with 3D mesh; these differences were statistically significant (p<0.05). Prior reports have also indicated that 3D meshes reduce operative time, complication rates, and postoperative analgesia requirements; however, higher costs remain a limiting factor in some regions (21). Our findings support the clinical advantages of 3D mesh, despite the increased expense.

The learning curve for laparoscopic inguinal hernia repair has been extensively studied. A meta-analysis estimated that proficiency is achieved after approximately 32 cases; however, this threshold has decreased over the past 15 years owing to earlier training, improvements in laparoscopic equipment, and increased case exposure (2).

In our cohort, 13 complications occurred (6 recurrences, 4 hematomas, 2 bowel injuries, 1 bladder hematoma); nine of these complications occurred within the first 100 cases. As experience increased, complication rates declined significantly, underscoring the effect of cumulative surgical expertise.

Although the anatomy of the right and left inguinal regions is symmetrical, subtle technical challenges may still arise. In our study, left-sided repairs averaged three minutes longer than right-sided repairs, although this difference was not statistically significant (p=0.512). Similarly, mesh placement times were slightly longer on the left, although this difference was not statistically significant (p=0.213). Among bilateral cases, operative times were comparable between sides, but mesh placement times were significantly longer on the left ($p<0.05$). Interestingly, four of the six recurrences occurred on the left side, suggesting technical difficulty related to surgeons' right-hand dominance. Dissection and mesh deployment may be less intuitive on the left side, a phenomenon that deserves further investigation. Given the relatively large proportion of bilateral cases in our series, future studies with larger sample sizes may help confirm these observations. To our knowledge, this is the first study to examine side-specific outcomes in such detail, making it a unique contribution to the literature.

Study Limitations

This study has several limitations that should be considered. First, its retrospective design may have introduced selection bias, as data were derived from existing medical records rather than prospective collection. Second, although our series includes 192 cases, the sample size may be insufficient to detect subtle differences between right- and left-sided procedures with strong statistical power particularly regarding rare outcomes such as recurrence patterns. Lastly, the lack of long-term follow-up beyond the early postoperative period limits our ability to assess late complications, chronic pain, or recurrence rates. Future prospective, multicenter cohort studies with standardized protocols and extended follow-up are needed to validate and expand upon our findings.

CONCLUSION

Laparoscopic techniques for inguinal hernia repair are sometimes still perceived as technically demanding, even among experienced surgeons. However, growing surgical expertise consistently leads to reductions in operative time and complication rates. Although laparoscopic equipment and the use of advanced mesh types may increase overall costs, the advantages for patient outcomes, faster recovery, and surgeon comfort are indisputable. Moreover, our findings suggest that left-sided repairs may be associated with slightly longer operative and mesh placement times, potentially reflecting the predominance of right-handed surgeons. While these differences did not reach statistical significance, they raise an important technical consideration. To validate these observations and provide stronger evidence-based guidance, further randomized controlled trials with larger patient cohorts are warranted.

ETHICS

Ethics Committee Approval: The study protocol was approved by the İstanbul Aydın University Non-Interventional Clinical Research Ethics Committee (approval no: 164/2025, date: 06.08.2025).

Informed Consent: Retrospective study.

Acknowledgements

The authors would like to express their gratitude to the staff of the Department of General Surgery at İstanbul Aydın University Faculty of Medicine for their contributions to patient management and data collection.

FOOTNOTES

Authorship Contributions

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

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

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

REFERENCES

1. Pulikkal Reghunandanan R, Ali Usman A, Basheer S, Kuttichi L, Els Jojo J, Abdul Rasheed MF. Laparoscopic versus open inguinal hernia repair: a comparative study. Cureus. 2023;15:e48619.
2. Sivakumar J, Chen Q, Hii MW, Cullinan M, Choi J, Steven M, et al. Learning curve of laparoscopic inguinal hernia repair: systematic review, meta-analysis, and meta-regression. Surg Endosc. 2023;37:2453-75.
3. Haladu N, Alabi A, Brazzelli M, Imamura M, Ahmed I, Ramsay G, et al. Open versus laparoscopic repair of inguinal hernia: an overview of systematic reviews of randomised controlled trials. Surg Endosc. 2022;36:4685-700.
4. Xu LS, Li Q, Wang Y, Wang JW, Wang S, Wu CW, et al. Current status and progress of laparoscopic inguinal hernia repair: a review. Medicine (Baltimore). 2023;102:e34554.
5. Yildiz A. Laparoscopic transabdominal preperitoneal and totally extraperitoneal in inguinal hernia surgery: comparison of intraoperative and postoperative early complications of two techniques. J Minim Invasive Surg. 2022;25:18-23.

6. Faessen JL, Duijsings ESR, Boerma EG, Broos PPHL, van Vugt R, Stoot JHMB. The first experience with the Dextile anatomical mesh in laparoscopic inguinal hernia repair. Hernia. 2023;27:1203-8.
7. Yang C, Deng S. Laparoscopic versus open mesh repair for the treatment of recurrent inguinal hernia: a systematic review and meta-analysis. Ann Palliat Med. 2020;9:1164-73.
8. Kler A, Sekhon N, Antoniou GA, Satyadas T. Totally extra-peritoneal repair versus trans-abdominal pre-peritoneal repair for the laparoscopic surgical management of sportsman's hernia: a systematic review and meta-analysis. Surg Endosc. 2021;35:5399-413.
9. Solaini L, Cavaliere D, Avanzolini A, Rocco G, Ercolani G. Robotic versus laparoscopic inguinal hernia repair: an updated systematic review and meta-analysis. J Robot Surg. 2022;16:775-81.
10. Andresen K, Rosenberg J. Transabdominal pre-peritoneal (TAPP) versus totally extraperitoneal (TEP) laparoscopic techniques for inguinal hernia repair. Cochrane Database Syst Rev. 2024;7:CD004703.
11. Bittner R, Montgomery MA, Arregui E, Bansal V, Bingener J, Bisgaard T, et al.; International Endohernia Society. Update of guidelines on laparoscopic (TAPP) and endoscopic (TEP) treatment of inguinal hernia (International Endohernia Society). Surg Endosc. 2015;29:289-321. Erratum in: Surg Endosc. 2015;29:1655-6.
12. Madion M, Goldblatt MI, Gould JC, Higgins RM. Ten-year trends in minimally invasive hernia repair: a NSQIP database review. Surg Endosc. 2021;35:7200-8.
13. Palser TR, Swift S, Williams RN, Bowrey DJ, Beckingham IJ. Variation in outcomes and use of laparoscopy in elective inguinal hernia repair. BJS Open. 2019;3:466-75.
14. Hidalgo NJ, Guillaumes S, Bachero I, Holguín V, Momblán D. Trends and predictors of laparoscopic bilateral inguinal hernia repair in Spain: a population-based study. Surg Endosc. 2023;37:4784-94.
15. Bullen NL, Massey LH, Antoniou SA, Smart NJ, Fortelny RH. Open versus laparoscopic mesh repair of primary unilateral uncomplicated inguinal hernia: a systematic review with meta-analysis and trial sequential analysis. Hernia. 2019;23:461-72.
16. Burcharth J, Andresen K, Pommergaard HC, Bisgaard T, Rosenberg J. Recurrence patterns of direct and indirect inguinal hernias in a nationwide population in Denmark. Surgery. 2014;155:173-7.
17. Nikkolo C, Lillsaar T, Vaasna T, Kirsimägi Ü, Lepner U. Ten-year results of inguinal hernia open mesh repair. J Abdom Wall Surg. 2025;4:14384.
18. Goksoy B, Yilmaz G, Azamat IF, Ozata IH, Duman K. Laparoscopic inguinal hernia repair-TAPP versus TEP: results of 301 consecutive patients. Surg Technol Int. 2021;39:191-5.
19. Şahin N, Yıldız UM. Comparison of transabdominal preperitoneal (TAPP) and total extraperitoneal (TEP) methods in adult inguinal hernia repair at a secondary-level state hospital: a retrospective cohort study. Compreh Med. 2025;17:115-23.
20. Wei FX, Zhang YC, Han W, Zhang YL, Shao Y, Ni R. Transabdominal preperitoneal (TAPP) versus totally extraperitoneal (TEP) repair: a meta-analysis of randomized controlled trials. Surg Laparosc Endosc Percutan Tech. 2015;25:375-83.
21. Mir IS, Nafae AA, Malyar AA, Nafae M, Watali Y, Farooq M, et al. An experience of short-term results of laparoscopic inguinal hernioplasty using 3D mesh in a developing country Int J Clin Med. 2015;6:64-9.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):398-404
DOI: 10.4274/BMJ.galenos.2025.2025.7-11

Research

Evaluation of the Prevalence and Activity of Thyroid Ophthalmopathy in Patients with Thyroid Disease (Graves' Disease)

Tiroid Hastalarında (Graves Hastalığı) Tiroid Oftalmopati Sıklığı ve Aktivitesinin Değerlendirilmesi

Mehmet Özbaş, Sümeyye Selay Ersoy Şenel

University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Ophthalmology, İstanbul, Türkiye

ABSTRACT

Objective: To evaluate the prevalence, activity status, and clinical spectrum of thyroid associated ophthalmopathy (TAO) in patients with Graves' disease (GD), and to investigate the incidence of TAO activation and treatment approaches employed during follow-up.

Methods: This retrospective observational cohort study included 371 patients with GD who were referred from the endocrinology clinic to an the ophthalmology department and followed for at least 6 months. Thyroid function status at presentation (euthyroid, hypothyroid, or hyperthyroid), treatments for GD, thyroid-stimulating hormone receptor antibody (TRAb) levels, systemic comorbidities (diabetes mellitus, hypertension), and smoking status were recorded. Ocular involvement was evaluated using the clinical activity score (CAS), with CAS ≥3 indicating active TAO.

Results: Active TAO developed in 17.5% of GD patients. While most of these patients were in the hyperthyroid phase, a considerable proportion were in the euthyroid or hypothyroid phases. Patients with active TAO exhibited higher TRAb levels. Smoking emerged as a significant risk factor for TAO development. The most commonly administered treatment was intravenous methylprednisolone, followed by thyroidectomy, orbital radiotherapy, radioactive iodine therapy, and orbital decompression surgery.

Conclusion: Patients with GD, there is a risk of developing TAO during follow-up, even in those asymptomatic at presentation. Regular ophthalmologic assessment is recommended, particularly for patients who smoke or have elevated TRAb levels. CAS remains a valuable tool for evaluating disease activity and guiding treatment decisions.

Keywords: Graves' disease, thyroid associated ophthalmopathy, clinical activity score

ÖZ

Amaç: Graves hastalığı (GH) olan hastalarda tiroid ilişkili oftalmopatinin (TİO) prevalansını, aktivite durumunu ve klinik spektrumunu değerlendirmek ve takip süresince TİO aktivasyon insidansını ve uygulanan tedavi yaklaşımlarını araştırmaktır.

Gereç ve Yöntem: Bu retrospektif gözlemsel kohort çalışmasına, endokrinoloji kliniğinden göz hastalıkları bölümüne yönlendirilmiş ve en az 6 ay takip edilmiş 371 GH hastası dahil edildi. Başvuru anındaki tiroid fonksiyon durumu (ötiroidi, hipotiroidi veya hipertiroidi), GH tedavileri, tiroid uyarıcı hormon reseptör antikoru (TRAb) düzeyleri, sistemik komorbiditeler (diabetes mellitus, hipertansiyon) ve sigara içme durumu kaydedildi. Oküler tutulum klinik aktivite skoru (KAS) kullanılarak değerlendirildi; KAS ≥3 olan hastalar aktif TİO olarak kabul edildi.

Bulgular: GH hastalarının %17,5'inde aktif TİO gelişti. Bu hastaların çoğu hipertiroidi evresinde olsa da, önemli bir kısmı ötiroidi veya hipotiroidi evresindeydi. Aktif TİO'su olan hastalarda TRAb düzeyleri daha yüksekti. Sigara içimi, TİO gelişimi için anlamlı bir risk faktörü olarak belirlendi. En sık uygulanan tedavi damardan verilen metilprednizolon olurken; bunu tiroidektomi, orbital radyoterapi, radyoaktif iyot tedavisi ve orbital dekompresyon cerrahisi izledi.

Sonuç: GH hastalarında, başvuru sırasında asemptomatik olsalar bile takip süresince TİO gelişme riski bulunmaktadır. Özellikle sigara içen veya yüksek TRAb düzeylerine sahip hastalarda düzenli oftalmolojik değerlendirme önerilmektedir. KAS, hastalık aktivitesinin değerlendirilmesinde ve tedavi kararlarının yönlendirilmesinde değerli bir araç olmaya devam etmektedir.

Anahtar Kelimeler: Graves hastalığı, tiroid ilişkili oftalmopati, klinik aktivite skoru

Address for Correspondence: Mehmet Özbaş, MD, University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Ophthalmology, İstanbul, Türkiye
E-mail: dr.mehmetozbas@gmail.com **ORCID ID:** orcid.org/0000-0002-5679-3762

Cite as: Özbaş M, Ersoy Şenel SS. Evaluation of the prevalence and activity of thyroid ophthalmopathy in patients with thyroid disease (Graves' disease). Med J Bakirkoy. 2025;21(4):398-404

Received: 28.07.2025
Accepted: 02.09.2025
Publication Date: 18.12.2025

INTRODUCTION

Thyroid associated ophthalmopathy (TAO) is a significant autoimmune disorder that substantially impairs quality of life. Although rare, it is the most common inflammatory orbital disease and is associated with autoimmune thyroid dysfunction (1). This condition arises from an antibody response against a shared antigen present in both the thyroid gland and retro-orbital tissues (2). TAO almost always occurs in conjunction with circulating thyroid-stimulating hormone receptor antibodies (TRAb) (3).

The incidence of TAO is approximately 1.9 cases per 10,000 people annually (4). Although TAO is most commonly associated with hyperthyroidism secondary to Graves' disease (GD), about 10% of TAO patients are euthyroid or hypothyroid (5). The overall prevalence of TAO among patients with GD can be as high as 50% (6).

Although TAO is mild and non-progressive in most patients, close monitoring of high-risk individuals is crucial to ensure timely and appropriate treatment based on disease severity and activity. Moderate-to-severe forms, which necessitate aggressive therapy, constitute only 5-6% of all cases (7).

Risk factors identified for TAO include smoking, female sex, advanced age, genetic predisposition, wide lateral orbital wall angle, elevated TRAb levels, high pre-treatment triiodothyronine and thyroxine levels, uncontrolled hypo- or hyperthyroidism, and radioactive iodine (RAI) therapy (8).

Reports in the literature indicate that GD and associated TAO occur more frequently in women than in men. The prevalence of TAO increases with age, particularly between 40 and 60 years, and peaks during the fifth and sixth decades of life (9-11).

TAO is generally diagnosed clinically. However, in atypical or unilateral cases, orbital imaging plays an important role in the diagnosis, differential diagnosis, and clinical or surgical follow-up (12). TAO is diagnosed when at least two of the following three findings are present: thyroid dysfunction of autoimmune origin, one or more ocular signs, or radiological evidence of fusiform enlargement of the bellies of the extraocular muscles, sparing the tendinous insertions (13).

To assess TAO activity and to monitor treatment response, the clinical activity score (CAS), which reflects inflammatory changes, is commonly used (14).

The management of TAO is determined by disease severity. Mild cases typically require only symptomatic treatment, whereas moderate-to-severe TAO may necessitate multiple medical, surgical, or combined interventions. Close monitoring and appropriately timed treatments are essential; early initiation of medical therapy during the active phase of the disease is critical (14).

General management strategies for TAO encompass correcting hyperthyroidism, monitoring and treating hypothyroidism, and encouraging smoking cessation when appropriate (15).

Upon diagnosis, treatment should promptly target the active inflammatory phase of the disease to reduce severity, especially as chronic or fibrotic phases tend to be more resistant to therapy (13).

This study examined the demographic characteristics of GD patients, the presence of comorbidities (diabetes mellitus and hypertension), smoking status, the prevalence and activity of TAO, and the treatment modalities applied.

METHODS

This retrospective observational study included 371 patients diagnosed with GD and followed at the endocrinology clinic between January 2019 and September 2023, who were subsequently referred to the Ophthalmology Clinic of University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital. The study was conducted in accordance with the Declaration of Helsinki, and approval by the University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital Clinical Research Ethics Committee (approval no: 2023-21-26, date: 06.11.2023) and informed consent from all patients were obtained.

Patients' medical records were retrospectively reviewed. Only those with a minimum follow-up duration of six months were included. The mean follow-up period was 20.07±17.05 months (range: 6-57 months). Recorded data included the duration of ophthalmologic follow-up, current thyroid treatment, presence of comorbidities (diabetes mellitus and hypertension), TRAb levels, smoking status, ophthalmic examination findings, presence of TAO, and, when TAO was present, the treatment modalities applied.

TRAb levels were measured using a standardized immunoassay analyzer with a measurement range of 0.3-40 IU/L and a cut-off value of 1.75 IU/L. Values above this threshold were considered positive.

Assessment of TAO activity was performed according to the European Group on Graves' Orbitopathy guidelines, using the CAS. CAS is a seven-parameter scoring system that evaluates inflammatory activity in the eye. The parameters include spontaneous retrobulbar pain, pain on eye movement, eyelid erythema, conjunctival hyperemia, eyelid edema, conjunctival edema (chemosis), and inflammation of the caruncle or plica. The CAS is scored from 0 to 7 based

on the presence of these inflammatory signs (Table 1). Patients with a CAS score of 3 or higher were classified as having active TAO, those with scores of 1 or 2 as having inactive TAO, and patients with a CAS score of 0 as having no disease activation.

An experienced ophthalmologist performed the ophthalmologic examinations. Each patient underwent a comprehensive ophthalmic evaluation, including clinical assessment of CAS parameters, measurement of visual acuity, assessment of intraocular pressure, anterior segment examination with slit-lamp biomicroscopy, and fundus examination.

Thyroid function status at presentation (hypothyroid, hyperthyroid, or euthyroid) and smoking history were obtained from patients' medical records. Patients were categorized into three groups based on smoking status: current smokers, former smokers, and non-smokers.

Initially, 400 GD patients were screened for the study. However, patients with incomplete examination data or follow-up periods shorter than six months were excluded. Additionally, individuals with orbital diseases other than TAO were not included. Consequently, 371 patients with sufficient data and follow-up duration were enrolled in the study.

Statistical Analysis

Data were analyzed using IBM SPSS Statistics 25.0 (IBM Corp., Armonk, NY, USA). The distribution of continuous variables was assessed using the Shapiro-Wilk test. Since TRAb levels were not normally distributed, the Mann-Whitney U test was used for group comparisons. A p-value of <0.05 was considered statistically significant.

Table 1. Clinical activity score criteria

Clinical activity score criteria*	Description	Score if present
Spontaneous retrobulbar pain	Pain behind the eye at rest	1
Pain on eye movement	Pain caused by moving the eye	1
Eyelid erythema	Redness of the eyelids	1
Conjunctival injection (redness)	Redness of the conjunctiva	1
Eyelid edema	Swelling of the eyelids	1
Chemosis (conjunctival edema)	Swelling of the conjunctiva	1
Inflammation of caruncle or plica	Redness or swelling of caruncle or plica	1

*: Clinical activity score criteria, used to assess inflammatory activity in thyroid eye disease, are based on the guidelines of the European Group on Graves' orbitopathy. Each criterion is scored as 1 if present, with a total possible score ranging from 0 to 7. A clinical activity score ≥3 indicates active disease

RESULTS

The majority of the 371 Graves disease patients included in the study were female, with a mean age of 47 years. General data on patients' smoking status, thyroid function, and accompanying systemic diseases are summarized in Table 2.

Of the 371 patients with GD included in the study, 65 (17.5%) with a CAS ≥3 constituted the active TAO group. Among the remaining patients, 146 (39.3%) with a CAS score of 1 or 2 were classified as the inactive TAO group, while 160 patients (43.2%) with a CAS score of 0 were classified as having no clinical activity. The mean TRAb levels were highest in the active TAO group, followed by the inactive TAO group, and lowest in patients without clinical activity. Furthermore, TRAb levels in the active TAO group were significantly higher than in the other groups (Table 3).

The demographic and clinical characteristics of the 65 patients diagnosed with active TAO—including age, sex, smoking status, thyroid function at presentation, systemic

Table 2. General demographic and clinical characteristics of the entire cohort (n*=371)

Variable	Value
Female, n (%)	289 (77.9%)
Male, n (%)	82 (22.1%)
Age, mean±SD (range)	47.01±11.05 years (22-84 years)
Smoking status	
Current smoker, n (%)	158 (42.6%)
Former smoker, n (%)	36 (9.7%)
Never smoker, n (%)	177 (47.7%)
Thyroid status at referral	
Hyperthyroid, n (%)	170 (45.8%)
Euthyroid, n (%)	162 (43.7%)
Hypothyroid, n (%)	39 (10.5%)
Systemic comorbidities	
HT, n (%)	16 (4.3%)
DM, n (%)	37 (10.0%)
Both HT and DM, n (%)	8 (2.2%)
Mean TRAb level, IU/L	10.63±11.45 (median: 4.37, range: 0.3-40.0)
Mean follow-up duration, months	20.70±17.05 (median: 12.0, range: 6-57)

*: n indicates the number of patients; percentages are calculated based on the total cohort (n=371), unless otherwise specified
HT: Hypertension, DM: Diabetes mellitus, SD: Standard deviation, TRAb: Thyroid-stimulating hormone receptor antibody

comorbidities, and TRAb levels—are summarized in Table 4. In this group, 3 patients had previously undergone thyroidectomy, 4 had received RAI therapy, and the remaining 58 patients were receiving antithyroid medication.

The frequencies of the clinical findings comprising the CAS criteria in patients diagnosed with active TAO are presented in Table 5.

The distribution of symptomatic, medical, surgical, and adjuvant treatments applied to patients with TAO is presented in Table 6.

Table 3. TRAb levels and clinical activity distribution according to clinical activity score -based classification in the study cohort

Disease activity group	Mean TRAb (IU/L*)	Median TRAb (IU/L)	Min-max TRAb (IU/L)	Number of patients (n)	Percentage (%)
Active thyroid associated ophthalmopathy (CAS≥3)	14.20**±14.20	8.36	0.3-40.0	65	17.5
İnactive thyroid associated ophthalmopathy (CAS 1-2)	8.72±9.70	5.11	0.3-40.0	146	39.3
No clinical activity (CAS=0)	6.84±8.39	3.95	0.3-40.0	160	43.2
Total cohort	10.63±11.45	4.37	0.3-40.0	371	100

*: IU/L: International units per liter, **: TRAb levels were significantly higher in the active TAO group compared to the inactive TAO and no clinical activity groups (Mann-Whitney U test, p=0.0017; significance level p<0.05 was considered)
TRAb: Thyroid-stimulating hormone receptor antibody, CAS: Clinical activity score, TAO: Thyroid associated ophthalmopathy

Table 4. Demographic and clinical features of the active thyroid associated ophthalmopathy patients (n*=65)

Variable	Value
Female, n (%)	42 (64.6%)
Male, n (%)	23 (35.4%)
Age, mean±SD (range)	48.8±10.8 years (22-67 years)
Smoking status	
Current smoker, n (%)	35 (53.8%)
Former smoker, n (%)	6 (9.2%)
Never smoker, n (%)	24 (36.9%)
Thyroid status at referral	
Hyperthyroid, n (%)	39 (60.0%)
Euthyroid, n (%)	18 (27.7%)
Hypothyroid, n (%)	8 (12.3%)
Systemic comorbidities	
HT, n (%)	2 (3.1%)
DM, n (%)	5 (7.7%)
Both HT and DM, n (%)	2 (3.1%)
Mean TRAb level, IU/L	14.20±14.20 (median: 8.36, range: 0.3-40.0)
Mean follow-up duration, months	26.57±17.90 (median: 19.0, range: 6-57)

*: Indicates the number of patients; percentages are calculated based on the total cohort (n=65), unless otherwise specified
HT: Hypertension, DM: Diabetes mellitus, SD: Standard deviation, TRAb: Thyroid-stimulating hormone receptor antibody

Table 5. Number and frequency of symptoms according to clinical activity score criteria in active thyroid associated ophthalmopathy patients

Clinical activity score criterion	Number of patients (n*)	Frequency (%)
Spontaneous retrobulbar pain	34	52.3%
Pain on eye movement	22	33.8%
Eyelid edema (swelling)	52	80.0%
Eyelid erythema (redness)	10	15.4%
Conjunctival injection (redness)	47	72.3%
Chemosis (conjunctival swelling)	36	55.4%
Caruncle or plica inflammation	23	35.4%

Since patients may exhibit multiple symptoms simultaneously, the total number of symptoms may exceed the number of patients. *: Number of patients in whom the symptom was observed

Table 6. Eye treatment approaches in active thyroid associated ophthalmopathy patients (n*=65)

Treatment	n* (%)
Symptomatic/lubricant treatment	65 (100%)
Intravenous methylprednisolone	40 (61.5%)
Thyroidectomy	9 (13.8%)
Orbital radiotherapy	6 (9.2%)
Radioactive iodine	4 (6.2%)
Orbital decompression surgery	1 (1.5%)
Only symptomatic treatment**	5 (7.7%)

*: Number of patients, **: Due to mild clinical findings, patient preferences, or the presence of systemic contraindications, symptomatic treatment was administered

DISCUSSION

This study retrospectively evaluated the prevalence, clinical features, and treatment approaches of TAO in patients with GD. According to our findings, TAO developed in 211 of 371 patients diagnosed with GD (56.9%). Among these patients with TAO, 146 (69.2%) were classified as having inactive TAO and 65 (30.8%) as having active TAO based on the CAS. Although the prevalence of TAO in our study appears higher than the 25-50% range reported in the literature (6,16,17), this can be explained by our cohort comprising patients referred to a tertiary care center. Therefore, the obtained rate can be considered acceptable within the study context. The rate of active TAO was within the previously reported range of 15-30% (17,18). These findings indicate that a significant proportion of patients with GD develop thyroid eye disease.

Approximately 60% of patients with active TAO in our study were in the hyperthyroid phase. However, a substantial proportion were euthyroid (28%), while a smaller proportion were hypothyroid (12%). This aligns with existing literature showing that TAO is not limited to the thyrotoxic phase but can also develop during euthyroid or hypothyroid states (19). In clinical practice, normal thyroid function does not exclude the risk of TAO, underscoring the need for vigilance, particularly in euthyroid patients.

Female sex has been reported in the literature as a risk factor for TAO (10). Similarly, in our study, the majority of patients who developed orbital involvement were female (64.6%).

We observed that TRAb levels increased with clinical activity in TAO patients. The mean TRAb level in the active TAO group was significantly higher than in the inactive and no-clinical-activity groups (p=0.0017) (Table 3). Similarly, the literature reports a significant association between elevated TRAb levels and active TAO. TRAb serves as a specific biomarker for TAO and is a valuable tool for the accurate management of this complex disease. Its measurement is important and clinically useful for the diagnosis, differential diagnosis, and monitoring of TAO and GD (20,21). Circulating TRAb have been correlated with clinical activity and severity of TAO (22). These data support TRAb as a reliable indicator of disease activity and suggest that it can be used alongside clinical assessment to guide treatment decisions.

According to the clinical findings comprising the CAS criteria (Table 5), the most frequently observed signs among active patients were eyelid edema (80%), conjunctival injection (72.3%), and chemosis (55.4%). These findings represent the main components of orbital inflammation and play a significant role in CAS scoring (23). In addition, the high prevalence of symptoms, such as spontaneous retrobulbar pain (52.3%) and pain on eye movement (33.8%), suggests that inflammation may adversely affect patients' quality of life. Inflammation of the caruncle or plica (35.4%) and eyelid erythema (15.4%) were other ocular signs observed in our patients and are also included in the CAS criteria. Our findings are consistent with the existing literature and reflect the clinical manifestations of the inflammatory phase of TAO (24,25).

Among the 65 patients with active TAO in our study, 53.8% (n=35) were current smokers, 9.2% (n=6) were former smokers, and 36.9% (n=24) had never smoked. This finding aligns with the literature, which consistently suggests that smoking increases the risk of TAO. Smoking has been identified as an independent risk factor for both the development and progression of TAO (26). Previous research has shown that smoking may exacerbate inflammation by stimulating orbital fibroblasts and activating hypoxia-inducible factors. Based on these findings, smoking cessation counseling should be considered a key component in the management of TAO (27).

Some of our patients also presented with systemic comorbidities such as diabetes mellitus and hypertension. Although the relationship between these comorbidities and TAO activity could not be analyzed in this study, their presence remains clinically relevant for systemic evaluation and potential treatment planning for patients.

All 65 patients diagnosed with active TAO received symptomatic treatment consisting of artificial tears and gel for ocular surface relief. Regarding systemic therapies (Table 6), intravenous methylprednisolone was the most commonly administered treatment (61.5%), supporting corticosteroids as the primary treatment modality for active TAO (19,28,29). Additionally, based on clinical and systemic factors, the following treatment approaches were administered: thyroidectomy in 9 patients, orbital radiotherapy in 6 patients, and RAI therapy in 4 patients. One patient underwent orbital decompression surgery due to marked proptosis, signs of orbital compression, and resistance to medical treatment.

Among five patients with active TAO, only symptomatic treatment was provided. This was attributed to mild clinical findings, patient preferences, or systemic contraindications. According to the literature, supportive therapy combined with close monitoring is considered appropriate for patients with active TAO and mild symptoms (20).

This therapeutic diversity highlights the need for individualized management of TAO based on disease activity, severity, and patient characteristics and demonstrates that clinical practice is shaped accordingly (30).

Except for the patient who underwent decompression surgery, ophthalmologic assessments during follow-up—including visual acuity, intraocular pressure, anterior segment examinations, and fundus examinations—showed no significant changes among the remaining patients with active orbitopathy. Overall, their ophthalmologic findings remained stable throughout the follow-up period.

Study Limitations

The primary limitation of our study is the absence of data regarding disease duration since the initial diagnosis of GD, which precluded assessment of its impact on ocular involvement. Additionally, the inability to evaluate the relationship between systemic comorbidities (hypertension and diabetes mellitus) and disease activity constitutes an additional limitation.

CONCLUSION

In conclusion, this study highlights the clinical spectrum of TAO in patients with GD and underscores the necessity of individualized treatment approaches. Our findings demonstrate that active TAO affects a significant proportion of patients and can occur not only during the hyperthyroid phase but also during the euthyroid and hypothyroid phases. The observed correlation between increasing TRAb levels and disease activity supports the use of this biomarker as an adjunct to clinical assessment. Furthermore, smoking was identified as an important risk factor, and individualized treatment strategies were found to be effective in achieving clinical stabilization. Early recognition of inflammatory signs and implementation of appropriate treatment strategies through a multidisciplinary approach are crucial for preventing serious TAO-related complications in patients with GD.

ETHICS

Ethics Committee Approval: This study approval by the University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital Clinical Research Ethics Committee (approval no: 2023-21-26, date: 06.11.2023).

Informed Consent: Informed consent from all patients were obtained.

Acknowledgments

We sincerely thank Dr. Özgür Örüm, MD, for his valuable contributions to the conduct of this study.

FOOTNOTES

Authorship Contributions

Surgical and Medical Practices: M.Ö., S.S.E.Ş., Concept: M.Ö., Design: M.Ö., Data Collection or Processing: M.Ö., S.S.E.Ş., Analysis or Interpretation: M.Ö., S.S.E.Ş., Literature Search: M.Ö., S.S.E.Ş., Writing: M.Ö.

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

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

REFERENCES

1. Szelog J, Swanson H, Sniegowski MC, Lyon DB. Thyroid eye disease. Mo Med. 2022;119:343-50.
2. Li Z, Cestari DM, Fortin E. Thyroid eye disease: what is new to know? Curr Opin Ophthalmol. 2018;29:528-34.
3. Kahaly GJ, Bartalena L, Hegedüs L, Leenhardt L, Poppe K, Pearce SH. 2018 European Thyroid Association guideline for the management of Graves' hyperthyroidism. Eur Thyroid J. 2018;7:167-86.
4. Smith TJ, Hegedüs L. Graves' disease. N Engl J Med. 2016;375:1552-65.
5. Smith TJ, Janssen J. Insulin-like Growth factor-I receptor and thyroid-associated ophthalmopathy. Endocr Rev. 2019;40:236-67.
6. Chin YH, Ng CH, Lee MH, Koh JWH, Kiew J, Yang SP, et al. Prevalence of thyroid eye disease in Graves' disease: a meta-analysis and systematic review. Clin Endocrinol (Oxf). 2020;93:363-74.
7. Bartalena L, Gallo D, Tanda ML, Kahaly GJ. Thyroid eye disease: epidemiology, natural history, and risk factors. Ophthalmic Plast Reconstr Surg. 2023;39:2-8.
8. Lombardi A, Menconi F, Greenberg D, Concepcion E, Leo M, Rocchi R, et al. Dissecting the genetic susceptibility to Graves' disease in a cohort of patients of Italian origin. Front Endocrinol (Lausanne). 2016;7:21.
9. Bartley GB. The epidemiologic characteristics and clinical course of ophthalmopathy associated with autoimmune thyroid disease in Olmsted County, Minnesota. Trans Am Ophthalmol Soc. 1994;92:477-588.
10. Laurberg P, Berman DC, Bülow Pedersen I, Andersen S, Carlé A. Incidence and clinical presentation of moderate to severe graves' orbitopathy in a Danish population before and after iodine fortification of salt. J Clin Endocrinol Metab. 2012;97:2325-32.
11. Tanda ML, Piantanida E, Liparulo L, Veronesi G, Lai A, Sassi L, et al. Prevalence and natural history of Graves' orbitopathy in a large series of patients with newly diagnosed graves' hyperthyroidism seen at a single center. J Clin Endocrinol Metab. 2013;98:1443-9.
12. Gonçalves AC, Gebrim EM, Monteiro ML. Imaging studies for diagnosing Graves' orbitopathy and dysthyroid optic neuropathy. Clinics (Sao Paulo). 2012;67:1327-34.
13. Hoang TD, Stocker DJ, Chou EL, Burch HB. 2022 update on clinical management of graves disease and thyroid eye disease. Endocrinol Metab Clin North Am. 2022;51:287-304.

14. Savku E, Gündüz K. Diagnosis, follow-up and treatment results in thyroid ophthalmopathy. Turk J Ophthalmol. 2015;45:156-63.
15. Khong JJ, Finch S, De Silva C, Rylander S, Craig JE, Selva D, et al. Risk factors for Graves' orbitopathy; the australian thyroid-associated orbitopathy research (ATOR) study. J Clin Endocrinol Metab. 2016;101:2711-20.
16. Hiromatsu Y, Eguchi H, Tani J, Kasaoka M, Teshima Y. Graves' ophthalmopathy: epidemiology and natural history. Intern Med. 2014;53:353-60.
17. Piantanida E, Tanda ML, Lai A, Sassi L, Bartalena L. Prevalence and natural history of Graves' orbitopathy in the XXI century. J Endocrinol Invest. 2013;36:444-9.
18. Subramaniam DL, Yadalla D, Rajagopalan J. Are severe forms of thyroid eye disease common in the Indian population? Clinical characteristics of 136 patients from a tertiary eye care centre. Indian J Endocrinol Metab. 2023;27:56-61.
19. Bartalena L, Kahaly GJ, Baldeschi L, Dayan CM, Eckstein A, Marcocci C, et al. The 2021 European Group on Graves' orbitopathy (EUGOGO) clinical practice guidelines for the medical management of Graves' orbitopathy. Eur J Endocrinol. 2021;185:G43-67.
20. Sarić-Matutinović M, Diana T, Nedeljković-Beleslin B, Ćirić J, Žarković M, Perović-Blagojević I, et al. Sensitivity of three thyrotropin receptor antibody assays in thyroid-associated orbitopathy. J Med Biochem. 2022;41:211-20.
21. George A, Diana T, Längericht J, Kahaly GJ. Stimulatory thyrotropin receptor antibodies are a biomarker for Graves' orbitopathy. Front Endocrinol (Lausanne). 2021;11:629925.
22. Vannucchi G, Campi I, Bonomi M, Covelli D, Dazzi D, Currò N, et al. Rituximab treatment in patients with active Graves' orbitopathy: effects on proinflammatory and humoral immune reactions. Clin Exp Immunol. 2010;161:436-43.
23. Mourits MP, Prummel MF, Wiersinga WM, Koornneef L. Clinical activity score as a guide in the management of patients with Graves' ophthalmopathy. Clin Endocrinol (Oxf). 1997;47:9-14.
24. Shah SS, Patel BC. Thyroid eye disease. [Updated 2023 May 22]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan. Available from: https://www.ncbi.nlm.nih.gov/books/NBK582134/
25. Scarabosio A, Surico PL, Singh RB, Tereshenko V, Musa M, D'Esposito F, et al. Thyroid eye disease: advancements in orbital and ocular pathology management. J Pers Med. 2024;14:776.
26. Stan MN, Bahn RS. Risk factors for development or deterioration of Graves' ophthalmopathy. Thyroid. 2010;20:777-83.
27. Pfeilschifter J, Ziegler R. Smoking and endocrine ophthalmopathy: impact of smoking severity and current vs lifetime cigarette consumption. Clin Endocrinol (Oxf). 1996;45:477-81.
28. Marcocci C, Marinò M. Treatment of mild, moderate-to-severe and very severe Graves' orbitopathy. Best Pract Res Clin Endocrinol Metab. 2012;26:325-37.
29. Längericht J, Krämer I, Kahaly GJ. Glucocorticoids in Graves' orbitopathy: mechanisms of action and clinical application. Ther Adv Endocrinol Metab. 2020;11:2042018820958335.
30. Bartalena L, Tanda ML. Current concepts regarding Graves' orbitopathy. J Intern Med. 2022;292:692-716.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):405-413
DOI: 10.4274/BMJ.galenos.2025.2024.4-1

Research

Comparison of PCR with Conventional Methods in the Diagnosis of *Pneumocystis jirovecii* and Atypical *Pneumonia* Agent *Chlamydia pneumoniae* and *Mycoplasma pneumoniae* in Intensive Care Patients

Yoğun Bakım Hastalarında *Pneumocystis jirovecii* ve Atipik Pnömoni Etkeni *Chlamydia pneumoniae* ve *Mycoplasma pneumoniae*'nin Tanısında PCR'nin Konvansiyonel Yöntemlerle Karşılaştırılması

Yener Özel[1], Nurhan Sarıoğlu[2], Gülhan Vardar Ünlü[1], Fuat Erel[2], Fatih Ugün[3], Mehmet Ünlü[1]

[1]Balıkesir University Faculty of Medicine, Department of Medical Microbiology, Balıkesir, Türkiye

[2]Balıkesir University Faculty of Medicine, Department of Chest Disease, Balıkesir, Türkiye

[3]Balıkesir University Faculty of Medicine, Department of Anesthesiology and Reamination, Balıkesir, Türkiye

ABSTRACT

Objective: *Pneumocystis jirovecii (P. jirovecii)* can colonize healthy people and other chronic lung patients as well as in immunocompromised or immunosuppressed patients. *Mycoplasma pneumoniae* (*M. pneumoniae*) and *Chlamydia pneumoniae* (*C. pneumoniae*) are the most common atypical *pneumonia* agents in the community. This study aimed to investigate the presence of *P. jirovecii*, *M. pneumoniae*, and *C. pneumoniae* colonization/infection by different methods in patients in intensive care units.

Methods: In this study, the presence of *P. jirovecii*, *M. pneumoniae* and *C. pneumoniae* in endotracheal aspirate specimens taken from 50 intensive care patients was investigated by polymerase chain reaction (PCR) method and conventional methods.

Results: Twenty (40%) of 50 patients were diagnosed with *pneumonia*. *P. jirovecii* DNA was detected in six specimens. *P. jirovecii* cysts were detected by Gram-Weigert and Gomori silver staining in one of the patients who were PCR positive. Of 20 patients diagnosed with *pneumonia*, 9 (45%) were *C. pneumoniae* PCR positive, 18 (90%) were *C. pneumoniae* immunoglobulin M (IgM) positive, 8 (40%) were *C. pneumoniae* IgG positive, and 5 (25%) were *M. pneumoniae* IgG positive. Of 30 patients (60%) who were not diagnosed with *pneumonia*, 1 (3.3%) were *C. pneumoniae* IgM positive, 12 (40%) were *C. pneumoniae* IgG positive, and 5 (16.7%) were *M. pneumoniae* IgG positive.

Conclusion: It is important to use the PCR method together with conventional methods for rapid and accurate diagnosis of *P. jirovecii*, *C. pneumoniae*, and *M. pneumoniae* in patients hospitalized in intensive care units, to differentiate colonization/infection, and to prevent the risk of infection development in immune-compromised patients.

Keywords: Atypical *pneumoniae*, *C. pneumoniae*, intensive care unit, *M. pneumoniae*, *P. jirovecii*

ÖZ

Amaç: *Pneumocystis jirovecii (P. jirovecii)* sağlıklı insanlarda ve diğer kronik akciğer hastalarında olduğu gibi immün sistemi baskılanmış hastalarda da kolonize olabilir. *Mycoplasma pneumoniae (M. pneumoniae) ve Chlamydia pneumoniae (C. pneumoniae)* toplumda en sık görülen atipik pnömoni etkenleridir. Bu çalışmada yoğun bakım ünitelerinde yatan hastalarda *P. jirovecii, M. pneumoniae* ve *C. pneumoniae* kolonizasyon/enfeksiyon varlığının farklı yöntemlerle araştırılması amaçlanmıştır.

Gereç ve Yöntem: Bu çalışmada 50 yoğun bakım hastasından alınan endotrakeal aspirat örneklerinde *P. jirovecii, M. pneumoniae* ve *C. pneumoniae* varlığı polimeraz zincir reaksiyonu (PCR) yöntemi ve geleneksel yöntemlerle araştırılmıştır.

Bulgular: Elli hastanın 20'sine (%40) pnömoni tanısı konulmuştur. Altı örnekte *P. jirovecii* DNA'sı tespit edilmiştir. PCR pozitifliği saptanan hastaların sadece birinde Gram-Weigert ve Gomori silver boyama ile *P. jirovecii* kisti tespit edilmiştir. Pnömoni tanısı alan 20 hastanın 9'unda (%45) *C.*

Address for Correspondence: Yener Özel, MD, Balıkesir University Faculty of Medicine, Department of Medical Microbiology, Balıkesir, Türkiye
E-mail: yener_ozel@hotmail.com **ORCID ID:** orcid.org/0000-0001-6618-8251

Cite as: Özel Y, Sarıoğlu N, Vardar Ünlü G, Erel F, Ugün F, Ünlü M. Comparison of PCR with conventional methods in the diagnosis of *Pneumocystis jirovecii* and atypical *pneumonia* agent *Chlamydia pneumoniae* and *Mycoplasma pneumoniae* in intensive care patients. Med J Bakirkoy. 2025;21(4):405-413

Received: 01.04.2024
Accepted: 11.02.2025
Publication Date: 18.12.2025

ÖZ

pneumoniae PCR pozitif, 18'inde (%90) *C. pneumoniae* immünoglobulin M (IgM) pozitif, 8'inde (%40) *C. pneumoniae* IgG pozitif ve 5'inde ise (%25) *M. pneumoniae* IgG pozitif olarak saptanmıştır. Pnömoni tanısı konamayan 30 (%60) hastanın 1'inde (%3,3) *C. pneumoniae* IgM pozitif, 12'sinde (%40) *C. pneumoniae* IgG pozitif ve 5'inde (%16,7) *M. pneumoniae* IgG pozitif olarak bulunmuştur.

Sonuç: Yoğun bakım ünitelerinde yatan hastalarda *P. jirovecii, C. pneumoniae* ve *M. pneumoniae*'nin hızlı ve doğru tanısı, kolonizasyon/enfeksiyon ayrımı ve bağışıklık sistemi baskılanmış hastalarda enfeksiyon riskinin önlenmesi için PCR yönteminin konvansiyonel yöntemlerle birlikte kullanılması önemlidir.

Anahtar Kelimeler: Atipik pnömoni, *C. pneumoniae,* yoğun bakım birimi, *M. pneumoniae, P. jirovecii*

INTRODUCTION

Pneumocystis spp. are opportunistic organisms that are unicellular, eukaryotic, and can be found in the lungs of many mammals. Five species of *Pneumocystis* have been identified, namely *Pneumocystis carinii* and *Pneumocystis wakefieldiae* in rats, *Pneumocystis murina* in mice, *Pneumocystis oryctolagi* in rabbits, and *Pneumocystis jirovecii* (*P. jirovecii*) in humans (1). *P. jirovecii* causes *Pneumocystis pneumonia* (PCP), which has a dangerous clinical status, especially in people with compromised immune systems. Initially considered a protozoan, *Pneumocystis* was recognized as a fungus in the late 1980s with the development of gene sequencing techniques (2). Microscopic identification of organisms in lower respiratory tract specimens is considered the gold standard for PCP. Cyst forms of *Pneumocystis* can be stained with Grocott-Gomori metamine silver and trophozoite forms with Giemsa and Diff-Quick, but experience with microscopy is important in the diagnostic examination. Polymerase chain reaction (PCR) is increasingly used in the molecular diagnosis of PCP (3).

Chlamydia pneumoniae (*C. pneumoniae*) is an obligate intracellular bacterium that infects humans and is one of the major causes of atypical pneumonia. *C. pneumoniae* was first isolated from the eye of a child during a trachoma vaccine study in Taiwan in 1965. *C. pneumoniae* pneumonia causes upper and lower respiratory tract symptoms within 1-3 weeks and can occur as an acute or subacute infection (4). Serological diagnosis of *C. pneumoniae* infection is based on a fourfold or greater increase in immunoglobulin M (IgM) or IgG antibodies. Micro-immunofluorescence (IFA) testing is the most preferred method and is species-specific (5).

Mycoplasma pneumoniae (*M. pneumoniae*) is a bacterium that is slow-growing, pleomorphic, non-motile, and without a cell wall. *M. pneumoniae* is easily transmitted, and the infection is most likely to be spread by droplets (6). The most common symptoms include sore throat, hoarseness, fever, cough; initially unproductive but later moderate bloodless sputum, headache, chills, myalgia, earache, and general malaise. A four-fold or greater increase in both IgM and IgG antibody titers in specimens collected at 2-3 week intervals suggests an existing or recent infection (7). Many commercial kits are available based on IFA testing and enzyme immunoassay methods. PCR is another preferred method for the detection of *M. pneumoniae* (8).

Pneumonia is responsible for a significant portion of morbidity and mortality in intensive care units (ICUs). Since atypical pneumonia agents are difficult to grow in culture, data on the frequency of these infectious agents are limited (9). In recent years, advances in molecular microbiology techniques have led to an increased recognition of the importance of atypical pneumonias. However, there are few studies on the etiology of atypical pneumonia in Türkiye (10).

This study aimed to compare the sensitivity of conventional staining methods with PCR for the diagnosis of *P. jirovecii* in ICUs, and to investigate the relationship between the presence of *M. pneumoniae/C. pneumoniae* DNA and serum IgM/IgG antibodies.

METHODS

Ethical Approval

The study was conducted with the permission of Balıkesir University Clinical Research Ethics Committee with the approval number 2018/94, date 09.05.2018. Patients and their first-degree relatives were informed about the content of the study and their consent was obtained.

Collection of Specimens

Fifty patients who were monitored in the ICU of Balıkesir University Research Hospital were included in the study. Patients admitted to the ICU with and without symptoms of pneumonia were included in the study. Endotracheal aspirate (ETA) was collected for the detection of *P. jirovecii, C. pneumoniae,* and *M. pneumoniae,* and serum samples were collected for the serological diagnosis of *C. pneumoniae* and *M. pneumoniae.*

Preparation of Specimens

ETA specimens were centrifuged at 13000 rpm for 10 min. The supernatant was discarded and the pellet was stored as aliquots for staining methods and PCR tests. The pellet was spread on slides by the cytocentrifugation method for Giemsa, Gram-Weigert, and Gomori-Metamine silver staining. For PCR, at least 200 µL of each sample was taken, put into sterile eppendorf tubes, and stored at -20 °C. Blood specimens were centrifuged at 4000 rpm for 10 min. The serum specimens obtained were transferred to sterile tubes and stored at -20 °C.

Giemsa Staining Method

Giemsa solution was prepared by mixing the stock Giemsa (Merck, Darmstadt, Germany) with PBS adjusted to pH 7.2 at a ratio of 1/11. Specimen slides were fixed with methanol for 3 min and stained for 25 min with Giemsa solution. The slides were then rinsed with low-flow tap water from the edges and left upright to dry. The preparations were examined for the presence of *P. jirovecii* trophozoite using immersion oil at 100x magnification under a light microscope.

Gram-Weigert Staining Method

Gram-Weigert dye solution was prepared according to the protocol (11) by preparing crystal violet (Merck, Darmstadt, Germany), Gram iodine (Merck, Darmstadt, Germany), and aniline-xylene solutions (Merck, Darmstadt, Germany) separately. After cytocentrifugation, the slides were dried at room temperature and immersed in 1% eosin Y stain (Merck, Darmstadt, Germany) for 5 minutes. It was then washed in distilled water for 2 min. They were kept in crystal violet and Gram iodine solution for five minutes and then in distilled water for two minutes. The slides were gently dipped in the aniline-xylene solution to remove excess dye and dried at room temperature. They were examined with immersion oil at 100x magnification for the presence of *P. jirovecii* trophozoites and cysts.

Gomori-Methamine Silver Staining Method

A commercial dye (GBL, İstanbul, Türkiye) was used for silver staining. Staining was performed according to the protocols prepared by the manufacturer. The slides were then closed with entellan (Merck, Darmstadt, Germany) and examined for *P. jirovecii* cysts using immersion oil at 100x magnification.

Positive control slides were used to validate the staining methods applied in the study. Additionally, each stained slide that was deemed suspicious was reviewed and verified by at least three researchers who are experts in the field.

DNA Isolation

Specimens stored at -20 °C for PCR were thawed at room temperature and vortexed. The thawed specimens were centrifuged at 13,000 rpm for 10 min, and the supernatant was discarded. 200 µL of animal tissue lysis buffer and 20 µL of proteinase K solution (Qiagen, Hilden, Germany) were added to the pellet and vortexed. The specimens incubated at 56 °C for 60 min and then at 95 °C for 10 min in the heat block were allowed to reach room temperature. DNA isolation was performed using the Qiagen-EZ1 Mini kit and the EZ1 Advanced XL (Qiagen, Hilden, Germany) device in accordance with the manufacturer's instructions.

Real Time PCR

Presence of *P. jirovecii, C. pneumoniae* and *M. pneumoniae* in ETA specimens was examined using a commercial kit (Fast-Track, Esch-sur-Alzette, Luxembourg) and a real-time PCR device (Qiagen Rotor-Gene Q, Hilden, Germany). The amplification protocol was performed according to the manufacturer's instructions. The commercially available *P. jirovecii* and *M. pneumoniae/C. pneumoniae* PCR kit contains a primer/probe mixture, a positive control, a negative control, an internal control (IC), an enzyme mixture, and a PCR buffer solution. Ribonucleic acid was transcribed into cDNA using a specific primer-mediated reverse transcription step, which was immediately followed, in the same tube, by PCR. An increase in fluorescence observed from the relevant dual-labeled probe indicated the presence of specific pathogen sequences in the reaction, which was reported as a cycle threshold value by the real-time thermocycler. The assay used equine arteritis virus (EAV) as both an extraction control and IC. EAV was introduced by the laboratory into each sample and the negative control at the lysis buffer stage of the extraction process. The kit contents were prepared in volumes suitable for the number of samples provided in Table 1.

Real-time PCR procedures were performed according to the following protocol: hold at 42 °C for 15 minutes, hold at 94 °C for 3 minutes, followed by 40 cycles of 94 °C for 8 seconds and 60 °C for 34 seconds. The analysis results were interpreted in cases where the negative control was negative, and both the positive and ICs were positive. For detailed validation data such as sensitivity, specificity, clinical studies and external quality panel results, please refer to the related validation file at: www.fast-trackdiagnostics.com

Immunofluorescent Antibody Method

C. pneumoniae and *M. pneumoniae* IgG and IgM antibodies in serum specimens were investigated by the IFA method using a commercial kit (Euroimmun, Lübeck,

Table 1. Required reagent volumes for PCR reaction based on the number of samples

Number of reactions	1	15	32	64
PPmix	1.5 µL	22.5 µL	48 µL	96 µL
Buffer	12.5 µL	187.5 µL	400 µL	800 µL
Enzyme	1 µL	15 µL	32 µL	64 µL
Total	15 µL	225 µL	480 µL	960 µL

PPmix: Primer/prop mixture, PCR: Polymerase chain reaction

Germany). IgG-type antibodies in the patient's serum were removed with Eurosorb (Euroimmun, Lübeck, Germany) reagent before the detection of IgM-type antibodies. The presence of IgG and IgM type antibodies was examined under a fluorescence microscope at 40x magnification by comparison with positive and negative controls.

The evaluation of IFA images was performed by comparing the positive and negative control images provided in the kit with the clinical sample images. The results were reviewed and validated by at least three researchers who are experts in the field.

Statistical Analysis

Data were analyzed using IBM SPSS 25.0 statistical software (IBM SPSS Inc., Chicago, IL, USA). Continuous variables are expressed with the average±standard deviation, and minimum and maximum values. The differences between categorical variables are examined by chi-square analysis and Fisher's exact test.

RESULTS

Distribution of Patient Specimens by Units

Specimens included in the study and their distribution according to ICUs are shown in Table 2.

Diagnosis and Clinical Findings of the Patients

Twenty of 50 patients were diagnosed with pneumonia. The most common findings in all patients were cough and shortness of breath.

Table 2. Distribution of ETA samples by intensive care units

Intensive care	Specimen
Anesthesia intensive care	26
General intensive care	14
Coroner intensive care	5
CVS intensive care	5
Total	50

ETA: Endotracheal aspirate, CVS: Cardiovasculer surgery

P. Jirovecii PCR Results in ETA Specimen

P. jirovecii DNA was detected in six specimens; the DNA load was high in one patient, while it was found to be low in the other five.

P. Jirovecii Staining Results of ETA Specimen

P. jirovecii trophozoite forms were not detected in any of the specimens using Giemsa staining. *P. jirovecii* cysts were detected by Gram-Weigert and Gomori Silver staining in only one of the patients who were positive by PCR. In the other five PCR-positive patients, *P. jirovecii* cyst or trophozoite form were not observed by staining methods. PCR-negative patients, were also negative by staining methods. Five of the PCR-positive patients were followed up in the anesthesia ICU, and one was in the cardiovascular ICU. While two of the six patients with a positive PCR test for PCP had symptoms of pneumonia, the patient with the highest parasite load had no symptoms of pneumonia. The distribution of *P. jirovecii* positive patients is given in Table 3 according to the diagnostic methods and pneumoniae findings. PCR and Gram Weigert and Gomori silver staining images of the patient with the highest parasite load are shown in Figure 1.

Table 3. Distribution of *P. jirovecii* positive patients by specimen type, method and diagnosis

No	Giemsa staining	GW staining	Silver staining	PCR	Pneumonia
1	-	+	+	+	-
2	-	-	-	+	+
3	-	-	-	+	-
4	-	-	-	+	-
5	-	-	-	+	+
6	-	-	-	+	-

GW: Gram-Weigert staining, PCR: Polymerase chain reaction

C. Pneumoniae/M. Pneumoniae PCR and Antibody Results of the Patients With and Without Pneumonia

Of the 50 patients included in the study, 20 (40%) were diagnosed with pneumonia. Of 20 patients diagnosed with pneumonia, 9 (45%) were *C. pneumoniae* PCR positive, 18 (90%) *C. pneumoniae* IgM positive, 8 (40%) *C. pneumoniae* IgG positive, and 5 (25%) *M. pneumoniae* IgG positive. Of 30 patients (%60) who were not diagnosed with *pneumonia*, 1 (3.3%) *C. pneumoniae* IgM positive, 12 (40%) *C. pneumoniae* IgG positive, and 5 (16.7%) *M. pneumoniae* IgG positive. While *C. pneumonia* PCR and IgM results were found to be statistically significant in the group diagnosed with pneumonia compared to the other group ($p<0.05$), for *C. pneumoniae* IgG there was no significant correlation between the groups ($p>0.05$). The data for *M. pneumonia* are not sufficient for statistical evaluation in terms of PCR, IgM, and IgG positivity in both groups. PCR and immunofluorescent antibody results of patients with and without pneumonia are shown in Table 4. *C. pneumoniae/M. pneumoniae* amplification curve and IgG/IgM type antibodies images are shown in Figures 2 and 3.

C. Pneumoniae Antibody and PCR Results of the Patients With/Without *Pneumonia*

Clinical findings such as respiratory distress, dyspnea, cough, and sputum were observed in the majority of patients with *C. pneumoniae* detected by DNA and IgM antibody presence. The comparisons of *C. pneumoniae* antibody and PCR results are shown in Table 5.

M. Pneumoniae Antibody and PCR Results of the Patients With/Without *Pneumonia*

There was only one patient who tested positive for *M. pneumoniae* PCR and had chronic obstructive pulmonary disease (COPD), cerebrovascular disease, and signs of unconsciousness. IgG antibody was positive in 10 patients, while *M. pneumoniae* IgM antibody was negative in all patients. The comparisons of *M. pneumoniae* antibody and PCR results is shown in in Table 6.

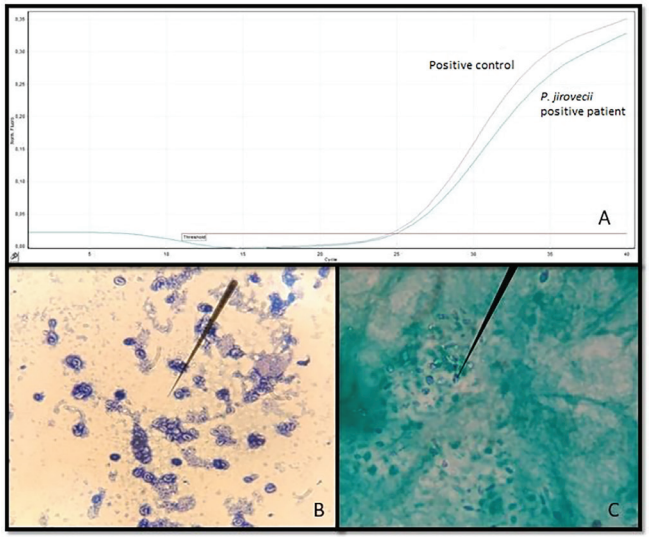


Figure 1. *P. jirovecii* amplification curve and staining images. **A)** *P. jirovecii* amplification curve in RT-PCR method, **B)** *P. jirovecii* cysts in Gram-Weigert staining, **C)** *P. jirovecii* cysts in Gomori-metamine silver staining
RT-PCR: Reverse transcription-polymerase chain reaction

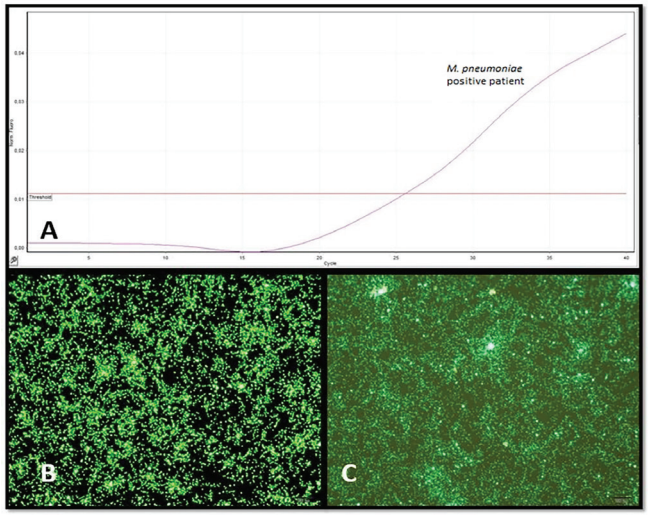


Figure 2. *M. pneumoniae* amplification curve and IFA images **A.** RT-PCR amplification curve, **B.** IgM type antibodies detected by IFA method (positive control), **C.** IgG type antibodies detected by IFA method
IFA: Immunofluorescence, RT-PCR: Reverse transcription-polymerase chain reaction, Ig: Immunoglobulin

Table 4. *C. pneumoniae/M. pneumoniae* PCR and antibody results of the patients with and without pneumonia

Atypical *pneumonia* agents	Pneumonia (n=20)		Non-pneumonia (n=30)		p-value
	n	%	n	%	
C. pneumoniae PCR	9	45	0	0	p<0.05
C. pneumoniae IgM	18	90	1	3.3	p<0.05
C. pneumoniae IgG	8	40	12	40	p>0.05
M. pneumoniae PCR	0	0	1	3.3	-
M. pneumoniae IgM	0	0	0	0	-
M. pneumoniae IgG	5	25	5	16.7	-

PCR: Polymerase chain reaction, Ig: Immunoglobulin

Table 5. Comparison of the *C. pneumoniae* antibody and PCR results of the patients with pneumonia

No	IgM	IgG	PCR	Pneumonia
1	+	-	+	+
2	-	+	-	-
3	+	-	-	+
4	+	-	+	+
5	-	+	-	-
6	+	-	-	+
7	+	-	-	+
8	-	+	-	-
9	+	-	-	+
10	-	+	-	-
11	-	+	-	-
12	-	+	-	-
13	+	-	-	+
14	+	+	+	+
15	-	+	-	-
16	+	-	-	+
17	-	+	-	-
18	+	-	+	+
19	+	+	+	+
20	+	-	+	+
21	+	+	+	+
22	-	+	-	-
23	-	+	-	-
24	+	+	+	+
25	-	+	-	+
26	+	+	-	-
27	+	-	-	+
28	+	-	-	+
29	-	+	-	-
30	+	+	-	+
31	-	+	-	+
32	+	+	+	+

PCR: Polymerase chain reaction, Ig: Immunoglobulin

Table 6. Comparison of the *M. pneumoniae* antibody and PCR results of the patients with pneumonia

No	IgM	IgG	PCR	Pneumonia
1	-	+	-	+
2	-	+	-	+
3	-	+	-	+
4	-	+	-	-
5	-	+	-	-
6	-	+	-	-
7	-	+	-	-
8	-	+	-	+
9	-	+	-	-
10	-	-	+	-
11	-	+	-	+

PCR: Polymerase chain reaction, Ig: Immunoglobulin

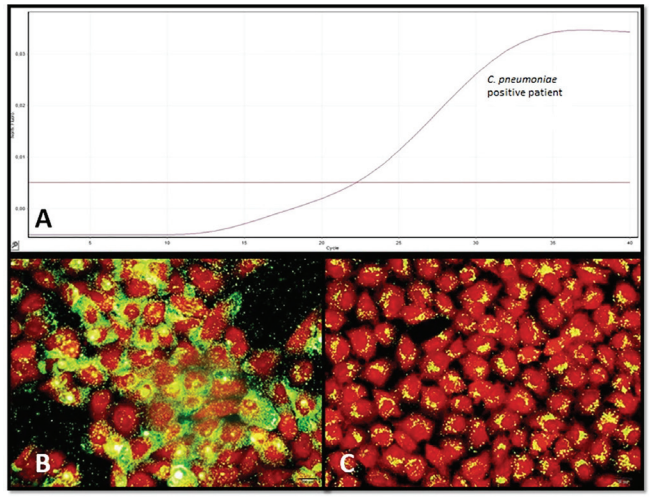


Figure 3. *C. pneumoniae* amplification curve and IFA images. **A.** RT-PCR amplification curve, **B.** IgM type antibodies detected by IFA method, **C.** IgG type antibodies detected by IFA method
IFA: Immunofluorescence, RT-PCR: Reverse transcription-polymerase chain reaction, Ig: Immunoglobulin

DISCUSSION

Community-acquired pneumonia (CAP) is a lung disease that threatens human health and has a high mortality rate, which is often associated with long hospital stays (12). More than two million children under five die from pneumonia each year (13). Detection of *P. jirovecii* DNA in patients without any symptoms has been defined as "*P. jirovecii* colonization/carrier", despite the risk of PCP (14). However, due to the low organismal load associated with colonization, the frequency of detection of colonization varies greatly from one study to another depending on the methods used, such as PCR or immunohistochemical staining of respiratory samples. *P. jirovecii* DNA has also been detected in individuals who are not immunocompromised, such as those with chronic lung disease, smokers, or pregnant women (15). In a study by Mori et al. (16), it was reported that PCP developed within 2-4 weeks in patients who had *P. jirovecii* DNA detected by PCR, did not have respiratory symptoms, and received immunosuppressive therapy for rheumatoid arthritis. Therefore, it has been reported that the detection of *P. jirovecii* DNA requires prophylaxis in conditions of clinical immune suppression that carry a risk of PCP risk (17).

In a study by Seyhan (18), the presence of *P. jirovecii* was investigated by comparing the PCR with various staining methods (Giemsa, Gram-weigert, and Gomori silver staining) in 100 immunocompromised patients. In that study, *P. jirovecii* DNA was detected by PCR in 4 of 100 patients, although *P. jirovecii* cysts and trophozoites could not be detected by Giemsa, Gram-Weigert, and Gomori silver staining methods in these patients. *P. jirovecii* cysts were observed with Gomori silver staining in a patient whose *P. jirovecii* DNA was not detected by PCR in the same study. In a study by Töz et al. (19), bronchoalveolar lavage (BAL) specimens of 42 patients with a pre-diagnosis of PCP were investigated using Giemsa, Gram-Weigert staining, and PCR. In this study, 13 patients were found to be positive by the PCR method, while three patients were found to be positive by both the staining and PCR methods, and, a total of 16 out of 42 patients were found to be positive for PCP. In a study conducted in İzmir, *P. jirovecii* positivity was found in 21 of 30 patients without pneumonia findings by PCR, six by direct fluorescent antibody (DFA), and one by staining. Four of these patients showed positivity with both PCR and DFA; furthermore, it was stated that PCR is a good method for demonstrating colonization (20).

In this study, two of the six patients with positive *P. jirovecii* DNA, followed up in the ICU, had pneumonia. In the patient who was PCR positive and had a high DNA load, *P. jirovecii* cysts were detected only in Gram-Weigert and Gomori methenamine silver staining, and there were no signs of pneumonia in this patient. This patient was thought to have *P. jirovecii* colonization. *P. jirovecii* was not detected in the three staining methods, which were applied to the samples from the other five patients with low DNA load. According to these data, the conclusion was that the probability of detecting *P. jirovecii* by staining methods is low. For this reason, it is thought that PCR will be suitable for investigating patients who may be at risk for PCP contamination, especially in ICUs, and will contribute to the identification of colonized patients.

In addition to cell culture and PCR, serological methods are also used in the diagnosis of *C. pneumoniae* infection. The serological diagnosis of *C. pneumoniae* infection is based on the detection of a 4-fold increase in IgG or IgA levels in serum specimens taken during the acute and convalescent periods. The retrospective nature of the diagnosis in this method means that serological results have little effect on treatment decisions. Specific IgM antibodies begin to form in 2-3 weeks, and IgG antibodies begin to form in 6-8 weeks. Therefore, it seems that serological methods are not effective in the diagnosis of acute infection (21). For this reason, PCR is thought to be the most sensitive and effective method in lower respiratory tract specimens.

Twenty of the 50 patients included in the study, (40%) had signs of pneumonia. *C. pneumoniae* PCR positivity was detected in 9 patients among those diagnosed with pneumonia. Of the 20 patients in the same group, 18 were positive for *C. pneumoniae* IgM and 8 were positive for *C. pneumoniae* IgG. In nine patients, *C. pneumoniae* DNA and IgM antibodies were found to be positive together. Similarly, in the thesis study conducted by Gökçınar (22), the lower respiratory tract and serum specimens, of 50 patients were investigated for *C. pneumoniae* using PCR and serological methods. Although IgM antibodies were detected in two of 50 patients in the related study, *C. pneumoniae* DNA was not detected in any of the patients by PCR. In this study, IgM antibody was positive in 18 of 20 patients (90%) with pneumonia symptoms. In 8 of these patients (40%), IgM was found to be positive alone, and in 9 (45%), it was positive with PCR. In the group without pneumonia symptoms, only IgM was found positive in one patient. IgM antibodies detected by the IFA method were observed to be associated with the presence of pneumonia symptoms. In this context, it is thought that studying IgM and PCR tests together, may be useful in indicating acute infection. In the patient group with pneumonia symptoms, IgM was found positive in 18 patients. However, PCR positivity was detected less frequently than expected. The potential reason for this result may be that ETA samples were used instead of BAL samples in the study. BAL samples were not preferred in this study because they require an invasive process, which is a limitation of our study.

Among the 15 *Mycoplasma* species that can infect humans, *M. pneumoniae* is the best-known atypical pneumonia agent (23). While IgM antibodies were not detected in any of the patient sera included in this study, IgG antibodies were detected in 10 of them. *M. pneumoniae* DNA was detected in one of these patients, and IgM and IgG antibodies were not detected in the same patient. *Mycoplasma* infection is frequently seen in children, but it has been reported that it may rarely be associated with certain diseases in the elderly. Many studies have shown that *M. pneumoniae* infection is associated with myocarditis, pericarditis, cerebral stroke, and vasculitis (24-26). In addition to playing a role in the pathogenesis of atherosclerosis, *M. pneumoniae* also induces cardiovascular disease and chronic inflammation (27). In this study, PCR was positive only in a 95-year-old patient who was followed up in the ICU, with the diagnosis of COPD and cerebrovascular disease.

Zhang et al. (28) extensively compared PCR and serological methods for the diagnosis of *M. pneumoniae* in a meta-analysis. Researchers have reported that commercial PCR tests exhibit high specificity but lower and more variable sensitivity. Despite their advantages, they still cannot replace serology, and PCR is a reliable and accurate method to be used together with serological diagnosis. Copete et al. (29) compared the presence of *M. pneumoniae* in children with and without CAP by serology and PCR; 13.9% of children with CAP were found positive by PCR and/or serology. While 10.3% of them were positive by PCR and 6.7% serologically, 2.8% of the cases were positive with both tests. In the acute phase, 32% of children with CAP and 38.3% of healthy children had positive IgM titers. It has been reported that IgM titer alone cannot be useful and PCR tests alone cannot distinguish between infection/colonization; therefore, it would be more reliable to use both tests together.

Study Limitations

The most significant limitation of the study is the small sample size. We believe that conducting similar studies with much larger sample sizes, within the context of financial resources and laboratory infrastructure, would be highly valuable. The data we have obtained can provide valuable preliminary information for larger-scale studies.

In this study, both PCR and serological methods were utilized. While these methods provide high sensitivity and specificity, they also have certain limitations. PCR, despite being a highly sensitive technique, is heavily reliant on sample quality and the preservation of genetic material. On the other hand, serological tests are valuable for detecting past or current infections but may have reduced sensitivity in the early stages of infection when antibody production is insufficient.

These limitations highlight the necessity of using complementary diagnostic tests, particularly in cases where a definitive diagnosis cannot be achieved with a single method. In this context, the combined use of PCR and serological methods can contribute to achieving more accurate and reliable diagnostic outcomes. Future studies should focus on evaluating the effectiveness of combining different diagnostic methods and further refining diagnostic protocols to enhance their utility.

CONCLUSION

As a result, the use of PCR together with conventional methods for the rapid and accurate diagnosis of *P. jirovecii* (PCP) and atypical pneumonia agents (*C. pneumoniae* and *M. pneumoniae*) in patients hospitalized in ICUs, guides the diagnosis. It is important because people with colonization act as reservoirs for spreading the agent within the community, and they also carry the risk of developing pneumonia if they have immune difficulties.

ETHICS

Ethics Committee Approval: The study was conducted with the permission of Balıkesir University Clinical Research Ethics Committee with the approval number 2018/94, date 09.05.2018.

Informed Consent: Patients and their first-degree relatives were informed about the content of the study and their consent was obtained.

FOOTNOTES

Authorship Contributions

Concept: Y.Ö., N.S., M.Ü., Design: Y.Ö., N.S., M.Ü., Data Collection or Processing: Y.Ö., N.S., F.E., F.U., Analysis or Interpretation: Y.Ö., G.V.Ü., F.E., F.U., M.Ü., Literature Search: Y.Ö., N.S., G.V.Ü., F.E., F.U., M.Ü., Writing: Y.Ö., M.Ü.

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

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

REFERENCES

1. Aliouat-Denis CM, Chabé M, Demanche C, Aliouat el M, Viscogliosi E, Guillot J, et al. Pneumocystis species, coevolution and pathogenic power. Infect Genet Evol. 2008;8:708-26.
2. Edman JC, Kovacs JA, Masur H, Santi DV, Elwood HJ, Sogin ML. Ribosomal RNA sequence shows pneumocystis carinii to be a member of the fungi. Nature. 1988;334:519-22.
3. Fan LC, Lu HW, Cheng KB, Li HP, Xu JF. Evaluation of PCR in bronchoalveolar lavage fluid for diagnosis of pneumocystis jirovecii pneumonia: a bivariate meta-analysis and systematic review. PLoS One. 2013;8:e73099.
4. Hahn DL. Chlamydia pneumoniae and chronic asthma: updated systematic review and meta-analysis of population attributable risk. PLoS One. 2021;16:e0250034.
5. Hagemann JB, Simnacher U, Marschall MT, Maile J, Soutschek E, Wellinghausen N, et al. Analysis of humoral immune responses to recombinant chlamydia pneumoniae antigens. Int J Infect Dis. 2020;91:232-9.
6. Esposito S, Argentiero A, Gramegna A, Principi N. Mycoplasma pneumoniae: a pathogen with unsolved therapeutic problems. Expert Opin Pharmacother. 2021;22:1193-202.
7. Lanao AE, Chakraborty RK, Pearson-Shaver AL. Mycoplasma Infections. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021.
8. Kurutepe S, Ecemiş T, Ozgen A, Biçmen C, Celik P, Aktoğu Özkan S, et al. Toplum kökenli pnömonisi olan erişkin hastalarda konvansiyonel ve multipleks PCR yöntemleriyle bakteriyel etiyolojinin araştırılması. Mikrobiyol Bul. 2012;46:523-31.
9. Nguyen C, Kaku S, Tutera D, Kuschner WG, Barr J. Viral respiratory infections of adults in the intensive care unit. J Intensive Care Med. 2016;1:427-41.

10. Kaygusuz S, Koksal I, Aydin K, Caylan R. Investigation of atypical bacteria and virus antigens in respiratory tract infections by use of an immunofluorescence method. Jpn J Infect Dis. 2004;57:33-6.
11. Özensoy Töz S. Aspirasyon, vücut sıvıları ve idrar incelemeleri. In: Korkmaz M, Ok ÜZ, editors. Parazitolojide laboratuvar. İzmir: Türkiye Parazitoloji Derneği; 2011. p. 58-9.
12. Arnold FW, Summersgill JT, Lajoie AS, Peyrani P, Marrie TJ, Rossi P, et al. Community-acquired pneumonia organization (CAPO) investigators. A worldwide perspective of atypical pathogens in community-acquired pneumonia. Am J Respir Crit Care Med. 2007;175:1086-93.
13. Wardlaw TM, Johansson EW, Hodge M. Pneumonia: the forgotten killer of children. New York: UNICEF; 2006. 40 s.
14. Xue T, Ma Z, Liu F, Du W, He L, Wang J, et al. Pneumocystis jirovecii colonization and its association with pulmonary diseases: a multicenter study based on a modified loop-mediated isothermal amplification assay. BMC Pulm Med. 2020;20:70.
15. Morris A, Norris KA. Colonization by pneumocystis jirovecii and its role in disease. Clin Microbiol Rev. 2012;25:297-317.
16. Mori S, Cho I, Sugimoto M. A followup study of asymptomatic carriers of pneumocystis jiroveci during immunosuppressive therapy for rheumatoid arthritis. J Rheumatol. 2009;36:1600-5.
17. Cordonnier C, Cesaro S, Maschmeyer G, Einsele H, Donnelly JP, Alanio A, et al. Pneumocystis jirovecii pneumonia: still a concern in patients with haematological malignancies and stem cell transplant recipients. J Antimicrob Chemother. 2016;71:2379-85.
18. Seyhan T. "Pneumocystis jirovecii'nin kantitatif PCR yöntemi ile tanımlanması ve diğer tanı yöntemleri ile karşılaştırılması." Selçuk Üniversitesi, Yüksek Lisans Tezi, Konya, Türkiye; 2015. Erişim: https://hdl.handle.net/20.500.12395/11827
19. Töz S, Gündüz C, Tetik A, Taşbakan M, Pullukçu H, Bacakoğlu F, et al. Pneumocystis jirovecii pnömonisi tanısında mikroskopi ve gerçek zamanlı polimeraz zincir reaksiyonu yöntemlerinin karşılaştırılması: klinik bulgular ile yorumlanması [The comparison of microscopy and real time polymerase chain reaction methods for the diagnosis of Pneumocystis Jirovecii pneumonia: evaluation of clinical parameters]. Tuberk Toraks. 2017;65:220-6.
20. Özkoç S, Bayram Delibaş S, Erbaycu AE, Ergüden C, Akısü Ç. Akciğer hastalığı olan hastalarda pneumocystis jirovecii kolonizasyonunun araştırılması [Investigation of pneumocystis jirovecii colonization in patients with pulmonary diseases]. Turkiye Parazitol Derg. 2014;38:214-9.
21. Puolakkainen M. Laboratory diagnosis of persistent human chlamydial infection. Front Cell Infect Microbiol. 2013;3:99.
22. Gökçınar B. Chlamydophila pneumoniae'ın alt solunum yolu enfeksiyonu görülen hastalarda serolojik ve moleküler yöntemlerle saptanması. Yüksek lisans tezi, Hacettepe Üniversitesi, Ankara, Türkiye; 2010.
23. American Academy of Pediatrics. Mycoplasma pneumoniae infections. In: Pickering LK, Baker CJ, Long SS, McMillan JA, editors. Red Book: 2006 Report of the Committee on Infectious Diseases. 27th ed. Elk Grove Village, IL: American Academy of Pediatrics; 2006. p. 468-70.
24. Chen SC, Tsai CC, Nouri S. Carditis associated with Mycoplasma pneumoniae infection. Am J Dis Child. 1986;140:471-2.
25. Farraj RS, McCully RB, Oh JK, Smith TF. Mycoplasma-associated pericarditis, case report. Mayo Clin Proc. 1997;72:33-6.
26. Perez C, Mendoza H, Hernandez R, Valcayo A, Guarch R. Leukocytoclastic vasculitis and polyarthritis associated with mycoplasma pneumoniae infection. Clin Infect Dis. 1997;25:154-5.
27. Taylor-Robinson D, Thomas BJ. Chlamydia pneumoniae in arteries: the facts, their interpretation, and future studies. J Clin Pathol. 1998;51:793-7.
28. Zhang L, Zong ZY, Liu YB, Ye H, Lv XJ. PCR versus serology for diagnosing mycoplasma pneumoniae infection: a systematic review & meta-analysis. Indian J Med Res. 2011;134:270-80.
29. Copete AR, Vera C, Herrera M, Aguilar YA, Rueda ZV, Vélez LA. Mycoplasma pneumoniae in children with and without community-acquired pneumonia. What do PCR and serology say? Pediatr Infect Dis J. 2020;39:e104-8.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):414-418
DOI: 10.4274/BMJ.galenos.2025.2025.8-14

Research

An Inexpensive Way for Ovarian Stimulation with Only Clomiphene Citrate in Women with Diminished Ovarian Reserve

Azalmış Over Rezervi Olan Kadınlarda Overyan Stimülasyonu Amacıyla Yalnızca Klomifen Sitrat Kullanımının Ekonomik Bir Yaklaşım Olarak Değerlendirilmesi

Ramazan Özyurt[1], Sebile Güler Çekiç[2], Aret Kamar[3], Ozan Özolcay[3], Mustafa Barış Ata[4]

[1]University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Türkiye
[2]Koç University Hospital, Department of Obstetrics and Gynecology, İstanbul, Türkiye
[3]İstanbul IVF Center, Assisted Reproductive Technology Unit, İstanbul, Türkiye
[4]Koç University Faculty of Medicine, Department of Obstetrics and Gynecology, İstanbul, Türkiye

ABSTRACT

Objective: Clomiphene citrate (CC) is commonly used for various indications related to ovarian stimulation in assisted reproductive technology. There is limited information on the effectiveness of clomiphene for ovarian stimulation when used alone in women with diminished ovarian reserve (DOR).

Methods: This retrospective analysis included women stimulated with CC only; a total of 840 cycles from 381 couples were analyzed.

Results: The median age was 39 years. All women had an antral follicle count <5, and the median anti-Müllerian hormone was 0.34 ng/mL. CC was given for 5 days, and stimulation lasted for 8 days. Among 840 cycles, 344 (%) cycles were cancelled. The reasons for cancellation were: premature ovulation, 119 (14.2%); no follicular growth, 98 (11.7%); no oocyte yield at oocyte pick-up, 66 (7.9%); and no metaphase II oocyte yield, 61 (7.3%). The number of cycles with no embryos was 2 (0.4%). The remaining 494 cycles (58.8%) proceeded to embryo transfer (ET). The median numbers of oocytes collected, mature oocytes, and 2 pronuclei embryos were each 2. Frozen ET was performed in 443 (89.1%) cycles, and fresh ET was performed in 51 (10.3%) cycles. The pregnancy test was positive in 131 cycles; there were 96 live births (one twin birth), and the live birth rate per ET was 19.5%.

Conclusion: CC alone seems capable of preventing a premature luteinizing hormone surge in the majority of women with severely DOR. CC-only stimulation is an effective and inexpensive approach to ovarian stimulation in women with severely DOR.

Keywords: Ovarian stimulation, clomiphene citrate, decreased ovarian reserve

ÖZ

Amaç: Klomifen sitrat (CC), yardımcı üreme teknolojilerinde overyan stimülasyon için çeşitli endikasyonlarla yaygın olarak kullanılan bir ajandır. Bununla birlikte, düşük over rezervine (DOR) sahip kadınlarda yalnızca CC kullanılarak yapılan overyan stimülasyonun etkinliği hakkında sınırlı bilgi bulunmaktadır.

Gereç ve Yöntem: Yalnızca CC ile stimüle edilen kadınlar retrospektif olarak analiz edilmiştir. Toplamda 381 çifte ait 840 siklus dahil edilmiştir.

Bulgular: Dahil edilen kadınların median yaşı 39, antral folikül sayısı <5 ve anti-Müllerian hormon değeri 0,34 ng/mL idi. CC beş gün süreyle uygulanmış ve toplam stimülasyon süresi 8 gün olarak kaydedilmiştir. Toplam 840 siklusun 344'ü (%41) iptal edilmiştir. İptal nedenleri şunlardır: prematür ovulasyon: 119 (%14,2); foliküler gelişim olmaması: 98 (%11,7); oosit toplama işlemi sırasında oosit elde edilememesi: 66 (%7,9); olgun (metafaz II) oosit elde edilememesi: 61 (%7,3). Embriyo oluşmayan siklus sayısı yalnızca 2 (%0,4) idi. Geriye kalan 494 siklus (%58,8) embriyo transferine (ET) ilerlemiştir. Toplanan oositlerin medyan sayısı 2, metafaz II oosit sayısı 2 ve 2 pronükleus embriyo sayısı da 2 idi. ET yapılan 494 siklusun 443'ünde (%89,1) donmuş transfer, 51'inde (%10,3) taze transfer uygulanmıştır. Toplam 131 siklusta gebelik testi pozitif bulunmuş ve 96 canlı doğum (biri ikiz) gerçekleşmiştir. ET başına canlı doğum oranı %19,5 olarak hesaplanmıştır.

Address for Correspondence: Sebile Güler Çekiç, MD, Koç University Hospital, Department of Obstetrics and Gynecology, İstanbul, Türkiye
E-mail: sebileguler@gmail.com **ORCID ID:** orcid.org/0000-0002-1632-1170

Cite as: Özyurt R, Güler Çekiç S, Kamar A, Özolcay O, Ata MB. An inexpensive way for ovarian stimulation with only clomiphene citrate in women with diminished ovarian reserve. Med J Bakirkoy. 2025;21(4):414-418

Received: 03.09.2025
Accepted: 27.09.2025
Publication Date: 18.12.2025

ÖZ

Sonuç: Yalnızca CC kullanımı, şiddetli DOR olan kadınların büyük çoğunluğunda erken luteinizan hormon yükselmesini önlemede etkili görünmektedir. Bu yaklaşım, etkili ve düşük maliyetli bir overyan stimülasyon yöntemi olarak değerlendirilebilir.

Anahtar Kelimeler: Overyan stimulasyon, klomifen sitrat, düşük over rezervi

INTRODUCTION

Ovarian stimulation (OS) is employed to induce multifollicular growth for in vitro fertilization (IVF) treatments. As the number of oocytes retrieved increases, the probability of live birth increases. Diminished ovarian reserve (DOR) limits the response to OS. Women with DOR constitute 9-24% of infertile patients (1,2). They have increased follicle-stimulating hormone (FSH) level and decreased anti-Müllerian hormone (AMH) levels, and their antral follicle count (AFC) is 5-7 or fewer follicles (3,4). In DOR, multifollicular growth during OS is not possible (5). Since ovarian response to stimulation treatments is inefficient, clinicians have experimented with various protocols to achieve satisfactory results.

For mild stimulation, alternative treatments include: i) low-dose gonadotropins; ii) oral treatments such as letrozole or clomiphene citrate (CC) (6). With mild stimulation, the cost of treatment and the number of drug injections are reduced, ultimately rendering treatment more accessible for these patients who may need several treatment cycles.

CC is a selective estrogen receptor modulator (7). It acts as an antiestrogen at normal estrogen levels. CC blocks the estrogen receptors in the hypothalamus, thereby inhibiting endogenous estrogen's negative-feedback effect, which results in an increased pulse frequency of gonadotropin-releasing hormone (GnRH) secretion and a subsequent rise in FSH and luteinizing hormone (LH) (8). This eventually stimulates follicular growth and ovulation. The advantages of CC used in OS during IVF include reduced gonadotropin administration, lower treatment cost, fewer side effects, and no decrease in oocyte quality (9). However, the anti-estrogenic effect of CC may impair endometrial growth.

Few studies have examined CC stimulation in DOR. Herein, we report a large series of women with DOR who underwent OS with CC.

METHODS

This retrospective descriptive study was conducted at an assisted reproductive technology center. Electronic records of the Clinique were screened to identify who underwent OS with CC only from January 2016 to December 2019. The study protocol was approved by the Koç University Clinical Research Ethics Committee (protocol no: 2022.431.IRB1.157, date: 05.12.2022).

CC-only stimulation was offered when the patient was not expected to develop more than 3 follicles, e.g., AFC ≤4 and/or serum AMH ≤1.1 ng/mL.

All women who met the above criteria and underwent CC-only OS were included in the study. Exclusion criteria were a body mass index over 30 kg/m^2 or oocyte cryopreservation. Those who received CC with gonadotropins or GnRH antagonists were also excluded.

Stimulation Protocols and Pituitary Suppression

A baseline scan was performed on the second or third day of the menstrual period to exclude any follicles greater than 12 mm in diameter. CC 100 mg/day orally (Klomen®, Koçak Farma, Türkiye) was administered for 5 days. Ovarian response was monitored by ultrasound and by measurement of serum estradiol, LH , and progesterone levels every 2-3 days, according to the physician's evaluation. If there were growing follicles >10 mm, monitoring was continued spontaneously; however, if there was no follicular activation, CC was continued for up to 10 days at the treating physician's discretion.

When the leading follicle was 18 mm, final oocyte maturation was induced with 250 µg recombinant human chorionic gonadotropin (250 µg/0.5 mL; Merck, Italy).

Transvaginal oocyte pick-up (OPU) under general anesthesia was performed 32 hours after the human chorionic gonadotropin injection. Generally, embryos were cultured until the day-5 stage. However, embryos with early-stage fragmentation were frozen on day 3 or day 4. Luteal phase support with vaginal micronized progesterone gel (crinone 8% vaginal gel, Merck, England) 90 mg twice a day, was started on the evening of embryo transfer (ET), and continued until a negative pregnancy test or 6th week of gestation. Additionally, weekly depot progesterone (500 mg) was administered intramuscularly (Proluton Depot, Bayer, Germany) until a negative pregnancy test or until the 10th week of gestation.

The outcomes were: duration of stimulation, total CC dosage used, number of oocytes retrieved, number of metaphase II oocytes retrieved, oocyte maturation rate, number of embryos, pregnancy rate, and live birth rate. Oocyte maturation rate was defined as the proportion of metaphase II oocytes among all collected oocytes per woman.

Clinical pregnancy was defined as visualization of a gestational sac on ultrasound 4-6 weeks after ET. Ongoing pregnancy was defined as pregnancy continuing beyond 24 weeks' gestation. Live birth was defined as any pregnancy that resulted in the birth of an infant with heartbeat after the 24th gestational week. Electronic records of all patients during the aforementioned time periods were screened for demographic, stimulation, laboratory, and clinical data.

Statistical Analysis

Variable distributions were evaluated visually using histograms. Continuous variables were summarized as mean (standard deviation) or median (25th-75th percentiles), depending on their distributional characteristics. Categorical variables were defined with numbers and percentages. Sample size calculation was not performed for this retrospective analysis; however, all cycles meeting the inclusion criteria within the specified period were included.

RESULTS

A total of 840 cycles from 381 women were included in the analysis. Median age was 39 years. Median AFC and AMH were 3 and 0.34 ng/mL, respectively. CC was administered for a median of 5 days, and stimulation lasted a median of 8 days. Baseline characteristics are summarized in Table 1.

Of 840 cycles, 217 (25.9%) were cancelled before OPU due to premature ovulation and the absence of follicular growth (Table 2). The remaining 623 cycles (74.1%) proceeded to oocyte retrieval. No oocytes were retrieved in 66 cycles (7.9%), and in 61 cycles (7.3%) all retrieved oocytes were immature (Table 2).

The median number of oocytes collected, mature oocytes, and 2 pronuclei embryos was 2 in 496 cycles. Two cycles (0.4%) had no embryos (Table 3).

Endometrial thickness on the day of ET was 9.0 mm. Frozen ET was performed in 443 cycles (89.1%) and fresh ET in 51 cycles (10.3%). One ET was performed in 229 cycles, and two ETs were performed in 265 cycles. Pregnancy tests were positive in 131 cycles (15.6% per started cycle; 21.0% per OPU; 26.6% per ET). There were 96 live births (one was a twin birth); the live birth rate per ET was 19.5% (11.4% per started cycle; 15.4% per OPU).

DISCUSSION

CC alone seems to be an inexpensive and effective agent for OS herein in patients with severe DOR. As highlighted by meta-analyses and other studies, no single treatment approach has been established as superior for women with DOR (6,9,10). The number of oocytes retrieved and

Table 1. Baseline characteristics of the patients

	Median (25th-75th quartiles)
Age	39 (35-43)
Parity	0
Number of abortions	0
Number of previous IVF treatments	3 (2-6)
Infertility duration (months)	48 (24-90)
BMI (kg/m²)	24.1 (22.0-27.1)
AFC	3 (2-4)
AMH (ng/dL)	0.34 (0.21-0.49)
Day 3 estradiol (pg/mL)	38 (29-59.3)
Day 3 progesterone (pg/mL)	0.2 (0.1-0.5)

IVF: In vitro fertilization, BMI: Body mass index, AFC: Antral follicle count, AMH: Anti-müllerian hormone

Table 2. Cancellation reasons and rates in all cycles

	Number of cycles, (% in total number of cycles)
Number of cancelled cycles before OPU	217 (25.9)
No growing follicles	98 (11.7)
Premature ovulation	119 (14.2)
Number of cancelled cycles after OPU	127 (15.2)
No oocytes in OPU	66 (7.9)
All immature oocytes	61 (7.3)

Total number of cycles: 840. OPU: Oocyte pick-up

Table 3. Fertility outcomes of the clomiphene induced cycles

	Median (25th-75th quartiles)
HCG day estradiol (pg/mL)	464.5 (295.0-693.0)
HCG day progesterone (pg/mL)	0.14 (0.1-0.27)
Endometrial thickness at ET day (mm)	9.0 (8.5-11.0)
Starting gonadotropin dose	100
Duration of treatment (days)	5
Total gonadotropin dose	500
Number of total oocytes	2 (1-3)
Number of MII oocytes	2 (1-3)
Number of 2PN embryos	2 (1-2)
Number of blasts	0 (0-1)
Number of day 5 blasts	0
Ratio of MII oocytes to total number of oocytes	1
Fertility rate	1 (0.81-1)

PN: Pronuclei, HCG: Human chorionic gonadotropin, ET: Embryo transfer, MII: Metaphase II

pregnancy rates with CC-only or mild stimulation protocols are comparable to those achieved with standard protocols. Moreover, CC-only and mild-stimulation approaches offer financial advantages (3). However, CC has some undesirable effects, including premature ovulation, insufficient follicular growth, and endometrial thinning (11). In a randomized controlled trial (RCT), Ragni et al. (9) reported similar cycle-cancellation rates in the CC-only and high-dose gonadotropin groups, attributable to premature ovulation or follicular arrest/absence. They used, as treatment groups, 150 mg/day for 5 days starting on the third day of the cycle, and 450 IU recombinant follicle-stimulating hormone plus 0.1 mg GnRH agonist. In our study, we administered CC at 100 mg/day for 5 days, with treatment extended to 10 days for selected patients.

Compared with Ragni et al.'s (9) findings, our study observed a higher incidence of premature ovulation, while the rate of inadequate follicular growth was similar. Although oocyte retrieval rates were higher in the high-dose gonadotropin group, live birth rates remained comparable between groups (9). The average number of oocytes retrieved was 1.1±1.1 in Ragni et al.'s (9) study, ours was 2 (median). Notably, their pregnancy rates were low across both treatment arms, with rates per cycle initiation, per OPU, and per ET of 5%, 6%, and 14%, respectively, and live birth rates of 3%, 4%, and 9%. In contrast, the live birth rate per ET in our study was 19.5%.

Revelli et al. (12) compared mild stimulation with the long GnRH agonist protocol and reported higher cancellation rates (13%) in the mild-stimulation group, along with a shorter duration of stimulation, lower total gonadotropin consumption, and fewer metaphase II oocytes and ETs. However, clinical pregnancy rates, ongoing pregnancy rates (OPR), and OPR per ET were similar between the two protocols. Their OPR per ET rate of 17.8% closely aligns with our findings.

One of the earliest studies comparing mild stimulation with the GnRH agonist long protocol, conducted by D'Amato et al. (13), found a higher cancellation rate in the long protocol group than in the mild stimulation group, particularly among patients older than 35 years. In this age group, the number of oocytes retrieved was lower with the long protocol, and pregnancy rates were lowest, although the differences did not reach statistical significance.

Karimzadeh et al. (14) compared mild stimulation with the microdose flare-up protocol and found that, in the mild stimulation group, endometrial thickness and follicle counts were higher, whereas total gonadotropin consumption was lower. Clinical pregnancy rates per ET and cycle cancellation rates were similar between groups. Because of the adverse effects of CC on endometrial thickness, frozen ET was performed more frequently in our study. Similarly, Ochin et al. (15) reported a higher rate of frozen ET among normoresponsive patients undergoing mild stimulation than among those receiving the long protocol, primarily attributable to inadequate endometrial thickness.

With respect to endometrial thickness, Shakerian et al. (16) demonstrated that it is not predictive of live birth; this finding was later confirmed by additional studies (17). In a multicenter RCT investigating different OS protocols in intrauterine insemination cycles for unexplained infertility, endometrial thickness was lower in CC cycles compared to gonadotropin cycles; however, this difference did not impact live birth rates (18).

Study Limitations

This is a retrospective, descriptive study of patients who underwent OS with only CC. The inclusion of a control group receiving gonadotropins would have strengthened the study's design and enhanced its scientific validity.

CONCLUSION

Given the high cost of gonadotropins and the comparable effectiveness of various OS protocols in women with DOR, CC represents a viable and cost-effective alternative in this clinical setting. A limitation of our study is the absence of a control group undergoing conventional OS with gonadotropins; however, as a descriptive study, it provides valuable insights. If future comparative prospective trials confirm the efficacy of CC, it may serve as a cost-effective alternative, reducing the financial and treatment burdens for women with DOR.

ETHICS

Ethics Committee Approval: The study protocol was approved by the Koç University Clinical Research Ethics Committee (protocol no: 2022.431.IRB1.157, date: 05.12.2022).

Informed Consent: Retrospective study.

FOOTNOTES

Authorship Contributions

Surgical and Medical Practices: R.Ö., A.K., O.Ö., Concept: R.Ö., A.K., M.B.A., Design: S.G.Ç., A.K., M.B.A., Data Collection or Processing: R.Ö., S.G.Ç., A.K., O.Ö., Analysis or Interpretation: S.G.Ç., M.B.A., Literature Search: R.Ö., S.G.Ç., Writing: R.Ö., S.G.Ç., M.B.A.

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

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

REFERENCES

1. Keay SD, Liversedge NH, Mathur RS, Jenkins JM. Assisted conception following poor ovarian response to gonadotrophin stimulation. Br J Obstet Gynaecol. 1997;104:521-7.
2. Pandian Z, McTavish AR, Aucott L, Hamilton MP, Bhattacharya S. Interventions for 'poor responders' to controlled ovarian hyper stimulation (COH) in in-vitro fertilisation (IVF). Cochrane Database Syst Rev. 2010:CD004379.
3. Ferraretti AP, La Marca A, Fauser BC, Tarlatzis B, Nargund G, Gianaroli L; ESHRE working group on Poor Ovarian Response Definition. ESHRE consensus on the definition of 'poor response' to ovarian stimulation for in vitro fertilization: the Bologna criteria. Hum Reprod. 2011;26:1616-24.
4. Hofmann GE, Sosnowski J, Khoury J, Kearns L, Thie J. Obstetric implications for and identification of women with a normal clomiphene citrate challenge test result who have a poor response to gonadotropins. Fertil Steril. 2004;82:908-12.
5. Siristatidis CS, Hamilton MP. What should be the maximum FSH dose in IVF/ICSI in poor responders? J Obstet Gynaecol. 2007;27:401-5.
6. Song D, Shi Y, Zhong Y, Meng Q, Hou S, Li H. Efficiency of mild ovarian stimulation with clomiphene on poor ovarian responders during IVF\ICSI procedures: a meta-analysis. Eur J Obstet Gynecol Reprod Biol. 2016;204:36-43.
7. Glasier AF. Clomiphene citrate. Baillieres Clin Obstet Gynaecol. 1990;4:491-501.
8. Kerin JF, Liu JH, Phillipou G, Yen SS. Evidence for a hypothalamic site of action of clomiphene citrate in women. J Clin Endocrinol Metab. 1985;61:265-8.
9. Ragni G, Levi-Setti PE, Fadini R, Brigante C, Scarduelli C, Alagna F, et al. Clomiphene citrate versus high doses of gonadotropins for in vitro fertilisation in women with compromised ovarian reserve: a randomised controlled non-inferiority trial. Reprod Biol Endocrinol. 2012;10:114.
10. Kyrou D, Kolibianakis EM, Venetis CA, Papanikolaou EG, Bontis J, Tarlatzis BC. How to improve the probability of pregnancy in poor responders undergoing in vitro fertilization: a systematic review and meta-analysis. Fertil Steril. 2009;91:749-66.
11. Nakagawa K, Kaneyama M, Nishi Y, Sugiyama R, Motoyama H, Sugiyama R. Clomiphene citrate affects the receptivity of the uterine endometrium. Reprod Med Biol. 2015;14:73-8.
12. Revelli A, Chiado A, Dalmasso P, Stabile V, Evangelista F, Basso G, et al. "Mild" vs. "long" protocol for controlled ovarian hyperstimulation in patients with expected poor ovarian responsiveness undergoing in vitro fertilization (IVF): a large prospective randomized trial. J Assist Reprod Genet. 2014;31:809-15.
13. D'Amato G, Caroppo E, Pasquadibisceglie A, Carone D, Vitti A, Vizziello GM. A novel protocol of ovulation induction with delayed gonadotropin-releasing hormone antagonist administration combined with high-dose recombinant follicle-stimulating hormone and clomiphene citrate for poor responders and women over 35 years. Fertil Steril. 2004;81:1572-7.
14. Karimzadeh MA, Mashayekhy M, Mohammadian F, Moghaddam FM. Comparison of mild and microdose GnRH agonist flare protocols on IVF outcome in poor responders. Arch Gynecol Obstet. 2011;283:1159-64.
15. Ochin H, Ma X, Wang L, Li X, Song J, Meng Y, et al. Low dose clomiphene citrate as a mild stimulation protocol in women with unsuspected poor in vitro fertilization result can generate more oocytes with optimal cumulative pregnancy rate. J Ovarian Res. 2018;11:37.
16. Shakerian B, Turkgeldi E, Yildiz S, Keles I, Ata B. Endometrial thickness is not predictive for live birth after embryo transfer, even without a cutoff. Fertil Steril. 2021;116:130-7.
17. Ata B, Linan A, Kalafat E, Ruiz F, Melado L, Bayram A, et al. Effect of the endometrial thickness on the live birth rate: insights from 959 single euploid frozen embryo transfers without a cutoff for thickness. Fertil Steril. 2023;120:91-8.
18. Quaas AM, Gavrizi SZ, Peck JD, Diamond MP, Legro RS, Robinson RD, et al. Endometrial thickness after ovarian stimulation with gonadotropin, clomiphene, or letrozole for unexplained infertility, and association with treatment outcomes. Fertil Steril. 2021;115:213-20.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):419-429
DOI: 10.4274/BMJ.galenos.2025.2024.9-4

Research

Evaluation of the Effect of Dietary Habits on the Health Profile in Women with Endometriosis

Endometriozisli Kadınlarda Beslenme Alışkanlıklarının Sağlık Profili Üzerine Etkisinin Değerlendirilmesi

Tuğçe Deveci[1], Nihan Çakır Biçer[2,3], Pınar Yalçın Bahat[4]

[1]Acıbadem Mehmet Ali Aydınlar University Institute of Health Sciences, Department of Nutrition and Dietetics, İstanbul, Türkiye
[2]Acıbadem Mehmet Ali Aydınlar University Faculty of Health Sciences, Department of Nutrition and Dietetics, İstanbul, Türkiye
[3]Kocaeli University Institute of Gastroenterology and Hepatology, Department of Molecular Gastroenterology and Hepatology, Kocaeli, Türkiye
[4]University of Health Sciences Türkiye, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Türkiye

ABSTRACT

Objective: Healthy eating habits in women with endometriosis play a role in the treatment and improvement of quality of life. This study aimed to evaluate the effect of adherence to the Mediterranean diet (MD) and polyphenol intake on health profile in endometriosis.

Methods: The study was conducted with 34 endometriosis patients admitted to Gynecology and Obstetrics Clinic of University of Helath Sciences Türkiye, İstanbul Kanuni Sultan Süleyman Training and Research Hospital. Demographic characteristics and food consumption were recorded, the Healthy Eating Index (HEI)-2015 scores were calculated, and the polyphenol consumption questionnaire, the 14-item MD Adherence Scale (MEDAS), and the Endometriosis Health Profile-30 (EHP-30) were applied.

Results: The mean body mass index of the participants was 22.92±2.94 kg/m^2, total polyphenol intake was 657.70±304.56 mg/day, and MEDAS, HEI-2015, and EHP-30 scores were 7.85±2.26, 61.71±18.51, and 41.25±20.77, respectively. 26.5% of participants had low; 44.1% had medium; 29.4% had high adherence to the MD; and 44.1% had low HEI-2015 scores. While no relationship was found between the MEDAS score and EHP-30 total and subscale scores, an inverse relationship was found between the HEI-2015 score and EHP-30 pain subscale score (r=-0.34, p=0.04).

Conclusion: Despite the role of nutrition in endometriosis, the adherence of the participants to the MD and their HEI-2015 scores were found to be low. It is predicted that a multidisciplinary treatment approach for endometriosis with the implementation of the MD and the promotion of healthy lifestyle habits will improve symptom management and quality of life.

Keywords: Endometriosis, Mediterranean diet, polyphenols, health profile

ÖZ

Amaç: Endometriozisli kadınlarda sağlıklı beslenme alışkanlıklarının hastalığın tedavisinde rol oynayacağı ve yaşam kalitesinin artırılmasına destek olacağı bildirilmektedir. Bu çalışmada Akdeniz diyetine uyumun ve polifenol alımının endometriozis sağlık profili üzerine etkisinin değerlendirilmesi amaçlanmıştır.

Gereç ve Yöntem: Sağlık Bilimleri Üniversitesi, İstanbul Kanuni Sultan Süleyman Eğitim ve Araştırma Hastanesi Kadın Hastalıkları ve Doğum Polikliniği'ne başvuran 34 endometriozis hastasının demografik özellikleri ve besin tüketimleri kaydedilmiş, Sağlıklı Yeme İndeksi (HEI)-2015 puanları hesaplanmış ve katılımcılara polifenol tüketim sıklığı anketi, 14 maddeli Akdeniz Diyeti Uyum Ölçeği (MEDAS), Endometriozis Sağlık Profili Anketi (EHP-30) uygulanmıştır.

Bulgular: Katılımcıların ortalama beden kütle indeksi 22,92±2,94 kg/m^2, MEDAS puanı 7,85±2,26, HEI-2015 puanı 61,71±18,51 ve EHP-30 toplam puanı 41,25±20,77 ve toplam polifenol alımı 657,70±304,56 mg/gün'dür. %26,5'inin Akdeniz diyetine düşük, %44,1'inin orta, %29,4'ünün yüksek düzeyde uyum gösterdiği, %44,1'inin HEI-2015 puanının düşük olduğu bulunmuştur. MEDAS puanı, EHP-30 toplam ve alt boyut puanları arasında ilişki saptanmazken, HEI-2015 puanı ile EHP-30 ağrı alt boyut puanı arasında ters yönlü ilişki belirlenmiştir (r=-0,34, p=0,04).

Address for Correspondence: Assist, Prof, Nihan Çakır Biçer, PhD, Acıbadem Mehmet Ali Aydınlar University Faculty of Health Sciences, Department of Nutrition and Dietetics, İstanbul, Türkiye; Kocaeli University Institute of Gastroenterology and Hepatology, Department of Molecular Gastroenterology and Hepatology, Kocaeli, Türkiye
E-mail: nihan.bicer@acibadem.edu.tr, nihancakir@gmail.com **ORCID ID:** orcid.org/0000-0002-7581-9083

Cite as: Deveci T, Çakır Biçer N, Yalçın Bahat P. Evaluation of the effect of dietary habits on the health profile in women with endometriosis. Med J Bakirkoy. 2025;21(4):419-429

Received: 17.09.2024
Accepted: 04.03.2025
Publication Date: 18.12.2025

ÖZ

Sonuç: Endometrioziste beslenmenin rolüne dikkat çekilmesine rağmen, çalışmaya katılan kadınların sağlıklı bir beslenme modeli olan Akdeniz diyetine uyumu ve HEI-2015 puanları düşük saptanmıştır. Endometriozis tedavisinde multidisipliner bir tedavi yaklaşımı ile Akdeniz diyetinin uygulanmasının ve sağlıklı yaşam tarzı alışkanlıklarının teşvik edilmesinin semptom yönetimini ve yaşam kalitesini iyileştireceği ön görülmektedir.

Anahtar Kelimeler: Endometriozis, Akdeniz diyeti, polifenol, sağlık profili

INTRODUCTION

Endometriosis, which usually affects women of reproductive age, is defined as a chronic gynecological disease associated with estrogen-dependent infertility and pelvic pain. It has been reported that the prevalence of endometriosis varies according to ethnicity, with black women having a higher prevalence than Asian women (1). 10% of women of reproductive age (~190 million women) and 18% of Turkish women are affected by endometriosis (2,3). At least one-third of women with chronic pelvic pain and 5% of women without pain or infertility have endometriosis (4). Endometriosis may be accompanied by fibromyalgia, allergic rhinitis and asthma, iron deficiency anemia, lipedema, depression, fibrocystic breasts, polycystic ovaries, adenomyosis, and central-peripheral nervous system diseases (5-7). In addition to medical, pharmacological, and surgical treatment in endometriosis, healthy eating habits and positive lifestyle changes have been reported to be effective in alleviating the symptoms of the disease (8).

According to Casas et al. (9) and Cirillo et al. (10), the Mediterranean diet (MD) has a protective and therapeutic effect on many chronic diseases and can reduce inflammation and associated pain symptoms. Vegetables and fruits are the main antioxidant sources of MD due to their rich content of vitamins, minerals, and phytochemicals, and these foods have been associated with a reduced risk of developing the disease. In the management of endometriosis, while saturated fat, red meat, and alcohol consumption have been associated with an increased risk of the disease (11). Additionally, the olive oil composition and omega-3 fatty acids of the MD has been associated with a reduced risk of endometriosis (8,10). It has been reported that dietary fatty acid composition has the potential to affect endogenous hormone metabolism and endogenous estrogen levels (12). Polyphenols, which are secondary metabolites of plants, are defined as compounds that are naturally found in foods and are generally considered safer than traditional pharmaceutical drugs. They also have antioxidant properties. Polyphenols are divided into different categories such as flavonoids, stilbenes, lignans, and phenolic acids according to the number of phenol rings they contain and the structural elements that connect these rings (13). It has been reported that these components may improve the quality of life by modulating estrogen networks, thanks to their anti-inflammatory and potential phytoestrogenic effects, without causing serious adverse effects, unlike anti-estrogenic treatment (14). Nutrition also contributes to the management of endometriosis, a disease characterized by inflammation and affecting life in many ways. This study aimed to evaluate the relationship between the nutritional status of endometriosis patients as assessed by the Healthy Eating Index (HEI), their adherence with the MD, their polyphenol intake, and their health status assessed by the Endometriosis Health Profile-30 (EHP-30).

METHODS

Research Design and Sample of the Study

Acıbadem Mehmet Ali Aydınlar University and Acıbadem Healthcare Institutions Medical Research Ethics Committee (ATADEK) approval was received (approval no: 2023-13/447, date: 17.08.2023). All patients were informed about the study, and written informed consent was obtained for the protocol. This prospective, cross-sectional study was conducted with women who applied to the Obstetrics and Gynecology Clinic of University of Helath Sciences Türkiye, İstanbul Kanuni Sultan Süleyman Training and Research Hospital between September and December 2023. Women aged 18-55 years, who were not pregnant or breastfeeding, not in menopause, and diagnosed with endometriosis stage I, II, or III by a physician, were included in the study. Individuals on a special diet were not included in the study. The sample of the study was determined as 34 participants, based on the research of Asencio et al. (15), by performing a power analysis using the R v3.6.1 program (Vienna, Austria), taking the margin of error as 5% and the beta error as 20%.

Data Collection Tools

Descriptive characteristics form: Sociodemographic and menstrual characteristics and health status were questioned. Anthropometric measurements (height, body weight, waist, and hip circumference) were taken by the researcher. Body mass index (BMI) was calculated and grouped by the recommendations of the World Health Organization (WHO) (16).

14-Point Mediterranean Diet Adherence Screener (MEDAS): The MEDAS was developed by Martínez-González et al. (17) in 2012 to assess adherence with the MD, and its Turkish validation was conducted by Bekar and Goktas (18) in 2023. The MEDAS consists of 14 questions that evaluate the daily consumption of vegetables, fruits, meat, butter, margarine, cream, alcohol and sugary drinks, olive oil, and the weekly consumption of legumes, fish, oilseeds, pastries, sauces, and the type of oil used in meals. A score of <7 indicates low adherence with the MD, 7-9 points indicate moderate adherence, and >9 points indicate high adherence. The Cronbach's alpha value of the MEDAS was reported as 0.829 (18).

Food Records: Food consumption was evaluated using a 24-hour recall method, and energy, macronutrient, and micronutrient intakes were calculated using the Nutrition Information System software (Türkiye, İstanbul).

Polyphenol Consumption Frequency Survey: To determine polyphenol intake, the Phenol-Explorer (version 3.6) database and literature were scanned, and a questionnaire was created focusing on the frequency and amount of consumption of 64 foods with high polyphenol content in the last year. According to the survey, daily intake was determined, and polyphenol consumption was calculated using Phenol-Explorer (19).

HEI-2015: The HEI-2015 was developed by the Center for Nutrition Policy and Promotion of the United States Department of Agriculture to measure adherence to dietary guidelines. The HEI-2015 consists of 13 components, the first 9 of which determine the adequacy of the diet and the last 4 of which should be consumed in limited amounts. The total score is calculated out of 100 points by adding the adequacy and limited consumption components and is defined as "poor diet quality" if ≤50, "dietary quality that needs improvement" if 51-80, and "good diet quality" if >80 (20). In addition, letter grades are given to the score ranges: 90-100 points: A, 80-89 points: B, 70-79 points: C, 60-69 points: D, and 0-59 points: F (21).

Endometriosis Health Profile-30: EHP-30 was developed by Jones et al. (22) to determine health-oriented quality of life criteria for women with endometriosis. The Turkish validation was conducted by Şahin (23). The survey consists of 53 questions and 11 subscales, including pain, control, powerlessness, emotional well-being, social support, self-image, work, sexual intercourse, relationship with children, the medical profession, treatment, and infertility. The survey is answered by considering the last four weeks, and is scored between 0 to 100, with 0 indicating the best and 100 indicating the worst health status. The Cronbach's alpha value of the EHP-30 was reported as 0.980 (23).

Statistical Analysis

The data were analyzed using the SPSS statistics 25.0 program (Armonk, NY). The skewness and kurtosis values of the variables were calculated to assess conformity with the normal distribution, and if the values were within ±2, it was assumed that they showed a normal distribution. The Pearson correlation analysis and the Spearman correlation analysis were used to evaluate the relationship between continuous variables. $P<0.05$ was accepted as significant. For this study, Cronbach's alpha values for the MEDAS and EHP-30 were found to be 0.747 and 0.963, respectively.

RESULTS

The demographic and anthropometric characteristics of the participants are shown in Table 1. The mean age of the 34 women was 35.44±6.72 years. It was determined that 17.6% (n=6) of the participants had secondary education, 79.4% (n=27) had higher education, 58.8% (n=20) were married, 88.2% (n=30) were employed, 35.3% (n=12) had income greater than their expenses, and 38.2% (n=13) had income equal to their expenses. The mean BMI of the participants was 22.92±2.94 kg/m^2, and 64.7% (n=22) of the women had normal weight, while 32.4% (n=11) were overweight (Table 1).

Table 2 presents the characteristics of disease status and menstrual cycle. 64.7% (n=22) of the women had a disease diagnosed by a physician other than endometriosis, and 17.6% (n=6) of these women had bowel problems, 5.9% (n=2), fibromyalgia, 5.9% (n=2), allergic rhinitis and asthma, 5.9% (n=2), fibrocystic breast, and 5.9% (n=2), PCOS and adenomyosis. 79.4% (n=27) of the participants reported regular menstruation (Table 2).

According to food consumption records, it was determined that participants consumed a mean of 1337.25±403.19 kcal/day of energy, 133.58±70.07 g/day carbohydrate, 59.68±17.28 g/day protein, 61.21±22.34 g/day fat, 15.29±6.03 g/day fiber, and 10.22±9.51 g/day added sugar. According to the recommendations of the Türkiye Dietary Guidelines (TDG), the ratio of energy intake from carbohydrates to total daily energy intake of the participants (38.45±12.74%) was below the daily recommendations (45-60%), and the ratio of lipid intake (41.00±12.07%) was above the recommendations (20-35%). In addition, fiber intake was below the recommendations, while added sugar intake (3.47±3.94%) was in line with them.

Table 1. Demographic and anthropometric characteristics of the participants

	Women with endometriosis (n=34)			
	Mean	**SD**	**Min**	**Max**
Age (year)	35.44	6.72	23.00	51.00
Height (cm)	163.09	5.98	150.00	175.00
Weight (kg)	60.98	8.69	44.00	82.00
BMI (kg/m^2)	22.92	2.94	16.98	29.05
Waist circumference (cm)	77.32	8.59	63.00	100.00
Hip circumference (cm)	100.94	7.71	86.00	118.00
Waist/hip ratio	0.77	0.07	0.57	0.93
Education status	**n**		**%**	
Primary school	1		2.9	
Secondary education	6		17.6	
Higher education	27		79.4	
Marital status				
Married	20		58.8	
Single	14		41.2	
Occupational status				
Employee	30		88.2	
Retired/non-employed	4		11.8	
BMI category				
Underweight (<18.5 kg/m^2)	1		2.9	
Normal (18.5-24.9 kg/m^2)	22		64.7	
Overweight (25.0-29.9 kg/m^2)	11		32.4	

BMI: Body mass index, Min: minimum, Max: Maximum, SD: Standard deviation

Table 2. Disease and menstrual cycle status of the participants

Women with endometriosis (n=34)		
Having a disease diagnosed by a physician	**n**	**%**
Yes	22	64.7
No	12	35.3
Diseases diagnosed by a physician* (n=22)		
Intestinal problems	6	17.6
Fibromyalgia	2	5.9
Allergic rhinitis and asthma	2	5.9
Iron deficiency anemia	2	5.9
Fibrocystic breast - cyst in the breast	2	5.9
PCOS-adenomyosis	2	5.9
Lipedema	1	2.9
Depression	1	2.9
Other diseases	10	29.4
Having a regular menstrual cycle		
Yes	27	79.4
No	7	20.6

*: Multiple response, PCOS: Polycystic ovary syndrome

Table 3. The energy, nutrient, and polyphenol intakes of the participants

Women with endometriosis (n=34)						
	Mean	**SD**	**Min**	**Max**	**TDG %***	**TDG recommendations**
Energy (kcal/day)	1337.25	403.19	577.30	2253.30	77.20[a]	-[a]
CHO (g/day)	133.58	70.07	28.30	298.80	-[b,c]	130[c]
CHO (%)	38.45	12.74	12.61	60.23	-[b]	45-60[b]
Protein (g/day)	59.68	17.28	24.10	93.00	112.7[a,b]	-[a]
Protein (%)	18.58	5.64	9.13	30.51	-[b]	10-20[b]
Lipid (g/day)	61.21	22.34	20.10	118.00	-[b]	-[d]
Lipid (%)	42.00	12.07	19.36	57.96	-[b]	20-35[b]
Fiber (g/day)	15.29	6.03	5.00	27.20	61.16[c]	25[c]
Added sugar (g/day)	10.22	9.51	0.00	32.50	-[e]	-[e]
Added sugar (%)	3.47	3.42	0.00	16.70	-[e]	<10[e]
SFA (g/day)	20.82	10.67	0.00	46.40	-[c]	-[c]
SFA (%)	14.91	4.73	7.47	24.56	-[c]	-[c]
MUFA (g/day)	21.60	8.51	6.80	39.20	-[d]	-[d]
PUFA (g/day)	11.87	9.02	2.90	49.90	-[d]	-[d]
Omega-3 (g/day)	1.47	1.11	0.40	4.30	-[d]	-[d]
Omega-6 (g/day)	10.2 3	8.44	2.10	46.00	-[d]	-[d]
Polyphenols (mg/day)						
Flavonoids	339.86	183.05	33.42	682.68	-[d]	-[d]
Flavonols	57.56	35.59	1.89	130.87	-[d]	-[d]
Flavanons	20.99	18.65	0.95	78.82	-[d]	-[d]
Flavanols	246.43	158.49	8.07	507.92	-[d]	-[d]
Flavons	8.43	6.26	2.49	31.61	-[d]	-[d]
Anthocyanidin	17.75	29.13	0.12	151.86	-[d]	-[d]
Stilbens	0.37	0.44	0.00	1.96	-[d]	-[d]
Lignans	10.51	6.44	2.98	31.20	-[d]	-[d]
Phenolic acids	207.66	118.03	37.38	440.50	-[d]	-[d]
Hydroxycinnamic acid	156.96	113.74	23.77	424.87	-[d]	-[d]
Hydroxybenzoic acid	50.70	32.19	0.75	107.25	-[d]	-[d]
Other polyphenols	94.70	153.68	5.09	769.48	-[d]	-[d]
Total polyphenols	657.70	304.56	146.85	1313.44	-[d]	-[d]

*: TDG %: The rate of meeting Türkiye Dietary Guidelines (TDG) recommendations.
a: Calculated according to age and the Türkiye Nutrition Guide adequate requirement of energy and AR/population reference intakes. b: Reference intake ranges for macronutrients (%) are recommended.
c: Population reference intakes is recommended as 130 g/day for CHO, and adequate intake is recommended as 25 g/day for fiber and as little as possible for SFA.
d: There is no specific intake recommendation.
e: It is recommended not to exceed 10% of energy intake.
CHO: Carbohydrate, SFA: Saturated fatty acids, MUFA: Monounsaturated fatty acids, PUFA: Polyunsaturated fatty acids, SD: Standard deviation, Min: Minimum, Max: Maximum

According to the polyphenol consumption frequency questionnaire, participants consumed a mean of 339.86 mg/day of flavonoids, 207.66 mg/day of phenolic acids, and 657.70 mg/day of total polyphenols (Table 3).

The mean MEDAS score was 7.85±2.26; the HEI-2015 score was 61.71±18.51 and the EHP-30 total score was 41.25±20.77. According to MEDAS scores, it was determined that 26.5% (n=9) of the participants had low, 44.1% (n=15) moderate, and 29.4% (n=10) high adherence with the MD; and according to HEI-2015 scores, in the distribution of letter categories, 5.9% (n=2) received A grade, 14.7% (n=5) received B grade, 20.6% (n=7) received C grade, 14.7% (n=5) received D grade and 44.1% (n=15) received F grade (Table 4).

The relationship between MEDAS, HEI-2015, and EHP-30 total, and subscales scores is shown in Table 5. A weak negative correlation was found between the EHP-30 pain subscale score and the HEI-2015 score ($r=-0.34$, $p<0.05$), but no correlation was found between the EHP-30 total and other subscales scores and the HEI-2015, and MEDAS scores ($p>0.05$, for all).

DISCUSSION

The MD is reported to be effective in relieving endometriosis-related pain, especially with its antioxidant content and anti-inflammatory effects. It is a nutritional model that has positive effects on the treatment process of endometriosis (10). This study aimed to evaluate the relationship between the nutritional habits in women diagnosed with endometriosis, and their health profile.

According to the Turkish Statistical Institute (TURKSTAT) 2023 data, the rate of higher education graduates among women was reported to be 20.9%, and the labor force participation rate for those over the age of 15 was reported to be 32.8% (24). The rate of higher education graduates or employed individuals in this study was found to be above the Turkish average, indicating that the women participating in the study were more informed and economically active individuals.

According to TURKSTAT 2023 data, 23.6% of women are obese (24). In a study conducted in Türkiye on endometriosis, the mean waist circumference (WC) was 75.02 cm, and BMI was 22.43 kg/m^2, and there were no underweight or obese subjects in the study (25). Similarly, in this study, WC was 77.32 cm and BMI was 22.92 kg/m^2, and although there were no obese women, the majority of women (64.7%) were found to be of normal weight.

Table 4. The MEDAS, HEI-2015, and EHP-30 scores of the participants

Women with endometriosis (n=34)	**Mean**	**SD**	**Min**	**Max**
MEDAS	7.85	2.26	2.00	11.00
HEI-2015	61.71	18.51	25.59	94.25
EHP-30	41.25	20.77	3.92	91.18
Pain	43.32	25.18	0.00	100.00
Control and powerlessness	48.04	31.69	0.00	100.00
Emotional well-being	40.44	29.36	0.00	100.00
Social support	44.36	28.39	0.00	87.50
Self-image	40.69	31.37	0.00	100.00
Work	31.62	27.95	0.00	100.00
Relationship with children	11.03	24.57	0.00	100.00
Sexual intercourse	35.48	35.59	0.00	100.00
Medical profession	15.99	21.22	0.00	100.00
Treatment	48.53	34.17	0.00	100.00
Infertility	36.95	32.87	0.00	100.00
MEDAS	**n**		**%**	
Low adherence	9		26.5	
Medium adherence	15		44.1	
High adherence	10		29.4	
HEI-2015 letter categories				
A (90-100 point)	2		5.9	
B (80-89 point)	5		14.7	
C (70-79 point)	7		20.6	
D (60-69 point)	5		14.7	
F (0-59 point)	15		44.1	

MEDAS: Mediterranean diet adherence screener, HEI-2015: Healthy eating index-2015, EHP-30: Endometriosis health profile-30, SD: Standard deviation, Min: Minimum, Max: Maximum

Table 5. Relationship between the MEDAS, HEI-2015, and EHP-30 scores

n=34		BMI	Total polyphenols	MEDAS	HEI-2015	EHP-30	1	2	3	4	5	6	7	8	9	10	11
MEDAS	r	-0.05	0.11	1.00													
	p	0.78	0.52	-													
HEI-2015	r	-0.18	-0.07	0.15	1.00												
	p	0.30	0.71	0.40	-												
EHP-30	r	0.19	0.06	-0.01	-0.19	1.00											
	p	0.29	0.74	0.94	0.29	-											
1. Pain	r	0.22	0.06	-0.23	-0.34	0.83	1.00										
	p	0.20	0.72	0.18	**0.04***	**<0.001***	-										
2. Control and powerlessness	r	0.21	0.04	-0.05	-0.21	0.87	0.71	1.00									
	p	0.24	0.82	0.80	0.23	**<0.001***	0.00	-									
3. Emotional well-being	r	0.02	-0.15	0.08	0.15	0.79	0.53	0.74	1.00								
	p	0.91	0.39	0.66	0.39	**<0.001***	**<0.001***	**<0.001***	-								
4. Social support	r	0.15	0.02	0.07	0.11	0.76	0.48	0.71	0.82	1.00							
	p	0.39	0.91	0.69	0.53	**<0.001***	**<0.001***	**<0.001***	**<0.001***	-							
5. Self-image	r	0.29	-0.07	0.09	0.09	0.52	0.41	0.48	0.46	0.45	1.00						
	p	0.09	0.69	0.63	0.61	**<0.001***	**0.02***	**<0.001***	**0.01***	**0.01***	-						
6. Work	r	0.16	0.09	-0.16	-0.08	0.64	0.64	0.48	0.45	0.26	0.28	1.00					
	p	0.38	0.60	0.35	0.65	**<0.001***	**<0.001***	**<0.001***	0.01*	0.14	0.11	-					
7. Relationship with children	r	0.26	0.35	0.11	0.06	0.16	0.18	0.30	0.04	0.18	-0.02	-0.05	1.00				
	p	0.37	**0.04***	0.55	0.75	0.37	0.32	0.09	0.84	0.31	0.93	0.78	-				
8. Sexual intercourse	r	0.05	0.28	0.14	-0.12	0.61	0.47	0.45	0.46	0.44	0.08	0.37	0.20	1.00			
	p	0.80	0.11	0.44	0.52	**<0.001***	**0.01***	**0.01***	**0.01***	**0.01***	0.64	**0.03***	0.27	-			
9. Medical profession	r	0.08	0.12	0.01	-0.10	0.45	0.27	0.32	0.27	0.19	0.20	0.05	0.02	0.35	1.00		
	p	0.91	0.51	0.94	0.58	**0.01***	0.13	0.07	0.13	0.28	0.27	0.80	0.89	0.04*	-		
10. Treatment	r	-0.14	0.12	0.18	-0.02	0.52	0.30	0.44	0.37	0.33	0.03	0.23	0.35	0.23	0.60	1.00	
	p	0.44	0.52	0.31	0.89	**<0.001***	0.09	0.01*	0.03**	0.06	0.86	0.18	**0.04***	0.19	**<0.001***	-	
11. Infertility	r	-0.06	-0.06	-0.05	-0.30	0.50	0.27	0.28	0.22	0.20	0.01	0.42	-0.18	0.26	0.30	0.38	1.00
	p	0.74	0.74	0.77	0.08	**<0.001***	0.12	0.11	0.21	0.26	0.96	**0.02***	0.30	0.14	0.08	**0.03***	-

BMI: Body mass index, MEDAS: Mediterranean diet adherence screener, HEI-2015: Healty eating index-2015, EHP-30: Endometriosis healt profile-30, *p<0.05

In addition, the WC of women was found to be below the WHO recommendations (<80 cm) (16). Although these results may be due to the demographic characteristics of women, they are similar to the literature also showing the anthropometric characteristics of women with endometriosis (26). Holdsworth-Carson et al. (27) reported that chronic pain symptoms related to endometriosis trigger gastrointestinal symptoms or emotional stress, leading to decreased appetite and food intake, which in turn leads to reduced body size in women with endometriosis. In this study, no significant difference was observed between BMI and EHP-30 subscale scores.

According to the Türkiye Nutrition and Health Survey 2019 data, the rate of women aged 15 years and older with a disease diagnosed by a physician was 48.5% (28). The rate of women with a disease diagnosed by a physician other than endometriosis was higher in this study (64.7%). It is thought that this may be due either to the characteristics of the sample group or to the educational level of the women, and differences in their access to health services.

The risk of endometriosis has been reported to be more than double in women with short menstrual cycles (≤27 days) and longer bleeding durations (≥1 week) (29). Consistent with the literature, more than two-thirds of women in this study reported regular menstruation, occurring at intervals of one month or less.

According to the food consumption records of the participants, the energy intake from carbohydrates is below the recommendations of the TDG 2022 (30). This is because fat consumption is above the recommendations. SFA intake at levels approximately three times higher than the recommended limit of 7% of daily energy intake also contributes to this result. A saturated fat intake of ≤8% of total energy has been reported to score the highest on the HEI-2015 (31). In this study, the number of individuals with high HEI-2015 scores is low, since SFA consumption is above the recommended limits. Furthermore, the recommendation for an optimal omega-6/omega-3 ratio varies from 1/1 to 4/1, depending on the disease, although in Western diets this ratio is 15/1 to 16.7/1 (32). In this study, the omega-6/omega-3 ratio that was determined as approximately 10/1 does not meet the recommendation. added sugar consumption (3.47%), which is associated with chronic diseases and inflammation, is below the WHO recommendations of "not exceeding 5-10% of daily energy" (33). It is thought that this result may be related to the socio-demographic characteristics of the population. Additionally, Bogusz and Górnicka (8) reported that 77.3% of women with endometriosis increased their consumption of anti-inflammatory foods after diagnosis. However, they had lower MD Score and Pro-Healthy Diet Index scores than the healthy control group and consumed fewer fruits, vegetables, dairy products, and whole grains than recommended. Similarly, in this study, the low number of individuals with adequate adherence to the MD according to the MEDAS score (29.4%) and the high number of individuals in the worst letter category according to the HEI-2015 score (44.1%) draw attention to the importance of the quality and nutrient diversity of the foods in the diet, as well as the daily energy and macronutrient consumption of individuals.

In adherence to the MD, polyphenol intake, which is one of the bioactive components in plant foods, also increases (9). Polyphenols have antioxidant properties, and a diet high in polyphenols may help relieve symptoms of endometriosis, an inflammatory disease (14).

Although there are no intake recommendations for polyphenols as there are for micronutrients, intake of polyphenol types has been evaluated in different studies. Polyphenol intake was reported to be higher in Japan (≥1277 mg/day) than in other regions. No significant differences were observed among European countries. Higher intakes were reported in Poland (≥500 mg/day) and France (≥1000 mg/day), intakes with lower Italy and Spain. Flavonoid and phenolic acid consumption was highest in Poland and Australia (34). In a study conducted in Türkiye, the total polyphenol intake of women was reported as 333.9 mg/day and the consumption of some polyphenol types, such as flavanols, flavonols, and phenolic acids, was lower in comparison to our study (13). In this study, total polyphenol consumption (657.70 mg/day) was higher than previously reported values. It is thought that the reason for this difference may be because the other study was conducted during the pandemic period, and with participants from different regions.

EHP-30 is a questionnaire developed to measure the quality of life of endometriosis patients and aims to evaluate the effects on their physical, mental, and social health (22). When the studies conducted with EHP-30 in the literature are evaluated, there are results both similar and different from those of this study (35-37). These differences may be due to the demographic or cultural characteristics of the participants. In this study, a positive correlation was found between the EHP-30 total score and all subscales except the relationship with children subscale. This shows that the EHP-30 is consistent among its subscales. In addition, the correlations between EHP-30 total and pain, control, and powerlessness subscales' scores were found to be very strong. This may be explained

by the fact that the subscales of the EHP-30, such as pain, control, and powerlessness, are the main factors that directly reflect the physical and emotional burden of the disease and seriously affect the quality of life. In the literature, these subscales are reported to have a significant impact on women with endometriosis (38). Another study reported that women with endometriosis experienced a loss of control and power due to pain and loss of functionality, which prevented them from participating in necessary and enjoyable activities that make life worth living (39). In a study on chronic pelvic pain, it was reported that women experienced a lack of social support and that this was associated with deep pain and suffering (40). In this study, a significant relationship was observed between pain and the scores of the emotional well-being and social support subscales' scores. All these results emphasize that pain management, especially, and psychosocial and holistic approaches are critical in coping with endometriosis. Studies have reported that pelvic pain due to endometriosis is associated with impaired body image, decreases quality of life and sexual satisfaction, and self-image is significantly weakened in women with dyspareunia (41,42). It was also reported that impaired body image was significantly associated with greater emotional loneliness and lower perceived social support (43). This study found positive correlations that are similar to those in the literature, specifically between pain and self-image and sexual life subscale scores, as well as between social support and self-image subscale scores. This suggests that women with higher perceived social support have a more positive self-image.

Recent studies on the role of the MD and its components in the treatment of endometriosis are notable for their contributions. In an endometriosis animal model study with *Urtica dioica* L. (above-ground parts of nettle), a significant and promising difference was found in TNF-alpha and IL-6 levels compared to the control group (44). In a study on clove and its components, clove and its components were reported to cause regression in endometrial implants without affecting fertility in animal models with endometriosis (45). In another study, clove, one of the important spices of the MD, showed a suppressive effect on pro-inflammatory cytokines (IL-6, TNF-alpha) with its antioxidant activity (46). Ott et al. (47) reported that the MD was effective in reducing endometriosis-associated pain. Another study reported an inverse association between adherence to the MD and pain, with adherence being associated with a higher intake of polyphenols (48). While Zhou et al. (49) found no association between pain and adherence to the MD in their study on middle-aged Chinese, Cirillo et al. (10) reported a significant association between the MD and pain-relieving effect. Dougan et al. (50) similarly reported that adherence to the alternative HEI would be beneficial for pelvic pain. In the present study, no correlation was found between the MEDAS scores assessing adherence to the MD and the HEI-2015 scores assessing diet quality, and the EHP-30 total score, but a negative correlation was found between the HEI-2015 and the pain subscale scores. This result points to the importance of nutrition in pain management in women with endometriosis. Although the effect of dietary changes varies individually, a period of several weeks to several months is usually required to reduce the inflammation and symptoms associated with endometriosis (51-53). The length of this period depends on the initial state of health, the level of inflammation and the degree of dietary compliance (52). Dietary patterns based on food restriction are feasible in the short term but are not sustainable in the long term.

Therefore, the literature recommends implementing such diets with personalized plans for a few weeks before switching to the MD for both its anti-inflammatory effects and sustainability (54).

Study Limitations

This study provides a detailed assessment of the nutritional status of women with endometriosis by applying MEDAS, HEI-2015, and polyphenol consumption frequency questionnaires. In addition, EHP-30, which provides assessments of many health-related areas for women with endometriosis, was applied and its relationship with nutrition was determined. While this situation constitutes a strength of the study, there are also some limitations. Considering the demographic characteristics of the individuals participating in this cross-sectional study, the sample does not reflect the general population. In addition, the calculation of the HEI-2015 score was based on the self-report of a one-day food consumption record. On the other hand, polyphenol consumption was determined through a retrospective consumption frequency questionnaire. The accuracy of the consumption frequency data depends on the information declared by the participants. As in studies in both national and international literature, when calculating daily polyphenol intake, the Phenol-Explorer database was used (19) as there is no polyphenol database that provides analysis results of foods specific to Türkiye. The polyphenol content of foods is affected by many factors, such as the conditions in which the food is grown, geographical region, season, post-harvest processing, and storage conditions.

Therefore, it is difficult to precisely determine the polyphenol intake of individuals, although international databases are used, there is a need for geographically specific food composition databases developed with a more consistent

approach. Although the results obtained in this study shed light on future studies on adherence to the MD, diet quality, and health profile of women with endometriosis, prospective studies with a more comprehensive sample and design are needed.

CONCLUSION

In conclusion, this study, which evaluated the nutritional status and health profile of women with endometriosis, draws attention to the importance of nutrition in the treatment of endometriosis. The findings indicate that women with endometriosis have low adherence to the MD a healthy dietary model and low diet quality. Nutrition should be emphasized, especially in the management of pain, which is a common complaint in endometriosis. At this point, nutrition education and counseling provided through a multidisciplinary approach to women with endometriosis after diagnosis will be an important step towards improving quality of life. More research is needed to develop nutritional strategies to improve the health and well-being of women with endometriosis.

ETHICS

Ethics Committee Approval: Acıbadem Mehmet Ali Aydınlar University and Acıbadem Healthcare Institutions Medical Research Ethics Committee (ATADEK) approval was received (approval no: 2023-13/447, date: 17.08.2023).

Informed Consent: All patients were informed about the study, and written informed consent was obtained for the protocol.

Acknowledgement

This study was carried out within the scope of Tuğçe Deveci's master's thesis in 2024 and was presented in 13th National Healthy Life Congress as an oral presentation. The authors would like to thank all participated in the study and to the research and the health professionals who facilitated the conduct of the study.

FOOTNOTES

Author Contributions

Concept: T.D., N.Ç.B., Desing: T.D., N.Ç.B., P.Y.B., Data Collection or Proccessing: T.D., N.Ç.B., P.Y.B., Analysis or Interpretation: T.D., N.Ç.B., Literature Search: T.D., N.Ç.B., Writing: T.D., N.Ç.B., P.Y.B.

Conflict of Interest: The authors declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Financial Disclosure: This research was supported by the Acıbadem Mehmet Ali Aydınlar University Scientific Research Projects Commission (ABAPKO) [TYL-2023-136].

REFERENCES

1. Bougie O, Yap MI, Sikora L, Flaxman T, Singh S. Influence of race/ethnicity on prevalence and presentation of endometriosis: a systematic review and meta-analysis. BJOG. 2019;126:1104-15.
2. World Health Organization. Endometriosis [Internet]. 2023 [cited 2024 August 19]. Available from: https://www.who.int/news-room/fact-sheets/detail/endometriosis
3. Nnoaham KE, Hummelshoj L, Webster P, d'Hooghe T, de Cicco Nardone F, de Cicco Nardone C, et al. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. Fertil Steril. 2011;96:366-373.e8.
4. Howard FM. Endometriosis and mechanisms of pelvic pain. J Minim Invasive Gynecol. 2009;16:540-50.
5. Sinaii N, Cleary SD, Ballweg ML, Nieman LK, Stratton P. High rates of autoimmune and endocrine disorders, fibromyalgia, chronic fatigue syndrome and atopic diseases among women with endometriosis: a survey analysis. Hum Reprod. 2002;17:2715-24.
6. Sepulcri Rde P, do Amaral VF. Depressive symptoms, anxiety, and quality of life in women with pelvic endometriosis. Eur J Obstet Gynecol Reprod Biol. 2009;142:53-6.
7. Holoch KJ, Savaris RF, Forstein DA, Miller PB, Higdon HL, Likes CE, et al. Coexistence of polycystic ovary syndrome and endometriosis in women with infertility. J Endometr Pelvic Pain Disord. 2014;6:79-83.
8. Bogusz A, Górnicka M. Low diet quality and nutritional knowledge in women with endometriosis: a pilot study. Healthcare (Basel). 2024;12:673.
9. Casas R, Sacanella E, Estruch R. The immune protective effect of the Mediterranean diet against chronic low-grade inflammatory diseases. Endocr Metab Immune Disord Drug Targets. 2014;14:245-54.
10. Cirillo M, Argento FR, Becatti M, Fiorillo C, Coccia ME, Fatini C. Mediterranean diet and oxidative stress: a relationship with pain perception in endometriosis. Int J Mol Sci. 2023;24:14601.
11. Piecuch M, Garbicz J, Waliczek M, Malinowska-Borowska J, Rozentryt P. I am the 1 in 10-what should I eat? A research review of nutrition in endometriosis. Nutrients. 2022;14:5283.
12. Minihane AM, Vinoy S, Russell WR, Baka A, Roche HM, Tuohy KM, et al. Low-grade inflammation, diet composition and health: current research evidence and its translation. Br J Nutr. 2015;114:999-1012.
13. Acar MN. Yetişkin bireylerde polifenol tüketim sıklığının ruh haline etkisi ve bireylerin polifenol tüketim alışkanlıklarının Covid-19 pandemi sürecindeki değişiminin incelenmesi [dissertation] Karabük: Karabük University; 2022.
14. Tassinari V, Smeriglio A, Stillittano V, Trombetta D, Zilli R, Tassinari R, et al. Endometriosis treatment: role of natural polyphenols as anti-inflammatory agents. Nutrients. 2023;15:2967.
15. Asencio FA, Fins RJP, Mitie CK, Ussia A, Wattiez A, Ribeiro HS, et al. Segmental rectum resection for deep endometriosis and excision similarly improve sexual function and pain. Clin Pract. 2023;13:780-90.
16. World Health Organization. (2011). Waist circumference and waist-hip ratio: report of a WHO expert consultation, 1st ed. Geneva, 2008. Available from: https://iris.who.int/handle/10665/44583
17. Martínez-González MA, García-Arellano A, Toledo E, Salas-Salvadó J, Buil-Cosiales P, Corella D, et al. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects: the PREDIMED trial. PLoS One. 2012;7:e43134.
18. Bekar C, Goktas Z. Validation of the 14-item Mediterranean diet adherence screener. Clin Nutr ESPEN. 2023;53:238-43.
19. Rothwell JA, Perez-Jimenez J, Neveu V, Medina-Remón A, M'hiri N, García-Lobato P, Manach C, et al. Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on

the effects of food processing on polyphenol content. Database (Oxford). 2013;2013:bat070.

20. Reedy J, Lerman JL, Krebs-Smith SM, Kirkpatrick SI, Pannucci TE, Wilson MM, et al. Evaluation of the Healthy Eating Index-2015. J Acad Nutr Diet. 2018;118:1622-33.
21. National Cancer Institute. Visualizing and interpreting Healthy Eating Index (HEI) scores [Internet]. Bethesda (MD): National Cancer Institute; [cited 2024 August 19]. Available from: https://epi.grants.cancer.gov/hei/interpret-visualize-hei-scores.html
22. Jones G, Jenkinson C, Kennedy S. Evaluating the responsiveness of the endometriosis health profile questionnaire: the EHP-30. Qual Life Res. 2004;13:705-13.
23. Şahin S. Endometriozisli kadınlarda yaşam kalitesinin belirlenmesi [master's thesis]. İstanbul: Haliç Üniversitesi Sağlık Bilimleri Enstitüsü; 2010. p. 5-73.
24. Türk İstatistik Kurumu. İstatistiklerle Kadın, 2023 [internet]. Ankara: Türkiye İstatistik Kurumu; 2023 [2023; cited 2024 Aug 22]. Available from: https://data.tuik.gov.tr/Bulten/Index?p=Istatistiklerle-Kadin-2023-53675
25. Mahmutoğlu M. Endometriyozisli kadınların beslenme alışkanlıkları ile yaşam kaliteleri arasındaki ilişkinin değerlendirilmesi [master's thesis]. İstanbul: Biruni University; 2023.
26. Hong J, Yi KW. What is the link between endometriosis and adiposity? Obstet Gynecol Sci. 2022;65:227-33.
27. Holdsworth-Carson SJ, Dior UP, Colgrave EM, Healey M, Montgomery GW, Rogers PA, et al. The association of body mass index with endometriosis and disease severity in women with pain. J Endometr Pelvic Pain Disord. 2018;10:79-87.
28. T.C. Sağlık Bakanlığı. Türkiye Beslenme Sağlığı Araştırması (TBSA). Ankara: T.C. Sağlık Bakanlığı Halk Sağlığı Genel Müdürlüğü; 2019. Report No: 1132.
29. Cramer DW, Wilson E, Stillman RJ, Berger MJ, Belisle S, Schiff I, et al. The relation of endometriosis to menstrual characteristics, smoking, and exercise. JAMA. 1986;255:1904-8.
30. T.C. Sağlık Bakanlığı. Türkiye Beslenme Rehberi (TÜBER) 2022. Ankara: T.C. Sağlık Bakanlığı Halk Sağlığı Genel Müdürlüğü; 2022. Report No: 1031.
31. Krebs-Smith SM, Pannucci TE, Subar AF, Kirkpatrick SI, Lerman JL, Tooze JA, et al. Update of the Healthy Eating Index: HEI-2015. J Acad Nutr Diet. 2018;118:1591-602.
32. Simopoulos AP. The importance of the ratio of omega-6/omega-3 essential fatty acids. Biomed Pharmacother. 2002;56:365-79.
33. Dünya Sağlık Örgütü. Uluslararası fonksiyon, engellilik ve sağlık sınıflaması [Internet]. Cenevre: Dünya Sağlık Örgütü; 2001. Available from: https://iris.who.int/bitstream/handle/10665/42407/9241545429.pdf
34. Del Bo' C, Bernardi S, Marino M, Porrini M, Tucci M, Guglielmetti S, et al. Systematic review on polyphenol intake and health outcomes: is there sufficient evidence to define a health-promoting polyphenol-rich dietary pattern? Nutrients. 2019;11:1355.
35. Jia SZ, Leng JH, Sun PR, Lang JH. Translation and psychometric evaluation of the simplified Chinese-version Endometriosis Health Profile-30. Hum Reprod. 2013;28:691-7.
36. Pontoppidan K, Olovsson M, Grundström H. Clinical factors associated with quality of life among women with endometriosis: a cross-sectional study. BMC Womens Health. 2023;23:551.
37. Marí-Alexandre J, García-Oms J, Agababyan C, Belda-Montesinos R, Royo-Bolea S, Varo-Gómez B, et al. Toward an improved assessment of quality of life in endometriosis: evaluation of the Spanish version of the Endometriosis Health Profile 30. J Psychosom Obstet Gynaecol. 2022;43:251-7.
38. Jones G, Jenkinson C, Taylor N, Mills A, Kennedy S. Measuring quality of life in women with endometriosis: tests of data quality, score reliability, response rate and scaling assumptions of the Endometriosis Health Profile Questionnaire. Hum Reprod. 2006;21:2686-93.
39. Warwick R, Joseph S, Cordle C, Ashworth P. Social support for women with chronic pelvic pain: what is helpful from whom? Psychology & Health. 2004;19:117-34.
40. Volker C, Mills J. Endometriosis and body image: comparing people with and without endometriosis and exploring the relationship with pelvic pain. Body Image. 2022;43:518-22.
41. Pazmany E, Bergeron S, Van Oudenhove L, Verhaeghe J, Enzlin P. Body image and genital self-image in pre-menopausal women with dyspareunia. Arch Sex Behav. 2013;42:999-1010.
42. Tripoli TM, Sato H, Sartori MG, de Araujo FF, Girão MJ, Schor E. Evaluation of quality of life and sexual satisfaction in women suffering from chronic pelvic pain with or without endometriosis. J Sex Med. 2011;8:497-503.
43. Calvi C, Sherman KA, Pham D. Loneliness and perceived social support in endometriosis: the roles of body image disturbance and anticipated stigma. Int J Behav Med. 2024;31:433-44.
44. Ilhan M, Ali Z, Khan IA, Taştan H, Küpeli Akkol E. Bioactivity-guided isolation of flavonoids from Urtica dioica L. and their effect on endometriosis rat model. J Ethnopharmacol. 2019;243:112100.
45. Taghipour Z, Bahmanzadeh M, Rahimi R. The effects of clove and its constituents on reproductive system: a comprehensive review. Reprod Sci. 2023;30:2591-614.
46. Talib WH, AlHur MJ, Al Naimat S, Ahmad RE, Al-Yasari AH, Al-Dalaeen A, et al. Anticancer effect of spices used in Mediterranean diet: preventive and therapeutic potentials. Front Nutr. 2022;9:905658.
47. Ott J, Nouri K, Hrebacka D, Gutschelhofer S, Huber J, Wenzl R. Endometriosis and nutrition-recommending a Mediterranean diet decreases endometriosis-associated pain: an experimental observational study. J Aging Res Clin Pract. 2012;1:162-6.
48. Ortolá R, García-Esquinas E, Sotos-Prieto M, Struijk EA, Caballero FF, Lopez-Garcia E, et al. Mediterranean diet and changes in frequency, severity, and localization of pain in older adults: the seniors-ENRICA cohorts. J Gerontol A Biol Sci Med Sci. 2022;77:122-30.
49. Zhou YF, Song XY, Wu J, Chen GC, Neelakantan N, van Dam RM, et al. Association between dietary patterns in midlife and healthy ageing in Chinese adults: the Singapore Chinese health study. J Am Med Dir Assoc. 2021;22:1279-86.
50. Dougan MM, Fest S, Cushing-Haugen K, Farland LV, Chavarro J, Harris HR, et al. A prospective study of dietary patterns and the incidence of endometriosis diagnosis. Am J Obstet Gynecol. 2024;231:443.e1-10.
51. Moore JS, Gibson PR, Perry RE, Burgell RE. Endometriosis in patients with irritable bowel syndrome: specific symptomatic and demographic profile, and response to the low FODMAP diet. Aust N Z J Obstet Gynaecol. 2017;57:201-5.
52. Vennberg Karlsson J, Patel H, Premberg A. Experiences of health after dietary changes in endometriosis: a qualitative interview study. BMJ Open. 2020;10:e032321.
53. Marziali M, Venza M, Lazzaro S, Lazzaro A, Micossi C, Stolfi VM. Gluten-free diet: a new strategy for management of painful endometriosis related symptoms? Minerva Chir. 2012;67:499-504.
54. Nirgianakis K, Egger K, Kalaitzopoulos DR, Lanz S, Bally L, Mueller MD. Effectiveness of dietary interventions in the treatment of endometriosis: a systematic review. Reprod Sci. 2022;29:26-42.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):430-435
DOI: 10.4274/BMJ.galenos.2025.2025.1-22

Research

Does the Menstrual Cycle Affect Skin Puncture Pain During Spinal Anesthesia? Prospective, Observational Study

Menstural Siklus Spinal Anestezi Sırasında Deri Delinme Ağrısını Etkiler mi? Prospektif, Gözlemsel Çalışma

Mahmut Sami Tutar[1], Rabia Korkmaz[1], Şule Yaşayacak[1], Mustafa Atcı[2], Muhammed Halit Satıcı[1], Bülent Hanedan[1], Munise Yıldız[1], Betül Kozanhan[1]

[1]Konya City Hospital, Clinic of Anesthesiology and Reanimation, Konya, Türkiye

[2]Kozan State Hospital, Clinic of Anesthesiology and Reanimation, Adana, Türkiye

ABSTRACT

Objective: Skin puncture pain during spinal anesthesia is a significant concern for many patients and may lead to reluctance to choose this anesthetic method, affecting patient satisfaction and acceptance of the procedure. This study aimed to evaluate the effect of menstrual cycle phase on needle-insertion pain during spinal anesthesia.

Methods: This study included 80 female patients, aged 18-45 years, who underwent spinal anesthesia using a 25G Quincke-tipped needle at the L3-L4 or L4-L5 vertebral levels in the sitting position. Patients were divided into two groups: follicular phase and luteal phase. Upon completion of the procedure, pain during needle insertion was assessed using the Numerical Rating Scale (NRS).

Results: No statistically significant differences were observed between menstrual cycle phases and NRS scores during needle insertion (p=0.804). Additionally, no significant correlation was found between NRS scores and body mass index (r=-0.109, p=0.335). However, a weak but statistically significant negative correlation was identified between NRS scores and age (r=-0.246, p=0.028).

Conclusion: The findings indicate that the phase of the menstrual cycle does not influence needle insertion pain during spinal anesthesia and emphasize that other factors, such as age, might have a minor impact on pain perception.

Keywords: Oestrogen, menstrual cycle, spinal anesthesia, skin puncture pain

ÖZ

Amaç: Spinal anestezi sırasında deri delme ağrısı birçok hasta için önemli bir endişe kaynağıdır ve bu anestezi yöntemini seçmede isteksizliğe yol açarak hasta memnuniyetini ve prosedür kabulünü etkileyebilir. Bu çalışma, adet siklusunun spinal anestezi sırasında iğne yerleştirme ağrısı üzerindeki etkisini değerlendirmeyi amaçladı.

Gereç ve Yöntem: Bu çalışmaya, oturma pozisyonunda L3-L4 veya L4-L5 vertebra seviyelerinde 25G Quinke uçlu iğne kullanılarak spinal anestezi uygulanan 18-45 yaş arası 80 kadın hastalar dahil edildi. Hastalar foliküler ve luteal faz olmak üzere iki gruba ayrıldı. İşlem tamamlandıktan sonra, iğne yerleştirme sırasındaki ağrı Sayısal Derecelendirme Ölçeği (SDÖ) kullanılarak değerlendirildi.

Bulgular: İğne yerleştirme sırasında adet siklusu ve SDÖ skorları arasında istatistiksel olarak anlamlı bir fark gözlenmedi (p=0,804). Ek olarak, SDÖ skorları ve vücut kitle indeksi arasında anlamlı bir korelasyon bulunmadı (r=-0,109, p=0,335). Ancak, SDÖ puanları ile yaş arasında zayıf ancak istatistiksel olarak anlamlı bir negatif korelasyon tespit edildi (r=-0,246, p=0,028).

Sonuç: Bulgular, adet siklusunun spinal anestezi sırasında iğne batırma ağrısını etkilemediğini, yaş gibi diğer faktörlerin ağrı algısı üzerinde küçük bir etkiye sahip olabileceğini vurgulamaktadır.

Anahtar Kelimeler: Östrojen, adet siklusu, spinal anestezi, deri delinme ağrısı

Address for Correspondence: Muhammed Halit Satıcı, MD, Konya City Hospital, Clinic of Anesthesiology and Reanimation, Konya, Türkiye
E-mail: halit_satici@hotmail.com **ORCID ID:** orcid.org/0009-0005-0652-3624

Cite as: Tutar MS, Korkmaz R, Yaşayacak Ş, Atcı M, Satıcı MH, Hanedan B, et al. Does the menstrual cycle affect skin puncture pain during spinal anesthesia? Prospective, observational study. Med J Bakirkoy. 2025;21(4):430-435

Received: 27.01.2025
Accepted: 13.05.2025
Publication Date: 18.12.2025

INTRODUCTION

Pain perception varies significantly among individuals and is influenced by biological, psychological, and social factors. Among these, sex-based differences in pain sensitivity have been extensively studied. Women consistently demonstrate greater pain sensitivity than men, attributed to variations in the endogenous opioid system, hormonal fluctuations, and psychosocial factors (1). Hormonal cycles unique to women, particularly the perimenstrual phase, are associated with heightened pain sensitivity, systemic inflammatory responses, and conditions such as vocal cord edema and peripheral edema (2). Additionally, women are more likely to experience postoperative complications, including nausea, vomiting, sleep disturbances, reduced intravenous anesthetic requirements, and delayed gastric emptying, particularly during the luteal phase, when cyclical fluctuations in hormones such as progesterone occur (2-4).

Despite evidence linking menstrual cycles to pain and physiological changes, limited research has explored their impact on procedural pain. Spinal anesthesia, a cornerstone in neuraxial blocks for lower body surgeries, is effective and widely used (5). However, the potential influence of the menstrual cycle on pain during spinal needle insertion remains underexplored. This study aims to investigate the relationship between the menstrual cycle and pain during spinal needle insertion in women of reproductive age undergoing elective surgery, with the aim of providing insights to improve patient comfort and guide clinical practice.

METHODS

Ethics Approval and Registration

Ethical approval for the study was obtained from the Ethics Committee of University of Health Sciences Türkiye, Hamidiye Scientific Research Ethics Committee (approval no: 16/2, date: 17.06.2022). It was registered on ClinicalTrials.gov (identifier: NCT05481255). The study was conducted using the principles of the Declaration of Helsinki and adhered to CONSORT guidelines (6,7). Written informed consent was obtained from all participants.

Patient Population and Inclusion/Exclusion Criteria

The inclusion criteria for the study consist of female patients aged 18-45 years who are undergoing a single-attempt spinal anesthesia with a 25G Quincke-tipped needle at the L3-L4 or L4-L5 vertebral levels in the sitting position, who are classified as American Society of Anesthesiologists physical status I or II, and who have had regular menstrual cycles for at least the past six months. The exclusion criteria include pregnancy, use of oral contraceptives, an anxiety subscale score on the Hospital Anxiety and Depression Scale (HADS) of 10 or higher, or a depression subscale score of 7 or higher (8,9).

Study Protocol

The patients' anxiety and depression levels were assessed using the 14-item HADS, which was administered by one of the researchers (M.Y.) on the morning of surgery prior to spinal anesthesia. Each participant was asked about the first day of their last menstrual period and the duration of their menstrual cycle to calculate the specific day within that cycle. Based on this information, patients were categorized into two groups: the follicular-phase group (cycle days 8-10) and the luteal-phase group (cycle days 18-20). Spinal anesthesia was administered by an experienced anesthesiologist (B.H.) who was not involved in either the data collection or group allocation processes. After the procedure, a second researcher (Ş.Y.), who was blinded to the patient's menstrual cycle phase, asked the patient to rate the pain during the skin puncture using the Numerical Rating Scale (NRS), where 0 indicated no pain and 10 represented the worst imaginable pain. Finally, a third researcher (B.K.) independently verified the menstrual cycle phase by re-evaluating the date of the last menstrual period and cycle duration provided by the patient.

Statistical Analysis

Sample size analysis was performed based on the patient's NRS score for skin-puncture pain during the spinal procedure. A 30% score difference between the groups was considered significant. Using a Cohen's d effect size of 0.670 in the independent-groups t-test model based on pilot data, it was calculated that 36 patients in each group (a total of 72 patients) should participate in the study to achieve 80% power with a maximum 5% type I error. Accounting for a 10% dropout rate, the required sample size was 80 participants.

The data obtained in the study were analyzed using the International Business Machines statistical package for the social sciences (IBM SPSS Inc., Chicago, IL, USA), version 26.0. The normality of the data distribution was evaluated using the Shapiro-Wilk test. Continuous variables were presented as mean and standard deviation or median (25^{th}-75^{th} percentiles), while categorical variables were expressed as numbers and percentages. For the analysis of continuous variables, the Independent samples t-test was employed when parametric test assumptions were met; otherwise, the Mann-Whitney U test was utilised. The relationship between NRS scores and patient characteristics was analysed using Spearman's rho. Statistical significance was accepted at $p<0.05$.

RESULTS

In this study, 105 cases were screened for eligibility based on the inclusion criteria. Fifteen patients were excluded for lack of consent to participate, and an additional ten were excluded for exceeding the cut-off on the HADS. Specifically, these patients had elevated HADS scores prior to spinal anesthesia. Consequently, the remaining 80 patients were included in the final analysis, with 40 allocated to the follicular-phase group and 40 to the luteal-phase group. A flowchart outlining the study process is presented in Figure 1.

An evaluation of patient characteristics based on menstrual cycle phase revealed no statistically significant differences in age, height, weight, or body mass index (BMI) between the follicular-phase and luteal-phase groups (p>0.05). Similarly, there were no significant differences between groups in preoperative HADS anxiety and depression scores (p=0.464 and p=0.661, respectively). The NRS scores for needle insertion pain were also comparable between groups (p=0.804; Table 1).

Further analysis of the relationship between NRS scores and patient characteristics revealed a weak but statistically significant negative correlation between NRS scores and age (r=-0.246, p=0.028). However, no significant correlations were observed between NRS scores and height, HADS anxiety scores, or HADS depression scores (r=0.04, p=0.725; r=0.085, p=0.455; and r=0.135, p=0.233, respectively). Similarly, weight and BMI showed weak negative correlations with NRS scores that were not statistically significant (weight: r=-0.082, p=0.469; BMI: r=-0.109, p=0.335) (Table 2).

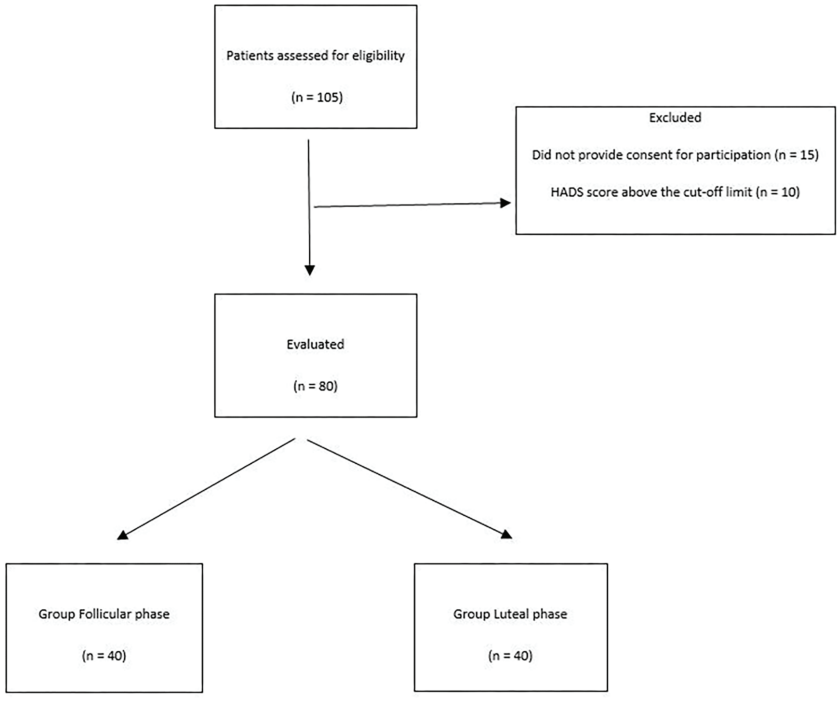


Figure 1. The flowchart of the study
HADS: Hospital Anxiety and Depression Scale

Table 1. Evaluation of patient characteristics according to menstrual cycle

Characteristic	Folicular phase (n=40)	Luteal phase (n=40)	p-value
Age	31.38±6.63	33.53±6	0.133
Height (cm)	162.62±5.57	163.80±6.59	0.392
Weight (kg)	69.85±12.62	69.25±11.65	0.826
BMI (kgm^2)	26.39±4.40	25.92±4.69	0.649
HADS anxiety score	3 (2-5)	3 (2-5)	0.464
HADS depression score	3 (2-5)	3 (2-4.5)	0.661
NRS score	2 (1-3.5)	2 (1-3)	0.804

Statistical test applied: independent samples t-test and the Mann-Whitney U test. Values are presented as mean±standard deviation or median (25-75 percentiles).
BMI: Body mass index, HADS: Hospital Anxiety and Depression Scale and NRS: Numerical Rating Scale

Table 2. Relationship between NRS score and patient characteristics

NRS score		
Characteristic	**r-value**	**p-value**
Age	-0.246	0.028
Height (cm)	0.04	0.725
Weight (kg)	-0.082	0.469
BMI (kgm^2)	-0.109	0.335
HADS anxiety score	0.085	0.455
HADS depression score	0.135	0.233

Statistical test applied: Spearman rho correlation test.
BMI: Body mass index, HADS: Hospital Anxiety and Depression Scale, NRS: Numerical Rating Scale and r: correlation coefficient

DISCUSSION

This study demonstrated that the menstrual cycle phase did not significantly influence pain intensity measured by the NRS. This finding suggests that hormonal fluctuations associated with the menstrual cycle may not play a clinically relevant role in modulating acute procedural pain during spinal anesthesia. Additionally, a weak but statistically significant negative correlation was observed between NRS scores and age, indicating that older patients tend to report slightly lower pain levels.

Spinal anesthesia is a widely used neuraxial technique in which local anesthetic is injected into the cerebrospinal fluid within the subarachnoid space of the lumbar spine. While it is generally well tolerated, pain during spinal needle insertion remains a concern for some patients, especially those with increased pain sensitivity. Spinal needle insertion activates nociceptive pathways by stimulating peripheral nerve endings in the skin and underlying tissues. Pain perception during such procedures is primarily mediated by A-delta and C fibers, which transmit noxious stimuli to the dorsal horn of the spinal cord (10,11). These impulses are then relayed through second-order neurons via the spinothalamic tract to higher brain regions, including the thalamus and somatosensory cortex, where pain is consciously perceived. The periaqueductal gray matter and descending inhibitory pathways also play a key role in modulating the pain response (10,11).

Hormonal fluctuations during the menstrual cycle, particularly changes in estrogen and progesterone levels, may influence the sensitivity of these pathways. Estrogen is known to modulate both peripheral and central nociceptive processing by interacting with opioid receptors, N-methyl-D-aspartate receptors, and serotonin pathways (12). Some studies suggest that high or fluctuating estrogen levels may enhance pain perception, while others indicate a protective analgesic effect depending on the context (12-14). Understanding these complex hormonal influences is essential when evaluating procedural pain in women during different phases of the menstrual cycle.

Pain associated with spinal needle puncture is a critical concern. These concerns highlight the importance of addressing procedural pain and improving patient education to alleviate fears related to spinal anesthesia (15).

Skin puncture can cause both physical and psychological distress, with multiple factors influencing pain perception. Physiologically, puncture activates nociceptive pathways, leading to the sensation of pain. Psychological factors such as fear of needles, anticipation of pain, and procedural unpredictability play significant roles. Fear and anxiety can heighten pain perception by activating the body's stress response, which amplifies the central nervous system's sensitivity. Moreover, pain experienced during a procedure can reinforce anxiety, creating a cycle of heightened distress in future interventions (16). It is well-documented that women generally exhibit lower pain thresholds compared to men (17,18). Women also report higher rates of discomfort and demonstrate reduced pain tolerance across various settings (17). These differences arise from biological and neurophysiological variability in nociceptive processing between the sexes. Females may display heightened sensitivity to experimentally induced pain and increased temporal summation of mechanically evoked pain, indicating a more pronounced response to repeated or sustained nociceptive stimuli (19).

One explanation for these differences lies in the influence of sex hormones, such as oestrogen and progesterone, which play a critical role in modulating pain perception. Oestrogen, in particular, interacts with opioid and serotonin receptors, impacting both the peripheral and central pain pathways. These hormonal fluctuations may contribute to variations in pain sensitivity at different points in the menstrual cycle, as well as to the increased prevalence of specific pain syndromes in women. For instance, research has identified a potential link between hormonal changes and the occurrence of particular types of headaches, such as migraines. Oestrogen's role in regulating vascular tone, inflammation, and nociception may explain its involvement in triggering or exacerbating headache disorders (20).

These findings highlight the importance of considering sex-based differences in pain perception and the role of hormonal influences when developing personalised pain management strategies for men and women. Numerous studies have investigated the relationship between the menstrual cycle and the prevalence or intensity of migraine

headaches, given the well-established influence of hormonal fluctuations on pain perception (21,22). However, findings from these studies have been inconsistent, reflecting the complexity of hormonal interactions with pain mechanisms (20).

Cao et al. (23) highlighted that oestrogen could enhance visceral pain sensitivity by modulating peripheral and central nociceptive pathways. This suggests that elevated oestrogen levels during certain phases of the menstrual cycle may increase pain perception in specific contexts. Conversely, Maleki et al. (24) proposed that decreased oestrogen levels, combined with elevated hypothalamic prostaglandin levels during the luteal phase, might trigger migraine episodes.

Tutar et al. (25) conducted a retrospective study of patients undergoing spinal anesthesia and found no significant.

Although hormonal fluctuations during the menstrual cycle have long been hypothesized to modulate nociceptive sensitivity, the findings of the present study did not reveal a significant association between the menstrual cycle and procedural pain during spinal anesthesia. This aligns with previous research suggesting that cyclical hormonal changes may not exert a clinically relevant effect on acute nociceptive responses, particularly in the context of localised, short-duration interventions such as needle penetration. The absence of a significant difference between the follicular and luteal phases indicates that factors other than hormonal status, such as individual pain thresholds, psychological state (e.g., anxiety), or procedural technique, may play a more substantial role in shaping pain perception during neuraxial anesthesia. Consequently, these results suggest that routine consideration of the menstrual cycle in anesthetic planning for women of reproductive age may not be necessary. However, the broader influence of sex hormones on pain remains a complex and multifactorial topic, and the present findings should not preclude individualized patient assessment. Instead, they highlight the importance of tailoring perioperative care to patient-specific characteristics rather than relying solely on biological cycles that may have limited predictive value in acute procedural settings.

Anxiety can influence pain perception through neurotransmitters such as gamma-aminobutyric acid, serotonin, noradrenaline, oxytocin, and endocannabinoids, which regulate anxiety and nociceptive pathways (26). Despite this connection, no significant correlation was identified between anxiety and needle penetration pain.

Age-related differences in pain perception remain inconsistent in the literature. While some studies suggest increased pain sensitivity in older adults due to changes in central processing, others propose a decline in pain sensitivity with age, possibly related to reduced nociceptive receptor density or slower nerve conduction (27,28). In this study, a weak but significant negative correlation between age and needle penetration pain was observed, indicating that older patients may experience slightly less pain. This finding aligns with research suggesting reduced nociceptive sensitivity in ageing populations and highlights the importance of considering demographic factors in procedural pain management.

Study Limitations

This study has several limitations. First, the oestrogen and progesterone levels of the patients were not measured using quantitative methods, which limits the ability to directly correlate hormonal fluctuations with pain perception. Second, as spinal anesthesia at the institution was predominantly performed using 25-gauge needles, this study exclusively included patients who underwent the procedure with this needle size.

Another limitation of this study is the lack of information regarding patients' prior experiences with spinal or epidural anesthesia. Since pain perception and behavioral responses can be shaped by previous exposure and learning, this may have influenced individuals' pain reporting.

CONCLUSION

In conclusion, this study demonstrated that the menstrual cycle phase did not significantly influence needle penetration pain during spinal anesthesia in women of reproductive age. While age exhibited a weak negative correlation with pain perception, the clinical significance of this finding requires further exploration. By advancing understanding of the factors affecting procedural pain, this research underscores the importance of personalized approaches to anesthesia care, ultimately aiming to improve patient satisfaction and outcomes.

ETHICS

Ethics Committee Approval: Ethical approval for the study was obtained from the Ethics Committee of University of Health Sciences Türkiye, Hamidiye Scientific Research Ethics Committee (approval no: 16/2, date: 17.06.2022).

Informed Consent: Patients provided both written and verbal consent.

FOOTNOTES

Authorship Contributions

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

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

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

REFERENCES

1. Bartley EJ, Fillingim RB. Sex differences in pain: a brief review of clinical and experimental findings. Br J Anaesth. 2013;111:52-8.
2. Kurdi MS, Ramaswamy AH. Does the phase of the menstrual cycle really matter to anaesthesia? Indian J Anaesth. 2018;62:330.
3. Tutar MS, Sekmenli N, Erol G, Karaarslan E, Yildiz M, Kozanhan B. Does the menstrual phase affect gastric emptying? Prospective, observational study. Eur Rev Med Pharmacol Sci. 2023;27:1989-95.
4. Clarke BL, Khosla S. Female reproductive system and bone. Arch Biochem Biophys. 2010;503:118-28.
5. Wulf HF. The centennial of spinal anesthesia. Anesthesiology. 1998;89:500-6.
6. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310:2191-4.
7. Moher D, Hopewell S, Schulz KF. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c869.
8. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983; 67:361-70.
9. Aydemir Ö, Guvenir T, Kuey L, Kultur S. Validity and reliability of Turkish version of hospital anxiety and depression scale. Turk Psikiyatri Derg. 1997;8:280-7.
10. Dowlati E. Spinal cord anatomy, pain, and spinal cord stimulation mechanisms. Semin Spine Surg. 2017;29:136-46.
11. Wahezi SE, Hanstein R, Spray DC. Understanding the pathways of pain. In: Sharan A, Tang SY, Vaccaro AR, editors. Basic Science of Spinal Diseases. New Delhi: Jaypee Brothers Medical Publishers; 2013. p. 95-118.
12. Aloisi AM. Gonadal hormones and sex differences in pain reactivity. Clin J Pain. 2003;19:168-74.
13. Craft RM. Modulation of pain by estrogens. Pain. 2007;132:S3-12.
14. Fillingim RB, King CD, Ribeiro-Dasilva MC, Rahim-Williams B, Riley JL. Sex, gender, and pain: a review of recent clinical and experimental findings. J Pain. 2009;10:447-85.
15. Gajraj NM, Sharma S, Souter AJ, Pole Y, Sidawi E. A survey of patients who refuse regional anesthesia. Anesth Analg. 1994;78:S126.
16. Kumar S, Gautam SKS, Gupta D, Agarwal A, Dhirraj S, Khuba S. The effect of Valsalva maneuver in attenuating skin puncture pain during spinal anesthesia: a randomized controlled trial. Korean J Anesthesiol. 2016;69:27-31.
17. Vallerand AH, Polomano RC. The relationship of gender to pain. Pain Manag Nurs. 2000;1(Suppl 1):8-15.
18. Wandner LD, Scipio CD, Hirsh AT, Torres CA, Robinson ME. The perception of pain in others: how gender, race, and age influence pain expectations. J Pain. 2012;13:220-7.
19. Sarlani E, Farooq N, Greenspan JD. Gender and laterality differences in thermosensation throughout the perceptible range. Pain. 2003;106:9-18.
20. Sun LH, Zhang WX, Xu Q, Wu H, Jiao CC, Chen XZ. Estrogen modulation of visceral pain. J Zhejiang Univ Sci B. 2019;20:628-36.
21. MacGregor EA, Igarashi H, Wilkinson M. Headaches and hormones: subjective versus objective assessment. Headache Q Curr Treat Res. 1997;8:126-36.
22. MacGregor EA, Frith A, Ellis J, Aspinall L, Hackshaw A. Prevention of menstrual attacks of migraine: a double-blind placebo-controlled crossover study. Neurology. 2006;67:2159-63.
23. Cao DY, Bai G, Ji Y, Traub RJ. Epigenetic upregulation of metabotropic glutamate receptor 2 in the spinal cord attenuates oestrogen-induced visceral hypersensitivity. Gut. 2015;64:1913-20.
24. Maleki N, Linnman C, Brawn J, Burstein R, Becerra L, Borsook D. Her versus his migraine: multiple sex differences in brain function and structure. Brain. 2012;135:2546-59.
25. Tutar MS, Korucu D, Ciftci C, Kozanhan B. The relationship between serum estrogen concentration and post-dural puncture headache: a retrospective study. Eur Rev Med Pharmacol Sci. 2023;27:2250-55.
26. Gross C, Hen R. The developmental origins of anxiety. Nat Rev Neurosci. 2004;5:545-52.
27. Rittger H, Rieber J, Breithardt OA, Dücker M, Schmidt M, Abbara S, et al. Influence of age on pain perception in acute myocardial ischemia: a possible cause for delayed treatment in elderly patients. Int J Cardiol. 2011;149:63-7.
28. Lautenbacher S, Kunz M, Strate P, Nielsen J, Arendt-Nielsen L. Age effects on pain thresholds, temporal summation and spatial summation of heat and pressure pain. Pain. 2005;115:410-8.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):436-443
DOI: 10.4274/BMJ.galenos.2025.2025.9-7

Research

Evaluation of the Association Between the CALLY Index, Functional Class, and Mortality in Patients with Heart Failure with Reduced Ejection Fraction

Düşük Ejeksiyon Fraksiyonlu Kalp Yetersizliği Hastalarında CALLY İndeksi ile Fonksiyonel Sınıf ve Mortalite Arasındaki İlişkinin Değerlendirilmesi

Evliya Akdeniz, Dilay Karabulut

University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Cardiology, İstanbul, Türkiye

ABSTRACT

Objective: Heart failure with reduced ejection fraction (HFrEF) remains a major global health challenge. The New York Heart Association (NYHA) functional classification remains an indispensable instrument for evaluating symptom severity and carries recognized prognostic value in this patient population. Nevertheless, there is a growing need for accessible, objective markers that can help predict outcomes more accurately. The C-reactive protein-albumin-lymphocyte (CALLY) index, derived from C-reactive protein (CRP), albumin, and lymphocyte count, offers a composite measure that reflects systemic inflammation, nutritional status, and immune function. This study aimed to evaluate the association between the CALLY index, NYHA class, and all-cause of mortality in patients with HFrEF.

Methods: Patients diagnosed with HFrEF from January 2023 to April 2025 were retrospectively evaluated. Patients were categorized as survivors or non-survivors. Clinical characteristics, laboratory findings, and NYHA functional class were systematically compared between the two groups. To explore factors independently associated with mortality, both univariate and multivariate logistic regression analyses were conducted. Additionally, the prognostic performance of the CALLY index was assessed using receiver operating characteristic (ROC) curve analysis.

Results: Among 146 patients diagnosed with HFrEF, 29 (19.8%) experienced all-cause of mortality. Compared with survivors, those who died were older, had significantly lower left ventricular ejection fraction (LVEF), and showed higher levels of inflammatory and cardiac biomarkers, including CRP, troponin-T, and pro-brain natriuretic peptide. Advanced heart failure symptoms (NYHA class 3-4) were more common among non-survivors. Notably, the CALLY index—reflecting a combination of inflammation, nutritional status, and immune function—was markedly lower in the mortality group. In both univariate and multivariate logistic regression analyses, LVEF, NYHA class 3-4, and the CALLY index were independently associated with mortality. ROC curve analysis showed that the CALLY index had good predictive value for all-cause of mortality and moderate predictive value for NYHA class 3-4, supporting its potential as a practical tool in clinical risk assessment.

Conclusion: The CALLY index, which integrates markers of inflammation, nutrition, and immune status, was independently associated with all-cause of mortality and correlated with symptom severity in patients with HFrEF.

Keywords: HFrEF, CALLY index, mortality, NYHA

ÖZ

Amaç: Düşük ejeksiyon fraksiyonlu kalp yetersizliği (DEFKY), halen küresel ölçekte önemli bir sağlık sorunu olmaya devam etmektedir. New York Kalp Derneği (NYHA) fonksiyonel sınıflandırması, semptom şiddetinin değerlendirilmesinde vazgeçilmez bir araç olup, bu hasta grubunda prognostik değeri de kabul görmüştür. Bununla birlikte, sonuçları daha doğru tahmin etmeye yardımcı olabilecek erişilebilir ve objektif belirteçlere olan ihtiyaç giderek artmaktadır. C-reaktif protein (CRP), albümin ve lenfosit sayısından türetilen C-reaktif protein-albümin-lenfosit (CALLY) indeksi, sistemik enflamasyon, beslenme durumu ve bağışıklık fonksiyonunu yansıtan bileşik bir ölçüttür. Bu çalışma, DEFKY hastalarında, CALLY indeksi ile NYHA sınıfı ve tüm nedenlere bağlı mortalite arasındaki ilişkiyi değerlendirmeyi amaçlamıştır.

Address for Correspondence: Evliya Akdeniz, MD, University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Cardiology, İstanbul, Türkiye
E-mail: evliyakdeniz@gmail.com **ORCID ID:** orcid.org/0000-0002-4688-7992

Cite as: Akdeniz E, Karabulut D. Evaluation of the association between the CALLY index, functional class, and mortality in patients with heart failure with reduced ejection fraction. Med J Bakirkoy. 2025;21(4):436-443

Received: 03.10.2025
Accepted: 19.11.2025
Publication Date: 18.12.2025

ÖZ

Gereç ve Yöntem: Ocak 2023 ile Nisan 2025 arasında DEFKY tanısı alan hastalar retrospektif olarak incelendi. Hastalar sağ kalanlar ve mortalite grubuna ayrıldı. Klinik özellikler, laboratuvar bulguları ve NYHA fonksiyonel sınıfları iki grup arasında sistematik olarak karşılaştırıldı. Mortalitenin bağımsız belirleyicilerini saptamak amacıyla tek değişkenli ve çok değişkenli lojistik regresyon analizleri yapıldı. Ayrıca, CALLY indeksinin prognostik performansı receiver operating characteristic (ROC) eğrisi analizi ile değerlendirildi.

Bulgular: Çalışmaya dahil edilen 146 DEFKY hastasının 29'unda (%19,8) tüm nedenlere bağlı ölüm görüldü. Mortalite grubundaki hastalar, sağ kalanlara kıyasla daha yaşlı, daha düşük sol ventrikül ejeksiyon fraksiyonuna (LVEF) sahip ve CRP, troponin-T, pro-brain natriüretik peptid gibi enflamasyon ve kardiyak belirteçler açısından daha yüksek seviyelerdeydi. İleri kalp yetersizliği semptomları (NYHA sınıf 3-4) mortalite grubunda daha sık gözlendi. CALLY indeksi, enflamasyon, beslenme ve bağışıklık durumunu yansıtarak mortalite grubunda anlamlı derecede düşüktü. Hem tek değişkenli hem de çok değişkenli analizlerde LVEF, NYHA sınıf 3-4 ve CALLY indeksi mortalitenin bağımsız belirleyicileri olarak saptandı. ROC analizi CALLY indeksinin mortalite öngörmede iyi, NYHA sınıf 3,4 semptomlarını öngörmede ise orta düzeyde performans gösterdiğini ortaya koydu.

Sonuç: CALLY indeksi, enflamasyon, beslenme ve bağışıklık durumunu gösteren belirteçleri birleştirerek, DEFKY hastalarında tüm nedenlere bağlı mortalite ile bağımsız olarak ilişkili bulunmuş ve semptom şiddeti ile korelasyon göstermiştir.

Anahtar Kelimeler: DEFKY, CALLY indeksi, mortalite, NYHA

INTRODUCTION

Heart failure (HF), a multifaceted clinical condition afflicting 1-3% of the worldwide population, remains a prominent contributor to both mortality and morbidity. It imposes a considerable strain on healthcare systems worldwide, not only in terms of economic expenditure but also in terms of its profound psychosocial impact (1). Patients with HF are stratified into distinct subgroups according to left ventricular function. HF with reduced ejection fraction (HFrEF) is characterized by a left ventricular ejection fraction (LVEF) of ≤40%, accompanied by typical clinical manifestations or symptoms indicative of HF (2). HFrEF remains a substantial public health challenge due to its marked association with increased cardiovascular mortality and frequent HF-related hospitalizations, of which approximately 50% are attributable to this phenotype (3).

Although a decline in HF-related mortality has been observed in recent years, the incidence of cardiovascular death and HF-related hospitalization among patients with HFrEF remains substantially high-ranging from 21.8% to 26.5% in the PARADIGM-HF, despite significant advancements in medical technology and the introduction of novel therapeutic strategies (4-6). Since its introduction in 1928, the New York Heart Association (NYHA) functional classification has undergone several refinements; nevertheless, it remains the most widely used clinical framework for evaluating symptom severity in patients with HF (7). In addition to its utility in characterizing functional limitations, the NYHA classification possesses substantial prognostic significance. Considerable evidence has consistently demonstrated that advanced NYHA functional classification is an independent predictor of increased risk of adverse clinical outcomes, including all-cause of mortality and HF-related hospital admissions (8-11).

The C-reactive protein-albumin-lymphocyte (CALLY) index is a composite biomarker integrating C-reactive protein (CRP), serum albumin, and total lymphocyte count, and provides a comprehensive assessment of systemic inflammation, nutritional status, and immune function. It was first introduced by Iida et al. (12) as a prognostic tool for predicting survival in patients with hepatocellular carcinoma. Since its introduction, the CALLY index has attracted increasing attention in the field of cardiology, with several studies demonstrating its prognostic value across a broad spectrum of cardiovascular conditions. These findings imply that the index may serve as a useful, accessible marker to guide clinical decision-making and predict patient outcomes across diverse cardiac populations (13,14).

The current investigation was undertaken to elucidate the relationship between the CALLY index, NYHA functional classification, and overall mortality in a cohort of patients diagnosed with HFrEF. This investigation was designed to assess the prognostic utility of the CALLY index as a potential biomarker, with the overarching objective of augmenting risk stratification protocols and refining therapeutic decision-making processes within this clinically vulnerable population.

METHODS

Study Population

From January 2023 to April 2025, consecutive patients who presented to our hospital with symptoms of HF and were subsequently diagnosed with HFrEF were retrospectively included in the study. Eligibility criteria included individuals older than 18 years with a definitive diagnosis of HFrEF, based on clinical evaluation and echocardiographic assessment (LVEF ≤40%). Exclusion criteria included refusal to participate, incomplete echocardiographic data, acute

decompensated HF, HF with preserved ejection fraction (HFpEF) or HF with mildly reduced ejection fraction (HFmrEF), end-stage renal disease, active malignancy or active infection, chronic liver disease, chronic inflammatory diseases, and chronic hematological disorders.

Initial demographic variables, in conjunction with a broad range of clinical variables and laboratory parameters, were meticulously extracted from the institutional electronic health record system and the national digital health system. LVEF was quantified by transthoracic echocardiography using the modified Simpson's biplane method. The cohort was dichotomized according to mortality status into survivors and non-survivors. Demographic characteristics, clinical features, laboratory parameters, and NYHA functional classes were systematically compared between the two groups.

Ethical approval for the present investigation was formally conferred by the Ethics Committee of University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital (approval no: 2025-15-16, date: 20.08.2025) following a thorough review process. The investigation was conducted in strict accordance with the ethical principles of the Declaration of Helsinki, ensuring the protection of participant rights, welfare, and confidentiality throughout the study. All procedures adhered to established ethical standards governing biomedical research involving human subjects.

Definition

The CALLY index was defined as a composite biomarker integrating serum CRP, albumin concentration, and lymphocyte count, and is designed to provide a comprehensive assessment of both systemic inflammatory burden and nutritional status. The CALLY index was calculated as [serum albumin (g/L)×lymphocyte count (cells/μL)]/[CRP level (mg/L)×10^4] (12). HFrEF was defined as an LVEF ≤40%, accompanied by clinical manifestations and/or symptoms indicative of HF (2).

Statistical Analysis

Continuous data were expressed as means and their respective standard deviations, whereas categorical variables were presented as absolute counts and proportions. Comparative analyses between survivor and non-survivor cohorts were conducted using either the independent Student's t-test or the Mann-Whitney U test, depending on the underlying data distribution. Receiver operating characteristic (ROC) curve analyses were conducted to evaluate the predictive performance of the CALLY index for both mortality and NYHA class 3-4. An initial univariable logistic regression analysis was conducted to explore potential determinants of mortality. Predictor variables with a univariable p-value of <0.05 were subsequently incorporated into a multivariable logistic regression model to assess their independent contributions to the outcome while controlling for confounding influences that could bias the observed relationships. Collinearity was assessed using correlation coefficients; no substantial collinearity (defined as r>0.7) was detected. All inferential statistical procedures and data management processes were executed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA), in accordance with established methodological conventions and standards widely recognized in contemporary biomedical research.

RESULTS

A total of 146 individuals were enrolled in the current study, including 117 survivors and 29 who died, resulting in a mortality rate of 19.8%. While the gender distribution was similar between the two groups non-survivors were significantly older than survivors (66.55±9.35 vs. 61.82±11.02 years, p=0.035) and exhibited notably reduced LVEF (30.69±7.03% vs. 34.01±7.65%, p=0.033). Biomarkers indicative of systemic inflammation and myocardial injury—including CRP, troponin-T, and pro-brain natriuretic peptide—were substantially higher in non-survivors than in survivors. Other laboratory parameters did not differ significantly between groups. Additionally, the prevalence of advanced HF symptomatology, as defined by NYHA class 3-4, was markedly higher in the mortality group (31.0% vs. 10.3%, p=0.008). In contrast, the distributions of prevalent comorbid conditions [such as diabetes mellitus (DM), hypertension, and ischemic heart disease] and medication use were comparable between cohorts. Notably, the CALLY index—a composite measure integrating nutritional and inflammatory status—was significantly lower in non-survivors (1.33±2.86 vs. 4.02±9.11, p<0.001), highlighting its potential as a prognostic biomarker (Table 1).

Univariate logistic regression identified several key determinants associated with mortality. Notably, diminished LVEF was significantly associated with death, underscoring the critical role of cardiac function in patient outcomes [odds ratio (OR): 0.946; 95% confidence interval (CI): 0.896-0.998; p=0.044]. The presence of advanced HF symptoms, as indicated by NYHA class 3-4, was strongly predictive of mortality, emphasizing the clinical importance of symptom burden in this population (OR: 3.937; 95% CI: 1.466-10.573; p=0.007). The CALLY index was significantly inversely associated with mortality; higher values were associated

with improved survival (OR: 0.763; 95% CI: 0.626-0.981; p=0.034). Remaining variables failed to demonstrate statistically significant associations with mortality (Table 2).

Multivariate logistic regression analysis elucidated that several variables functioned as independent prognostic determinants of all-cause of mortality within the studied population. Notably, LVEF remained a significant protective factor: higher LVEF was associated with a reduced risk of mortality (OR: 0.958; 95% CI: 0.902-0.997; p=0.048), underscoring the vital role of preserved cardiac function in survival outcomes. In addition, HF symptoms classified as NYHA class 3-4 independently predicted increased mortality risk (OR: 2.845; 95% CI: 1.003-8.068; p=0.049), underscoring the significant influence of clinical symptom burden on patient prognosis. The CALLY index, reflecting the integrated status of nutrition and systemic inflammation, exhibited a significant inverse relationship with mortality (OR: 0.640; 95% CI: 0.438-0.935; p=0.021), highlighting its potential as a valuable prognostic biomarker (Table 3).

As presented in Table 4, patients classified in NYHA functional classes 3-4 (n=21) exhibited significantly lower CALLY index values compared with those in classes 1-2 (n=125) (1.22±1.53 vs. 3.80±8.91, p=0.016). This pronounced disparity highlights a robust correlation between diminished CALLY index values—indicative of compromised nutritional and inflammatory status—and more severe HF symptomatology, thereby underscoring the prognostic utility of the CALLY index in clinical evaluation.

ROC curve analysis was performed to rigorously determine the discriminative ability and prognostic efficacy of the CALLY index for both all-cause of mortality and advanced HF symptoms (defined as NYHA class 3-4). The CALLY

Table 1. Comparison of clinical, demographic, and laboratory parameters between survivors and non-survivors

Variable	Survival group (n=117) [mean±SD/n (%)]	Mortality group (n=29) [mean±SD/n (%)]	p-value
Age (years)	61.82±11.02	66.55±9.35	0.035
DM	42 (35.9)	9 (31)	0.620
HTN	73 (62.4)	15 (51.7)	0.297
IHD	65 (55.6)	17 (58.6)	0.765
ACEI/ARB	82 (70.1)	16 (55.2)	0.184
Beta-blocker	80 (68.4)	22 (75.9)	0.504
MRA	53 (45.3)	17 (58.6)	0.198
Furosemide	52 (44.4)	14 (48.3)	0.711
Antiplatelet	85 (72.6)	18 (62.1)	0.272
OAC	29 (24.8)	8 (27.6)	0.758
LVEF (%)	34.01±7.65	30.69±7.03	0.033
Glucose (mg/dL)	131.54±68.38	131.32±61.91	0.761
Creatinine (mg/dL)	1.19±0.89	1.27±0.90	0.165
Albumin (g/L)	45.12±34.70	40.75±6.63	0.136
CRP (mg/L)	1.32±2.78	1.96±2.41	0.006
Troponin-T (ng/L)	56.99±127.29	115.88±251.80	0.012
Pro-BNP (pg/mL)	2896.54±5366.33	5166.68±6430.90	0.011
LDL-C (mg/dL)	96.19±39.09	93.02±42.91	0.380
Triglyceride (mg/dL)	149.82±82.97	141.75±72.00	0.555
HGB (g/dL)	13.75±4.43	12.80±2.59	0.453
LEU (x10³/μL)	8.78±2.85	9.31±3.67	0.761
LYM (x10³/μL)	2.33±1.76	2.09±1.09	0.317
NYHA 3-4	12 (10.3)	9 (31.0)	0.008
CALLY index	4.02±9.11	1.33±2.86	<0.001

ACEI/ARB: Angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker, BNP: Brain natriuretic peptide, CALLY: C-reactive protein–albumin–lymphocyte, CRP: C-reactive protein, DM: Diabetes mellitus, HGB: Hemoglobin, HTN: Hypertension, IHD: Ischemic heart disease, LEU: Leukocytes, LDL-C: Low-density lipoprotein cholesterol, LVEF: Left ventricular ejection fraction, LYM: Lymphocytes, MRA: Mineralocorticoid receptor antagonist, NYHA: New York Heart Association, OAC: Oral anticoagulant, SD: Standard deviation

Table 2. Univariate analysis for factors associated with all-cause of mortality

	OR	95% CI	p-value
Age	1.032	0.993-1.074	0.111
LVEF	0.946	0.896-0.998	0.044
CRP	1.007	0.994-1.020	0.279
Pro-BNP	1.000	0.999-1.001	0.120
Troponin-T	1.002	1.000-1.004	0.111
NYHA 3-4	3.937	1.466-10.573	0.007
CALLY index	0.763	0.626-0.981	0.034

BNP: Brain natriuretic peptide, CALLY: C-reactive protein–albumin–lymphocyte, CI: Confidence interval, CRP: C-reactive protein, LVEF: Left ventricular ejection fraction, NYHA: New York Heart Association, OR: Odds ratio

Table 3. Multivariate analysis for determining independent predictors of all-cause of mortality

Variable	OR	95% CI	p-value
LVEF	0.958	0.902-0.997	0.048
CALLY index	0.640	0.438-0.935	0.021
NYHA 3-4	2.845	1.003-8.068	0.049

CALLY: C-reactive protein–albumin–lymphocyte, CI: Confidence interval, LVEF: Left ventricular ejection fraction, NYHA: New York Heart Association, OR: Odds ratio

Table 4. CALLY index between two groups of patients based on their NYHA classes

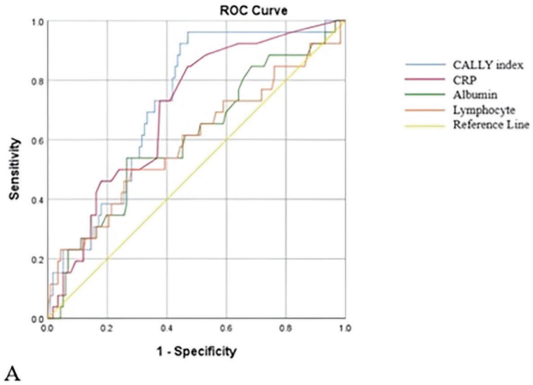
	NYHA class 1-2 (n=125)	NYHA class 3-4 (n=21)	p-value
CALLY index	3.80±8.91	1.22±1.53	0.016

CALLY: C-reactive protein–albumin–lymphocyte, NYHA: New York Heart Association

index exhibited a respectable predictive performance for all-cause of mortality, with an area under the curve (AUC) of 0.728 (95% CI: 0.633-0.823; p<0.001). A threshold value of 0.869 was established, providing a balanced sensitivity of 66.7% and a specificity of 71.4%. CRP demonstrated a moderate predictive ability (AUC=0.702, 95% CI: 0.600-0.803, p=0.001). In contrast, serum albumin (AUC=0.606, 95% CI: 0.483-0.729, p=0.090) and lymphocyte count (AUC=0.596, 95% CI: 0.464-0.727, p=0.127) exhibited lower discriminative capacity. Collectively, these findings suggest that the composite CALLY index provides a more robust and comprehensive prognostic indicator of mortality than any of its individual components. In the context of advanced HF symptoms (NYHA class 3-4), the CALLY index exhibited moderate discriminative ability with an AUC of 0.664 (95% CI: 0.544-0.784, p=0.016). A cut-off value of 0.681 was identified, corresponding to a sensitivity of 68.5% and a specificity of 57.1%. These findings highlight the potential utility of the CALLY index as a meaningful biomarker for risk stratification in patients with HF (Figure 1).

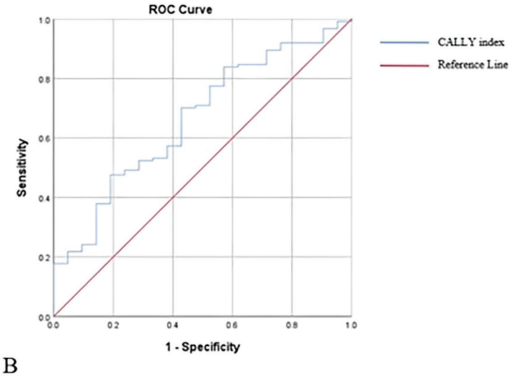
DISCUSSION

In our retrospective investigation aimed at elucidating the association among the CALLY index, NYHA functional classification, and all-cause of mortality in patients diagnosed with HFrEF, the following key findings emerged. Firstly, CALLY index was found to be associated with advanced functional impairment and independently predicted increased all-cause of mortality, demonstrating that lower CALLY values are associated with more severe clinical



A

	AUC	p-value	95% CI
CALLY index	0.728	<0.001	0.633-0.823
CRP	0.702	0.001	0.600-0.803
Albumin	0.606	0.090	0.483-0.729
Lympohocyte	0.596	0.127	0.464-0.727



B

	AUC	p-value	95% CI
CALLY index	0.664	0.016	0.544-0.784

Figure 1. ROC analysis of CALLY index and its individual components for all-cause of mortality and CALLY index for NYHA class 3-4. **A.** ROC analysis of CALLY index and its individual components for all-cause of mortality, **B.** ROC analysis for NYHA class 3-4

AUC: Area under curve, CALLY: C-reactive protein–albumin–lymphocyte, CI: Confidence interval, CRP: C-reactive protein, NYHA: New York Heart Association, ROC: Receiver operating characteristic

symptoms and a heightened risk of mortality in patients with HFrEF. In addition to the CALLY index, LVEF, with a mean value of approximately 30% in the non-survivors, was independently and inversely associated with all-cause of mortality. Ultimately, the investigation identified advanced functional deterioration (NYHA functional class 3-4) as the strongest independent prognostic determinant of all-cause of mortality within the cohort, with an odds ratio of 2.8.

Despite regional variations in the prevalence of HF attributable to differing socioeconomic conditions across countries, it remains a pervasive and significant global public health challenge. Affecting a large number of individuals worldwide, HF contributes substantially to higher mortality and morbidity rates. Furthermore, it imposes a profound economic burden on healthcare systems through heightened utilization of medical resources, frequent hospital admissions, and escalating costs (15). HF is traditionally stratified by LVEF into three principal groups: HFrEF, delineated by LVEF ≤40%; HFmrEF, characterized by LVEF 41-49%; and HFpEF, identified by LVEF ≥50% (2). Among these three classifications, HFrEF is the most frequently observed, representing approximately 60% of cases (16). Moreover, individuals diagnosed with HFrEF have a markedly higher incidence of cardiovascular mortality compared with those with HFmrEF or HFpEF, with documented rates ranging from 8.8% to 16.5%. Indeed, in the PARADIGM-HF trial, all-cause of mortality in the HFrEF subgroup reached as high as 19.8% (4,16). In comparison, the PARAGON-HF trial reported a lower all-cause of mortality rate, ranging from 14.2% to 14.6% (17). In our study, the all-cause of mortality rate was 19.8%, which was consistent with the mortality rates reported in the PARADIGM-HF trial.

Given the substantial impact of HFrEF on both mortality and morbidity, accurate prognostication in this patient population is of critical importance. Consequently, numerous clinical, biochemical, and functional parameters have been investigated to predict adverse cardiovascular outcomes and guide risk stratification. In a pooled analysis of the three distinct CHARM trials that collectively enrolled HF patients across the LVEF spectrum (>40% and ≤40%), age and insulin-dependent DM emerged as two of the most significant independent predictors of mortality (11). The absence of a quantitatively corroborated of age or DM with all-cause of mortality in our cohort may be ascribed to several methodological factors, including the limited sample size and the homogeneous nature of the study population, which consisted exclusively of patients with reduced ejection fraction (LVEF ≤40%). Additional independent clinical predictors of mortality in patients with HFrEF include body mass index, systolic blood pressure, heart rate, advanced symptom burden (NYHA class 3-4), chronic kidney disease, and peripheral arterial disease. Among the aforementioned parameters, advanced functional status, classified as NYHA class 3-4, emerges as the most robust predictor of mortality, conferring approximately a twofold increased risk of mortality (16). In our study, an advanced functional status, defined as NYHA class 3-4, was independently associated with all-cause of mortality (OR: 2.84).

Currently, the mortality benefit conferred by pharmacologic attenuation of the renin-angiotensin-aldosterone axis and antagonism of β-adrenergic receptors in individuals with HFrEF is unequivocally established, rendering these agents foundational to the modern therapeutic paradigm for HFrEF (2,18). Nonetheless, within our study, no statistically significant difference was observed between the survivor and non-survivor cohorts in the administration of these medications. The absence of a discernible difference may be attributable to the limited sample size, which could have reduced the statistical power required to detect subtle but clinically meaningful effects.

Inflammatory pathophysiology has garnered growing recognition as a pivotal determinant in the development and perpetuation of HF. Beyond its role in the underlying pathophysiology, systemic inflammatory activity exerts a profound influence on the clinical course of HF, modulating symptom burden, functional deterioration, and prognostic outcomes. Augmented circulating concentrations of CRP, a robust and extensively validated surrogate marker of systemic inflammatory activity, have been consistently associated with unfavorable cardiovascular outcomes, including heightened mortality risk, increased rates of readmission, and diminished functional capacity. In the Val-HeFT, CRP concentrations were found to be significantly associated with both advanced functional impairment—characterized by NYHA class 3-4—and increased mortality and morbidity among patients with HF (19-22). Beyond CRP, data from the EVEREST trial demonstrated that a relatively low lymphocyte count in patients with HFrEF was independently associated with an increased risk of all-cause of mortality, cardiovascular mortality, and HF-related hospitalization (23). In parallel with relative lymphocyte count, absolute lymphocyte count has demonstrated an inverse association with mortality (24).

Nutritional status also represents a fundamental clinical consideration in patients with HFrEF. Kałużna-Oleksy et al. (25) demonstrated an association between malnutrition and increased mortality in patients with HFrEF. Moreover, in the same study, patients who died had significantly lower serum albumin levels than those who survived. Albumin is

a negative acute-phase reactant, and it is well established that the acute-phase response is linked to deterioration in nutritional status. Hypoalbuminemia in HF can be linked to key underlying mechanisms such as inflammation and immune dysfunction, as well as malnutrition (26).

Considering the immunologic response, systemic inflammation, and nutritional status—each pivotal to the pathophysiology and prognosis of HF—our finding that the CALLY index, calculated from serum albumin concentration, lymphocyte count, and CRP levels, independently predicted all-cause of mortality as well as associated with advanced symptom burden (NYHA class 3-4) in patients with HFrEF appears to be consistent with the existing scientific literature. As albumin and lymphocyte levels decrease and CRP, a surrogate marker of systemic inflammatory response, increases, the CALLY index correspondingly declines. Our results likewise revealed a negative association between the CALLY index and both all-cause of mortality and advanced HF symptoms (NYHA class 3-4).

Furthermore, the complex interrelationships among systemic inflammation, immunologic condition, nutritional status, and myocardial impairment highlight the multifaceted etiology and prognostic landscape of HFrEF. The CALLY index serves as a holistic biomarker, integrating inflammatory, immunological, and nutritional dimensions into a singular measure that mirrors this intricate pathophysiological web. To address its novelty, the CALLY index was evaluated in the HFrEF population, among whom its association with mortality and functional class has not been comprehensively investigated. This study therefore provides novel evidence that the CALLY index may serve as a prognostic marker for clinical risk stratification in patients with HFrEF. The CALLY index complements the MAGGIC HF score by incorporating systemic inflammatory and nutritional parameters, thereby refining risk stratification beyond conventional clinical and demographic predictors. This integrated approach provides clinicians with a more nuanced tool to assess prognosis and guide individualized therapeutic strategies for patients with HFrEF (27).

Study Limitations

This study is subject to several notable limitations that should be acknowledged when interpreting the results. The retrospective and single-center design inherently limits external validity and may reduce the generalizability of the findings to broader or more diverse patient populations. The relatively small sample size likely diminished the statistical power to detect subtle but clinically meaningful associations, particularly with well-established prognostic factors such as age and DM. This limitation increases the potential for type II errors, possibly underestimating the influence of these variables on clinical outcomes. Additionally, the absence of key hemodynamic parameters—such as heart rate and systolic blood pressure—precluded a more comprehensive assessment of cardiovascular status and its relationship with mortality and morbidity. These omitted variables may have concealed important physiological mechanisms or confounding interactions relevant to outcome prediction in patients with HFrEF. Another limitation of our study is that data on sodium-glucose co-transporter-2 inhibitor use were lacking; therefore, the potential impact of these agents on clinical outcomes could not be evaluated. Collectively, these methodological constraints underscore the importance of conducting prospective, multicenter studies with larger, more heterogeneous cohorts and a more exhaustive set of clinical variables to enhance the robustness and applicability of the findings.

CONCLUSION

This study elucidated that the CALLY index serves as an independent predictor of all-cause of mortality and is associated with advanced functional impairment, as denoted by NYHA class 3-4, in individuals with HFrEF. Notably, diminished CALLY index values were significantly correlated with adverse clinical trajectories, underscoring its potential utility as a pragmatic, cost-effective, and readily accessible tool for risk stratification in everyday clinical settings. By consolidating multiple interdependent physiological dimensions into a singular, quantifiable metric, the CALLY index offers a more holistic appraisal of patient vulnerability and long-term prognosis. This integrative approach may be particularly valuable in contexts where conventional risk models fall short of capturing the multifaceted nature of disease burden in HFrEF. Nevertheless, to fully ascertain its prognostic fidelity and operational relevance, further corroboration through prospective, multicenter investigations involving larger and more heterogeneous cohorts is imperative.

ETHICS

Ethics Committee Approval: Ethical approval for the present investigation was formally conferred by the Ethics Committee of University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital (approval no: 2025-15-16, date: 20.08.2025) following a thorough review process.

Informed Consent: Retrospective study.

FOOTNOTES

Authorship Contributions

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

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

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

REFERENCES

1. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res. 2023;118:3272-87. Erratum in: Cardiovasc Res. 2023;119:1453.
2. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al.; ESC Scientific Document Group. 2023 Focused Update of the 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur J Heart Fail. 2024;26:5-17.
3. Murphy SP, Ibrahim NE, Januzzi JL Jr. Heart failure with reduced ejection fraction: a review. JAMA. 2020;324:488-504. Erratum in: JAMA. 2020;324:2107.
4. McMurray JJ, Packer M, Desai AS, Gong J, Lefkowitz MP, Rizkala AR, et al.; PARADIGM-HF Investigators and Committees. Angiotensin-neprilysin inhibition versus enalapril in heart failure. N Engl J Med. 2014;371:993-1004.
5. Gabet A, Juillière Y, Lamarche-Vadel A, Vernay M, Olié V. National trends in rate of patients hospitalized for heart failure and heart failure mortality in France, 2000-2012. Eur J Heart Fail. 2015;17:583-90.
6. Chen J, Normand SL, Wang Y, Krumholz HM. National and regional trends in heart failure hospitalization and mortality rates for medicare beneficiaries, 1998-2008. JAMA. 2011;306:1669-78.
7. The Criteria Committee of the New York Heart Association. Functional capacity and objective assessment. In: Dolgin M, editor. Nomenclature and criteria for diagnosis of diseases of the heart and great vessels. 9th ed. Boston, MA: Little, Brown and Company; 1994. p. 253-5.
8. Ahmed A. A propensity matched study of New York Heart Association class and natural history end points in heart failure. Am J Cardiol. 2007;99:549-53.
9. Briongos-Figuero S, Estévez A, Pérez ML, Martínez-Ferrer JB, García E, Viñolas X, et al. Prognostic role of NYHA class in heart failure patients undergoing primary prevention ICD therapy. ESC Heart Fail. 2020;7:279-83.
10. Holland R, Rechel B, Stepien K, Harvey I, Brooksby I. Patients' self-assessed functional status in heart failure by New York Heart Association class: a prognostic predictor of hospitalizations, quality of life and death. J Card Fail. 2010;16:150-6.
11. Pocock SJ, Wang D, Pfeffer MA, Yusuf S, McMurray JJ, Swedberg KB, et al. Predictors of mortality and morbidity in patients with chronic heart failure. Eur Heart J. 2006;27:65-75.
12. Iida H, Tani M, Komeda K, Nomi T, Matsushima H, Tanaka S, et al. Superiority of CRP-albumin-lymphocyte index (CALLY index) as a non-invasive prognostic biomarker after hepatectomy for hepatocellular carcinoma. HPB (Oxford). 2022;24:101-15.
13. He Q, Cao Y, Fan X, Li B, He Q, Zhang H. Long-term prognostic value of CRP-albumin-lymphocyte index in elderly patients with heart failure with preserved ejection fraction. Exp Gerontol. 2025;204:112744.
14. Ji H, Luo Z, Ye L, He Y, Hao M, Yang Y, et al. Prognostic significance of C-reactive protein-albumin-lymphocyte (CALLY) index after primary percutaneous coronary intervention in patients with ST-segment elevation myocardial infarction. Int Immunopharmacol. 2024;141:112860.
15. Hessel FP. Overview of the socio-economic consequences of heart failure. Cardiovasc Diagn Ther. 2021;11:254-62.
16. Chioncel O, Lainscak M, Seferovic PM, Anker SD, Crespo-Leiro MG, Harjola VP, et al. Epidemiology and one-year outcomes in patients with chronic heart failure and preserved, mid-range and reduced ejection fraction: an analysis of the ESC heart failure long-term registry. Eur J Heart Fail. 2017;19:1574-85.
17. Solomon SD, McMurray JJV, Anand IS, Ge J, Lam CSP, Maggioni AP, et al.; PARAGON-HF Investigators and Committees. Angiotensin-neprilysin inhibition in heart failure with preserved ejection fraction. N Engl J Med. 2019;381:1609-20.
18. Burnett H, Earley A, Voors AA, Senni M, McMurray JJ, Deschaseaux C, et al. Thirty years of evidence on the efficacy of drug treatments for chronic heart failure with reduced ejection fraction: a network meta-analysis. Circ Heart Fail. 2017;10:e003529.
19. Dick SA, Epelman S. Chronic heart failure and inflammation: what do we really know? Circ Res. 2016;119:159-76.
20. Van Linthout S, Tschöpe C. Inflammation-cause or consequence of heart failure or both? Current Heart Failure Reports. 2017;14:251-65.
21. Anand IS, Latini R, Florea VG, Kuskowski MA, Rector T, Masson S, et al.; Val-HeFT Investigators. C-reactive protein in heart failure: prognostic value and the effect of valsartan. Circulation. 2005;112:1428-34.
22. Alonso-Martínez JL, Llorente-Diez B, Echegaray-Agara M, Olaz-Preciado F, Urbieta-Echezarreta M, González-Arencibia C. C-reactive protein as a predictor of improvement and readmission in heart failure. Eur J Heart Fail. 2002;4:331-6.
23. Vaduganathan M, Ambrosy AP, Greene SJ, Mentz RJ, Subacius HP, Maggioni AP, et al.; EVEREST trial investigators. Predictive value of low relative lymphocyte count in patients hospitalized for heart failure with reduced ejection fraction: insights from the EVEREST trial. Circ Heart Fail. 2012;5:750-8.
24. Charach G, Grosskopf I, Roth A, Afek A, Wexler D, Sheps D, et al. Usefulness of total lymphocyte count as predictor of outcome in patients with chronic heart failure. Am J Cardiol. 2011;107:1353-6.
25. Kałużna-Oleksy M, Krysztofiak H, Migaj J, Wleklik M, Dudek M, Uchmanowicz I, et al. Relationship between nutritional status and clinical and biochemical parameters in hospitalized patients with heart failure with reduced ejection fraction, with 1-year follow-up. Nutrients. 2020;12:2330.
26. Gariballa S, Forster S. Effects of acute-phase response on nutritional status and clinical outcome of hospitalized patients. Nutrition. 2006;22:750-7.
27. Pocock SJ, Ariti CA, McMurray JJ, Maggioni A, Køber L, Squire IB, et al.; Meta-Analysis Global Group in Chronic Heart Failure. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. Eur Heart J. 2013;34:1404-13.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):444-448
DOI: 10.4274/BMJ.galenos.2025.2025.3-8

Research

The Relationship Between Hemoglobin to Albumin Ratio and Disease-Free Survival in Patients with Locally Advanced Nasopharyngeal Cancer Treated with Chemoradiotherapy

Kemoradyoterapi ile Tedavi Edilen Lokal İleri Nazofaringeal Kanserli Hastalarda Hemoglobin-Albümin Oranı ile Hastalıksız Sağkalım Arasındaki İlişki

Rumeysa Çolak, Caner Kapar, Ezgi Değerli, Seher Yıldız Tacar, Mesut Yılmaz

University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Medical Oncology, İstanbul, Türkiye

ABSTRACT

Objective: The hemoglobin-to-albumin ratio (HAR) has been linked to survival in various cancers; however, its significance in nasopharyngeal carcinoma (NPC) remains unclear. This study aims to explore the relationship between HAR and disease-free survival (DFS) in NPC patients undergoing concurrent chemoradiotherapy (CRT).

Methods: This retrospective study included 30 patients with NPC who received concurrent CRT from January 2018 to December 2024. HAR was calculated as the ratio of hemoglobin to albumin. We performed Cox regression analyses to identify variables associated with DFS.

Results: Patients were classified into low- and high-HAR groups using the cut-off value determined by maximally selected rank statistics. DFS was 51.7 months in patients with low HAR and 29.8 months in patients with high HAR. We found a significant relationship between HAR and DFS (p=0.009).

Conclusion: A high HAR was associated with a poor prognosis in NPC patients treated with CRT.

Keywords: HAR, DFS, nasopharyngeal cancer, chemoradiotherapy

ÖZ

Amaç: Hemoglobin-albümin oranı (HAR) çeşitli kanserlerde sağkalımla ilişkilendirilmiştir, ancak nazofarengeal karsinomdaki (NPC) önemi hala belirsizdir. Bu çalışma, eşzamanlı kemoradyoterapi (CRT) uygulanan NPC hastalarında HAR ile hastalıksız sağkalım (DFS) arasındaki ilişkiyi araştırmayı amaçlamaktadır.

Gereç ve Yöntem: Bu retrospektif çalışmaya Ocak 2018 ile Aralık 2024 arasında eşzamanlı CRT alan NPC'li 30 hasta dahil edildi. HAR, hemoglobinin albümine oranı olarak hesaplandı. DFS ile ilişkili değişkenleri tanımlamak için Cox regresyon analizleri yapıldı.

Bulgular: Hastalar, maksimum seçilen sıralama istatistikleriyle belirlenen kesme değerine göre düşük ve yüksek HAR gruplarına ayrıldı. Düşük HAR hastalarında DFS 51,7 ay, yüksek HAR hastalarında 29,8 ay idi. HAR ve DFS arasında anlamlı bir ilişki bulduk (p=0,009).

Sonuç: Yüksek HAR, CRT ile tedavi edilen NPC hastaları için kötü prognoz ile ilişkiliydi.

Anahtar Kelimeler: HAR, DFS, nazofarenks kanseri, kemoradyoterapi

Address for Correspondence: Rumeysa Çolak, MD, University of Health Sciences, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Medical Oncology, İstanbul, Türkiye
E-mail: colak.rmys@gmail.com **ORCID ID:** orcid.org/0000-0002-3867-0793

Cite as: Çolak R, Kapar C, Değerli E, Yıldız Tacar S, Yılmaz M. The relationship between hemoglobin to albumin ratio and disease-free survival in patients with locally advanced nasopharyngeal cancer treated with chemoradiotherapy. Med J Bakirkoy. 2025;21(4):444-448

Received: 07.04.2025
Accepted: 08.07.2025
Publication Date: 18.12.2025

INTRODUCTION

Nasopharyngeal carcinoma (NPC) is a head and neck epithelial malignancy that shows a distinct geographic distribution, with the highest incidence in East and Southeast Asia (1). Advancements in diagnostic and therapeutic approaches, along with the widespread use of concurrent chemoradiotherapy (CRT), have significantly reduced the mortality rate of NPC (2).

Research indicates that cancer prognosis is linked not only to tumor characteristics but also to the patient's nutritional status and systemic inflammation (3).

Recently, various blood-derived markers have been used to create prognostic models for different malignancies, including NPC (4). These biomarkers include hemoglobin, albumin, the hemoglobin-to-albumin ratio (HAR), the neutrophil-lymphocyte ratio, the platelet-lymphocyte ratio, the monocyte-lymphocyte ratio, and the systemic immune-inflammation index (5). Hemoglobin and albumin are the two most common indicators of a patient's nutritional status (6). Research indicates that low preoperative serum albumin and hemoglobin levels are strongly associated with poor outcomes in patients with malignant tumors (7). Additionally, lower levels of hemoglobin and albumin were significantly associated with poorer overall survival in a cohort study of 8,093 patients with NPC (8). Furthermore, the HAR in gastric cancer patients has shown predictive value for short-term prognosis (9). However, the role of HAR in NPC is still not well understood. In this study, we aimed to clarify the prognostic significance of HAR in patients with NPC undergoing concurrent CRT. This study aimed to investigate the importance of HAR in assessing survival in NPC patients undergoing CRT.

METHODS

This retrospective study included a consecutive cohort of NPC patients who received concurrent CRT between January 2018 and December 2024. All patients underwent induction chemotherapy, followed by concurrent CRT. Primary inclusion criteria employed in this study were as follows: (I) previously untreated, locally advanced NPC confirmed by histological and radiological assessments, without evidence of metastasis; (II) definitive CRT combined with tri-weekly platinum-based concurrent chemotherapy; (III) age ≥18 years; (IV) availability of clinical, histological, and follow-up information; (V) no previous antitumor therapy. The primary exclusion criteria in this study were as follows: (I) a history of secondary cancer; (II) chronic inflammatory diseases; (III) an Eastern Cooperative Oncology Group-performance status (ECOG-PS) score of ≤70.

Clinicopathological information was extracted from the patient's medical records. Based on previous studies, variables related to NPC patients' prognosis were incorporated into the current study, including age, gender, T stage, and N stage. HAR was calculated as the ratio of hemoglobin to albumin.

Tumor response was assessed clinically by gynaecological physical examination and by conventional imaging 6 weeks after completion of treatment.

The final follow-up was conducted in December 2024. All patients were reviewed every three months for the first two years after treatment and every six months thereafter.

Disease-free survival (DFS) refers to the time from the date of NPC diagnosis to the date of disease recurrence or metastasis, or to the date of the last follow-up.

This study was approved by the Ethics Committee of University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital (approval no: 2025-03-01, date: 05.02.2025), was conducted in accordance with the precepts established by the Declaration of Helsinki, and informed consent was obtained from all the participants.

Statistical Analysis

The descriptive statistics of the data include mean, standard deviation, median, minimum, maximum, frequency, and ratio values. The Kolmogorov-Smirnov and Shapiro-Wilk tests were used to assess the distribution of variables. The independent sample t-test was used to analyze quantitative independent data with a normal distribution. The Mann-Whitney U test was utilized to analyze independent quantitative data with a non-normal distribution. The chi-square test was employed to analyze independent qualitative data. Kaplan-Meier was used in the survival analysis. The SPSS 28.0 program was used in the analyses.

RESULTS

The patient characteristics are presented in the (Table 1). This retrospective study enrolled thirty patients with NPC. The mean age was 52.1±12.8 years. All patients had non-keratinizing squamous histology. Twenty-eight patients were male, and 23 had undifferentiated histology.

According to the tumor-node-metastasis (TNM) staging system, eight patients (26.7%) had stage II tumors, and 22 (73.3%) had stage III tumors. Twenty-six patients had an ECOG-PS of 0 and four had an ECOG-PS of 1.20% were positive for Epstein-Barr virus (EBV)-DNA and 6.7%

Table 1. Patient characteristics

Age (mean±standard deviation)	**52.1±12.8**
Gender (n, %)	
Male	28 (93.3)
Female	2 (6.7)
Histology (n, %)	
Differentiated	7 (23.3)
Undifferentiated	23 (76.7)
Stage (n, %)	
II	8 (26.7)
III	22 (73.3)
T stage (n, %)	
T2	12 (40.0)
T3	11 (36.7)
T4	7 (23.3)
N stage (n, %)	
N1	5 (16.7)
N2	24 (80.0)
N3	1 (3.3)
ECOG-PS (n, %)	
0	26 (86.7)
I	4 (13.3)
EBV-DNA (n, %)	
Positive	6 (20.0)
Negative	19 (63.3)
Unknown	5 (16.7)
HPV-DNA (n, %)	
Positive	2 (6.7)
Negative	9 (30.0)
Unknown	19 (63.3)
Type of induction agent (n, %)	
Cisplatin-gemcitabine	27 (90.0)
Carboplatin-gemcitabine	3 (10.0)
Treatment response (n, %)	
CR	18 (60.0)
PR	11 (36.7)
SD	1 (3.3)
Recurrence (n, %)	
Yes	21 (70.0)
No	9 (30.0)
Hemoglobin (mean)	14.3±1.9
Albumin (mean)	4.3±0.3
HAR (mean)	3.3±0.4

ECOG-PS: Eastern Cooperative Oncology Group-performance status, CR: Complete response, PR: Partial response, SD: Stable disease, HAR: Hemoglobin-to-albumin ratio, EBV: Epstein-Barr virus, HPV: Human papilloma virus

were positive for human papilloma virus (HPV)-DNA on pathological examination. All patients received induction chemotherapy (with cisplatin-gemcitabine or carboplatin-gemcitabine). Then, concurrent CRT, was administered with 29 patients showing an effective response complete response (CR)+partial response (PR) and one patient not responding as expected [stable disease (SD)], according to World Health Organization criteria. Recurrence (local, locoregional, or distant) occurred in nine patients (30.0%).

The mean hemoglobin and albumin values were 14.3±1.9 and 4.3±0.3, respectively. The mean HAR was 3.3±0.4. The median DFS was 43.1±7.4 months (Table 1).

In the univariate analysis, no significant differences were observed among age, gender, ECOG score, clinical stage, T stage, N stage, histological characteristics, and the presence of EBV or HPV with respect to DFS ($p>0.05$).

DFS was significantly lower ($p<0.05$) in the SD treatment response group (8.0 months) than in the CR (50.2 months) and PR (33.3 months) treatment response groups. DFS did not differ significantly between the treatment response groups with CR (50.2 months) and PR (33.3 months) ($p>0.05$) (Table 2).

No significant relationship existed between pretreatment hemoglobin and DFS, or between pretreatment albumin values and DFS ($p=0.228$ and $p=0.08$, respectively). DFS was 51.7 months in patients with low HAR and 29.8 months in patients with high HAR. We found a significant relationship between HAR and DFS ($p=0.009$; Figure 1).

Table 2. Univariate analysis for disease free survival

	Univariete model			
	HR	**95% CI**		**p-value**
Age	1.044	0.984	- 1.108	0.154
Gender	3.687	0.749	- 18.138	0.108
Histology	30.509	0.042	- >100	0.310
Stage	1.212	0.251	- 5.848	0.811
T stage	1.039	0.434	- 2.483	0.932
N stage	0.985	0.256	- 3.784	0.983
ECOG-PS	1.596	0.189	- 13.464	0.668
EBV-DNA	0.778	0.250	- 2.425	0.666
HPV-DNA	1.545	0.365	- 6.541	0.555
Type of induction agent	0.475	0.050	4.052	0.449
Treatment response	6.203	1.539	- 25.002	**0.010**
Hemoglobin	0.835	0.622	- 1.120	0.228
Albumin	8.301	0.777	- 88.740	0.080
HAR	0.319	0.102	- 1.000	**0.009**

ECOG-PS: Eastern Cooperative Oncology Group-performance status, HAR: Hemoglobin-to-albumin ratio, EBV: Epstein-Barr virus, HPV: Human papilloma virus, HR. Hazard ratio, CI: Confidence interval

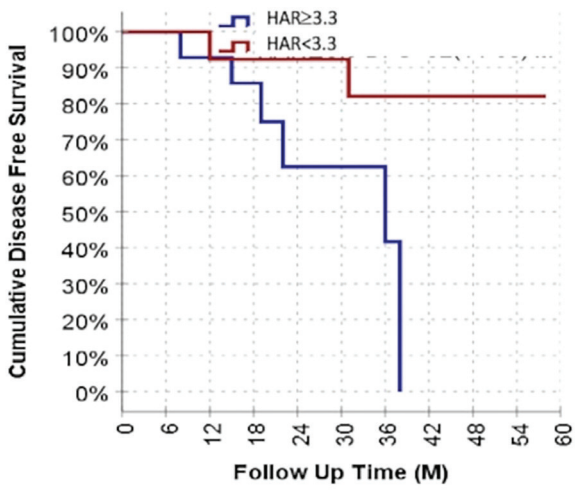


Figure 1. The relationship between HAR and DFS
HAR: Hemoglobin-to-albumin ratio, DFS: Disease-free survival

DISCUSSION

NPC is a relatively rare epithelial malignancy that exhibits significant regional variation. Historically, it has not received adequate attention or research (1). NPC exhibits considerable biological heterogeneity (10). According to the latest eighth edition of the TNM staging system, even among patients categorized in the same stage, survival outcomes can differ significantly when they undergo similar treatments (11). This indicates that the current anatomically based staging systems are insufficient for accurately predicting treatment effectiveness and patient prognosis. Our study showed no statistically significant differences in survival outcomes between stages.

Researchers are increasingly focusing on various clinicopathological factors and molecular biomarkers due to recent advancements in molecular biotechnology (12). EBV-DNA is a key marker with high sensitivity and specificity for the early diagnosis and screening of NPC (13). The predictive effectiveness of this method for staging and determining the prognosis of NPC is insufficient. In our trial, no statistically significant association was found between EBV-DNA and prognosis. These results may be due to the inability to evaluate the presence of EBV or HPV in all patients. Moreover, the fact that the study was conducted in a small, single center cohort may have influenced these results.

Genetic testing is costly, involves complex procedures, and has limited reproducibility. As a result, the widespread adoption of genetic testing for NPC still faces significant challenges. Ideal prognostic biomarkers are typically defined as characteristics that can independently identify an individual's prognosis. These biomarkers are designed to improve prognostic accuracy and provide multiple benefits. They are typically affordable, easily accessible and minimally invasive or non-invasive. Additionally, they provide clear indicators and can be scaled for implementation across various healthcare settings. Indicators identified through routine blood tests are associated with the most significant benefits (4).

The use of sub-parameters such as hemoglobin and albumin levels from routine tests—such as complete blood count and serum biochemistry—offers significant advantages over invasive histological examinations and costly genetic testing methods in clinical settings (4). Research suggests that hemoglobin levels affect oxygen transport, potentially leading to treatment-related complications and impacting the effectiveness of radiotherapy and chemotherapy (14).

Serum albumin is a crucial marker of nutritional status of cancer patients and an invaluable tool in their overall care and treatment (15). As cancer progresses, poor nutrition and inflammatory stress disrupt metabolism, leading to decreased albumin production and increased albumin consumption. This results in low serum albumin levels, a key predictor of prognosis in many types of cancer. Addressing these issues is vital for improving patient outcomes (4).

One trial revealed that TNM stage and HAR were significant independent prognostic factors for OS and PFS (16). We also found a significant relationship between HAR and DFS in our trial. High HAR was significantly associated with poor prognosis. However, there was no statistically significant association between the other clinical features and DFS.

Study Limitations

This study has some limitations. First, this was a single-center, small sample retrospective study. Second, several other inflammatory markers that correlate with prognosis were not included. Third, the hematological markers of NPC patients might fluctuate during treatment; for example, inflammation can reduce albumin levels, and chemotherapy can cause anemia, leading to variations in the HAR value at different time points. Therefore, multicenter, large-scale, prospective randomized controlled trials are necessary.

CONCLUSION

In this study, we showed that high HAR was associated with poor prognosis and was an independent predictor of outcome in NPC patients treated with CRT. Thus, the HAR score might be a convenient, non-invasive, affordable, and reliable tool to improve prognostication in NPC patients receiving CRT.

ETHICS

Ethics Committee Approval: This study was approved by the Ethics Committee of University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital (approval no: 2025-03-01, date: 05.02.2025), was conducted in accordance with the precepts established by the Declaration of Helsinki.

Informed Consent: Informed consent was obtained from all the participants.

FOOTNOTES

Authorship Contributions

Surgical and Medical Practices: R.Ç., E.D., Concept: M.Y., Design: R.Ç., Data Collection or Processing: R.Ç., Analysis or Interpretation: C.K., S.Y.T., Literature Search: R.Ç., S.Y.T., M.Y., Writing: C.K., E.D.

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

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

REFERENCES

1. Chen YP, Chan ATC, Le QT, Blanchard P, Sun Y, Ma J. Nasopharyngeal carcinoma. Lancet. 2019;394:64-80.
2. Wong KCW, Hui EP, Lo KW, Lam WKJ, Johnson D, Li L, et al. Nasopharyngeal carcinoma: an evolving paradigm. Nat Rev Clin Oncol. 2021;18:679-95.
3. Elinav E, Nowarski R, Thaiss CA, Hu B, Jin C, Flavell RA. Inflammation-induced cancer: crosstalk between tumours, immune cells and microorganisms. Nat Rev Cancer. 2013;13:759-71.
4. Zhang C, Wang SF, Zhang YL, Teng CX. Peripheral hemoglobin to albumin ratio predicts prognosis in patients with nasopharyngeal carcinoma underwent concurrent chemoradiotherapy. BMC Cancer. 2024;24:1012.
5. Kumarasamy C, Tiwary V, Sunil K, Suresh D, Shetty S, Muthukaliannan GK, et al. Prognostic utility of platelet–lymphocyte ratio, neutrophil–lymphocyte ratio and monocyte–lymphocyte ratio in head and neck cancers: a detailed PRISMA compliant systematic review and meta-analysis. Cancers (Basel). 2021;13:4166.
6. Oh SE, Choi MG, Seo JM, An JY, Lee JH, Sohn TS, et al. Prognostic significance of perioperative nutritional parameters in patients with gastric cancer. Clin Nutr. 2019;38:870-6.
7. Caro JJ, Salas M, Ward A, Goss G. Anemia as an independent prognostic factor for survival in patients with cancer: a systematic, quantitative review. Cancer. 2001;91:2214-21.
8. Zhang LL, Xu F, Song D, Huang MY, Huang YS, Deng QL, et al. Development of a nomogram model for treatment of nonmetastatic nasopharyngeal carcinoma. JAMA Netw Open. 2020;3:e2029882.
9. Hu CG, Hu BE, Zhu JF, Zhu ZM, Huang C. Prognostic significance of the preoperative hemoglobin to albumin ratio for the short-term survival of gastric cancer patients. World J Gastrointest Surg. 2022;14:580-93.
10. Chen YP, Yin JH, Li WF, Li HJ, Chen DP, Zhang CJ, et al. Single-cell transcriptomics reveals regulators underlying immune cell diversity and immune subtypes associated with prognosis in nasopharyngeal carcinoma. Cell Res. 2020;30:1024-42.
11. Chiang CL, Guo Q, Ng WT, Lin S, Ma TSW, Xu Z, et al. Prognostic factors for overall survival in nasopharyngeal cancer and implication for TNM staging by UICC: a systematic review of the literature. Front Oncol. 2021;11:703995.
12. Tan R, Phua SKA, Soong YL, Oon LLE, Chan KS, Lucky SS, et al. Clinical utility of Epstein-Barr virus DNA and other liquid biopsy markers in nasopharyngeal carcinoma. Cancer Commun. 2020;40:564-85.
13. Ramayanti O, Juwana H, Verkuijlen SA, Adham M, Pegtel MD, Greijer AE, et al. Epstein-Barr virus mRNA profiles and viral DNA methylation status in nasopharyngeal brushings from nasopharyngeal carcinoma patients reflect tumor origin. Int J cancer. 2017;140:149-62.
14. Gao J, Tao YL, Li G, Yi W, Xia YF. Involvement of difference in decrease of hemoglobin level in poor prognosis of Stage I and II nasopharyngeal carcinoma: implication in outcome of radiotherapy. Int J Radiat Oncol Biol Phys. 2012;82:1471-8.
15. Xu H, Zheng X, Ai J, Yang L. Hemoglobin, albumin, lymphocyte, and platelet (HALP) score and cancer prognosis: a systematic review and meta-analysis of 13,110 patients. Int Immunopharmacol. 2023;114:109496.
16. Chen Z, Ling J, Zhang S, Feng Y, Xie Y, Liu X, et al. Predicting the overall survival and progression-free survival of nasopharyngeal carcinoma patients based on hemoglobin, albumin, and globulin ratio and classical clinicopathological parameters. Head Neck. 2024;46:2600-15.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):449-457
DOI: 10.4274/BMJ.galenos.2025.2025.1-15

Research

Microsurgical Anatomy of the Insular Region and Its Connections: A Fiber Dissection Study

İnsular Bölgenin ve Bağlantılarının Mikrocerrahi Anatomisi: Bir Fiber Diseksiyon Çalışması

Ayşegül Esen Aydın, Merdan Balkansoy

University of Health Sciences Türkiye, Prof. Dr. Cemil Taşçıoğlu City Hospital, Clinic of Neurosurgery, İstanbul, Türkiye

ABSTRACT

Objective: The insular cortex is a structurally complex and functionally significant brain region, influencing a wide range of neurophysiological processes, including autonomic regulation and emotional cognition. This study aims to provide a detailed analysis of the microsurgical anatomy of the insula and its intricate white matter connections through meticulous fiber dissections of postmortem human brains.

Methods: Four postmortem human brains were preserved in a 10% formalin solution for at least two months. After removing the arachnoid mater, pia mater, and vascular structures, the brains were frozen at -16 °C for at least two weeks. Following rehydration in a 70% alcohol solution at room temperature, white matter dissections were performed using the Klingler technique.

Results: Dissection of the frontal, parietal, and temporal opercula revealed the insula, characterized by an inverted pyramidal shape at the base of the Sylvian fissure. The limen insula, an important anatomical landmark situated deep within the Sylvian fissure, was identified as a crucial reference point in surgical procedures. Upon removal of the insular cortex, detailed structural relationships between the insula, basal ganglia, and white matter tracts were demonstrated.

Conclusion: This study highlights the necessity of advanced surgical precision in addressing insular pathologies to improve neurosurgical outcomes.

Keywords: Insula, fiber dissection, neurosurgery, neuroanatomy

ÖZ

Amaç: Karmaşık bir yapı olan insüler korteks, otonom kontrolünden duygusal bilişe kadar geniş bir yelpazede nörofizyolojik işlevlerde kritik bir rol oynar. Bu çalışma, insülerin detaylı mikroskobik anatomisini ve karmaşık beyaz madde bağlantılarını, postmortem insan beyinlerinin titizlikle lif diseksiyonu yoluyla aydınlatmayı amaçlamaktadır.

Gereç ve Yöntem: Dört postmortem insan beyni, en az 2 ay boyunca %10 formalin çözeltisinde saklandı. Ardından araknoid mater, pia mater ve vasküler yapılar çıkarıldıktan sonra, beyinler en az 2 hafta boyunca -16 °C'de dondurulduu. Oda sıcaklığında %70 alkol çözeltisi içinde muhafaza edilen beyinler Klingler yöntemiyle insan kadavraları kullanılarak ak madde diseksiyonları yapıldı.

Bulgular: Frontal, parietal ve temporal operkulum, Silvian fissürünün tabanında yer alan ters piramit şeklindeki insüler yer almaktadır. Silvian fissürünün derininde yer alan ve insülerin anterobazal kısmını oluşturan limen insüler cerrahi açıdan önemli bir alandır. İnsüler korteks kaldırıldıktan sonra insülenin, bazal ganglionlar ve önemli ak madde yolları ile yakın ilişkisi ayrıntılı olarak gösterilmiştir.

Sonuç: Bu araştırma, insüler patolojilerin karmaşıklıklarını daha iyi ele almak ve nöroşirürji alanında hasta sonuçlarını iyileştirmek için gelişmiş cerrahi hassasiyetin kritik gerekliliğini vurgulamaktadır.

Anahtar Kelimeler: İnsula, lif diseksiyonu, nöroşirürji, nöroanatomi

Address for Correspondence: Ayşegül Esen Aydın, MD, University of Health Sciences Türkiye, Prof. Dr. Cemil Taşçıoğlu City Hospital, Clinic of Neurosurgery, İstanbul, Türkiye
E-mail: aysegul_esen@hotmail.com **ORCID ID:** orcid.org/0000-0001-7444-8156

Cite as: Esen Aydın A, Balkansoy M. Microsurgical anatomy of the insular region and its connections: a fiber dissection study. Med J Bakirkoy. 2025;21(4):449-457

Received: 17.04.2025
Accepted: 04.08.2025
Publication Date: 18.12.2025

INTRODUCTION

The insular cortex, a concealed structure deep within the lateral sulcus, is bounded by the frontal, parietal, and temporal opercula. Both experimental and clinical studies have underscored its multifaceted involvement in neurophysiological functions. Studies suggest that the insula plays a crucial role in autonomic regulation, emotional processing, and cognitive functions such as memory, mood, and sensory perception (1,2).

Predominantly affected by low-grade gliomas, the insular region is a common site of pathologies that present with epilepsy. The surgical management of insular gliomas poses significant challenges and a considerable risk of serious morbidity due to their propensity to infiltrate medially into critical paralimbic and limbic structures and laterally into key cortical areas such as the inferior frontal operculum and the supramarginal and angular gyri, particularly in the dominant hemisphere (3). This study meticulously details the microsurgical anatomy of the insula and elucidates its complex interconnections with the surrounding white matter fibers, aiming to enhance the precision and safety of neurosurgical interventions.

The insular cortex, a profound and enigmatic region of the brain, has significant implications for various neurophysiological and neuropsychiatric conditions. As a hub of convergence for sensory, emotional, and cognitive information, the insula orchestrates complex processes ranging from autonomic control to emotional regulation. Recognizing its centrality in human neurology, this study explores the intricate microsurgical anatomy of the insular region and its critical white matter pathways. Our exploration aims not only to delineate the anatomical landscape but also to underscore the clinical ramifications of its pathology, particularly in the context of low-grade gliomas and epilepsy. Such insights are pivotal as they enhance surgical precision and optimize patient outcomes, reflecting the necessity for a profound understanding of this region.

METHODS

Ethics Approval and Registration

Ethics committee approval was not required for this study, as it exclusively utilized cadaveric specimens obtained through authorized commercial sources. In accordance with relevant regulations, studies conducted on such materials do not necessitate ethical review. Since no living patients were involved, obtaining informed consent was not applicable.

Four postmortem human brains were kept in 10% formalin solution for at least (a total of eight hemispheres) months according to the method of Klingler. After the arachnoid mater, pia mater, and vascular structures were removed, they were frozen at -16 °C for at least two weeks. Then they were rinsed under running tap water and prepared for dissection (4). Between dissections, the brain hemispheres were kept in 70% alcohol solution at room temperature. Dissections were performed under a surgical microscope (Carl Zeiss AG, Oberkochen, Germany) at magnifications of 4× and 40×, using a Rhoton microsurgical set comprising toothless microforceps, a microhook, microscissors, a scalpel, and a dissector. Dissections began with decortication of the lateral and medial surfaces. Following the decortication, short association fibers (U-fibers) were removed and long major association fibers were accessed. Dissections were performed from the lateral to the medial aspect by removing fibers layer by layer. The interrelationships among the fibers and anatomical structures observed at each stage were demonstrated. All stages were recorded using a professional digital camera (Canon EOS 77D) equipped with a Canon 100 mm macro lens, a Canon ring flash, and 3D photography.

Statistical Analysis

This study is a qualitative and descriptive neuroanatomical investigation and does not include any quantitative or comparative dataset; therefore, statistical analysis is not required. As the aim of the study is to delineate microsurgical anatomical structures and demonstrate fiber pathways layer by layer, the use of statistical methods is not methodologically appropriate.

RESULTS

The arachnoid mater and pia mater, fixed using the Klingler method, were cleaned under a microscope using forceps (4). The lateral surfaces of the hemispheres were revealed, showing their cortical structures and all gyral and sulcal structures (Figure 1). Laterally, all cortical gray matter overlying the white matter was removed using dissectors. U-fibers connecting two adjacent gyri were exposed (Figure 2).

Language and Cognitive Connectivity

Long-range association fibers facilitating connectivity between distant cortical areas were visualized, including the superior longitudinal fasciculus (SLF) and the arcuate fasciculus (AF). The SLF was subdivided into SLF II and SLF III, with SLF II originating from the angular gyrus and terminating in the middle frontal gyrus, and SLF III extending from the supramarginal gyrus to the inferior frontal gyrus.

The fibers of the ventral component of the AF originate in the posterior portions of the superior, middle, and

transverse temporal gyri [Heschl's gyrus (HG)], make a turn at the level of the posterior insular point, run along the inferior aspect of the supramarginal gyrus just lateral to the corona radiata and along the superior insular limiting sulcus, and extend to the pars opercularis and triangular gyrus within the anterior limiting sulcus of the inferior frontal gyrus. The ventral component of AF was located inferolateral to SLF II and inferomedial to SLF III. SLF II and SLF III were removed, and the dorsal component of the AF was exposed. The dorsal component originated in the posterior one-third of the middle and inferior temporal gyri, turned at the angular gyrus, coursed inferior to SLF II, and terminated in the posterior and middle parts of the middle frontal gyrus, just anterior to the precentral sulcus, similar to SLF II (Figure 3).

The frontal, parietal, and temporal opercula were removed while protecting the HG, thereby exposing the inverted-pyramid-shaped insula located at the base of the Sylvian fissure. The sulci located around the insula, the circular sulcus and the anterior, superior, and inferior limiting sulci, were exposed. The anterior limiting sulcus separated the rostral insula from the inferior frontal gyrus, whereas the superior limiting sulcus separated the upper insula from the superior operculum. The inferior limiting sulcus separates the insula from the posterior part of the Sylvian fissure and the superior temporal gyrus. The limen insula, located deep within the Sylvian fissure and forming the anterobasal region of the insula, emerged as an important landmark.

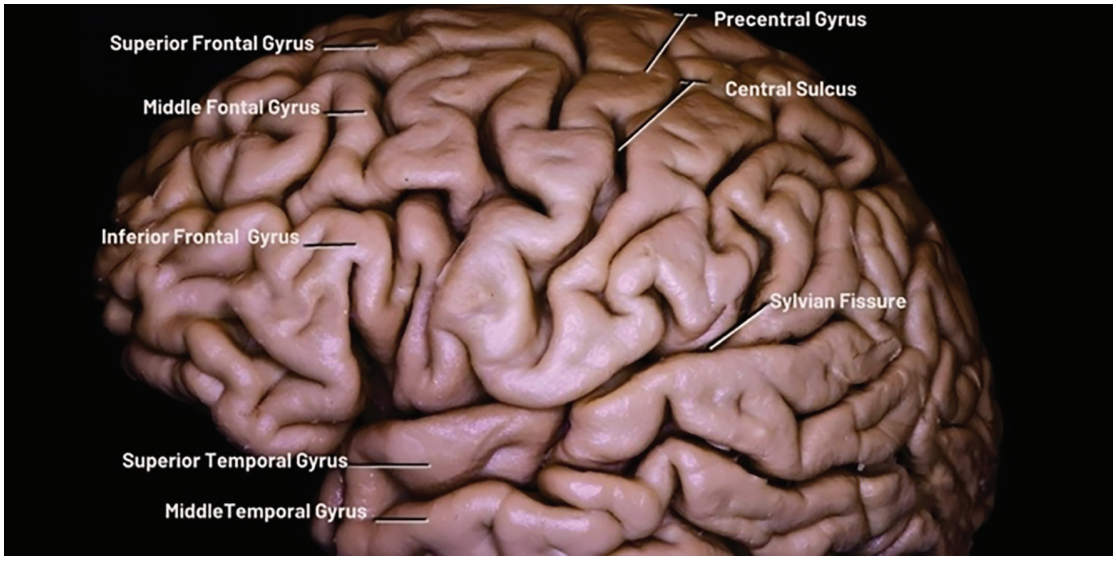


Figure 1. The lateral surface of the cerebral hemispheres

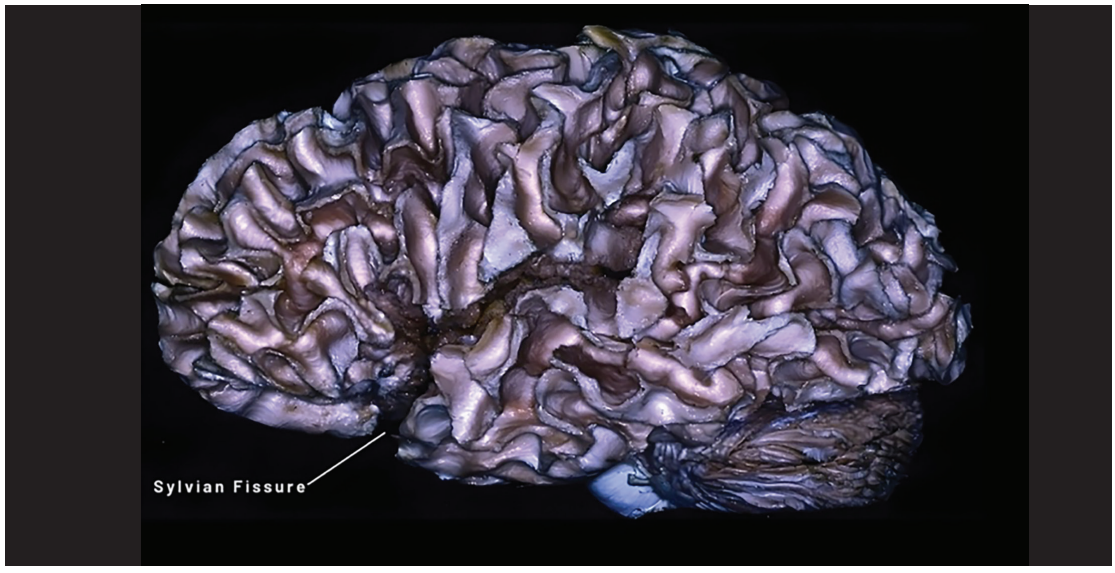


Figure 2. Lateral view of the cerebral hemisphere following the removal of cortical gray matter above the white matter using dissectors. U-fibers connecting adjacent gyri are distinctly visualized

The insula was divided into two parts by the central insular sulcus. The course of this sulcus was parallel to the central sulcus. The central insular sulcus divides the insula into anterior and posterior parts. Three gyri were present in the anterior insula and two in the posterior insula. The gyri anterior to the central insular sulcus are the anterior, middle, and posterior short gyri. The gyri posterior to the central insular sulcus are termed the anterior and posterior long gyri (Figure 4).

Upon removal of the insular cortex, the extreme capsule, which contains short association fibers linking the insula to adjacent opercular regions, was exposed. Further dissections revealed the middle longitudinal fasciculus, originating in the superior temporal gyrus and connecting to the sagittal stratum and the corona radiata. Following removal of the remaining portion of the extreme capsule, a gray layer—the claustrum—and the external capsule fibers associated with this gray matter were

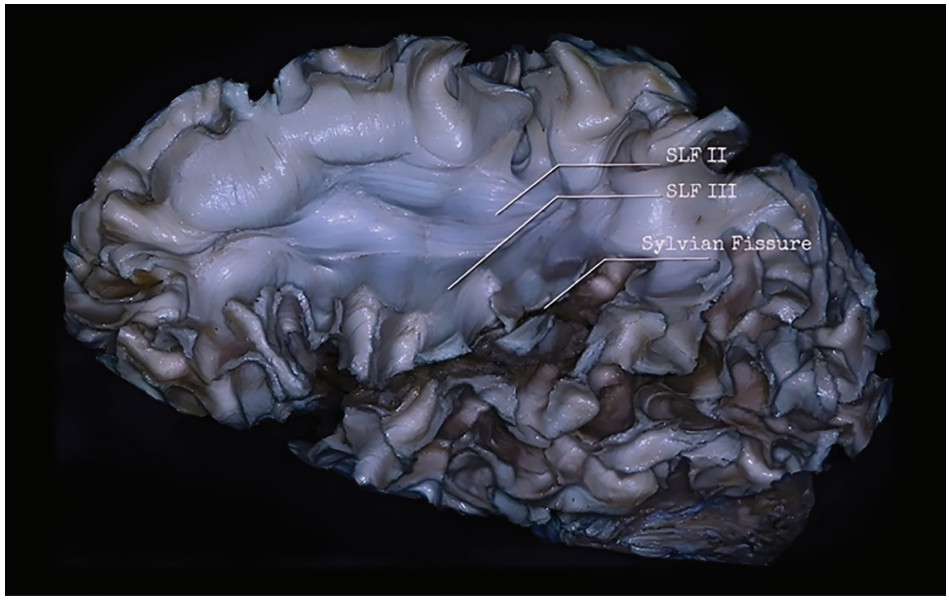


Figure 3. Dissection of the suprasylvian area showing SLF II and SLF III after U-fiber removal. SLF II connects the angular gyrus to the middle frontal gyrus, while SLF III extends between the supramarginal gyrus and the inferior frontal gyrus
SLF: Superior longitudinal fasciculus

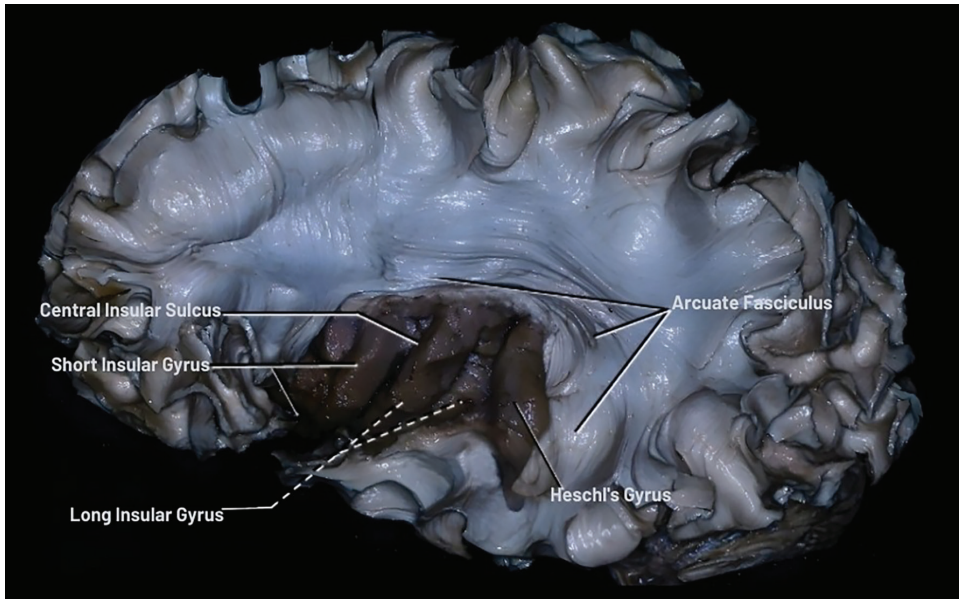


Figure 4. The insula exposed by removing the frontal, parietal, and temporal opercula, preserving Heschl's gyrus. The inverted pyramid-shaped insula, located at the base of the Sylvian fissure and surrounded by the circular sulcus, is visualized. The anterior, superior, and inferior limiting sulci demarcate the insula from neighboring structures. The central insular sulcus, which runs parallel to the central sulcus, divides the insula into anterior and posterior regions, which contain three short gyri and two long gyri, respectively

observed as dorsal and ventral components. Claustrocortical fibers, which constitute dorsal component of the external capsule, were traced to the corona radiata, where they merge with the internal capsule (IC) fibers; their fan-shaped structure extends from the claustrum to the supplementary motor area of the cortex and to the posterior parietal lobe. The ventral component of the external capsule was associated with two fiber bundles: the inferior fronto-occipital fasciculus (IFOF) and the uncinate fasciculus (UF). IFOF fibers were tracked anteriorly to the pars opercularis and pars triangularis within the inferior frontal gyrus. It was observed that the IFOF fibers crossed the AF posteriorly and joined the sagittal stratum. The fibers of the UF originating in the temporal region made a hook-shaped turn at the limen insula and reached the medial orbitofrontal, lateral orbitofrontal, and septal areas (Figure 5).

When the claustrocortical fibers were removed, the putamen, the most laterally located part of the lentiform nucleus (LN), was encountered. When the putamen was gradually removed, another component of the LN, the harder, lighter-colored, rounded globus pallidus (GP), was seen immediately medial to the putamen. The lateral part of the GP was completely covered by the putamen, and it was located lateral to the genu of the IC. The neighborhood of LN, including all IC segments, was shown. The posterior leg of the anterior commissure (AC), running inferior to the LN, was visualized. The posterior crural fibers of the AC were observed to fan out at the lateral border of the LN and to be directed toward their temporal and occipital terminations. After the SLF, AF, IFOF, and MLF fibers were removed, all parts of the IC were clearly revealed. The IC is a projection fiber bundle and continues its course alongside other fiber bundles that merge beyond the borders of the LN. After this combination, their names change during their journey. At the upper edge of the LN, IC fibers merge with claustrocortical fibers and form the corona radiata. Beyond the posterior border of the LN, it merges with the IFOF, AC, and optic radiation fibers associated with the ventral component of the external capsule and is termed the sagittal striatum. It is the continuation of the corona radiata. The frontal and parietal extensions are called the corona radiata, whereas the occipital and temporal extensions are called the sagittal striatum.

The IC fiber bundle is divided into three sections. These are the anterior limb, the posterior limb, and the genu. The anterior limb of the IC contains bundles of frontopontine and frontothalamic fibers. The IC genu contains the anterior parts of the corticobulbar and corticospinal fiber bundles, and the frontothalamic fiber bundle. The posterior limb of the IC is divided into three parts. These are the lenticulothalamic, retrolenticular, and sublenticular sections. The lenticulothalamic section contains the posterior portions of the corticospinal and corticobulbar fiber bundles, portions of the parietopontine and parietothalamic fiber bundles, and the frontothalamic fiber bundle. The retrolenticular section contains the parietopontine, occipitopontine, occipitothalamic, and parietothalamic fiber bundles.

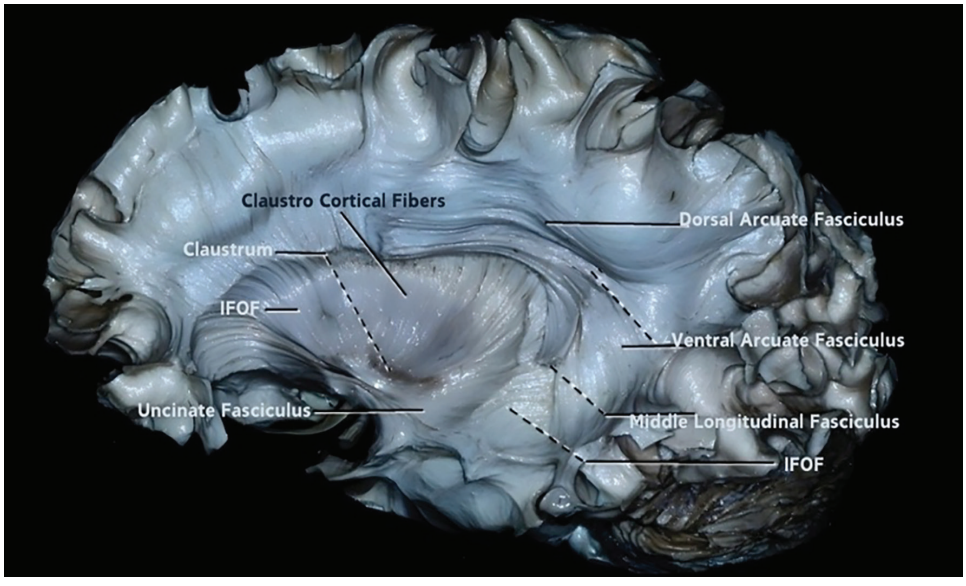


Figure 5. Exposure of the insular cortex revealing the extreme capsule, containing short association fibers connecting the insula with opercular regions. Removal of the extreme capsule exposed the middle longitudinal fasciculus, which originates from the superior temporal gyrus and joins the sagittal striatum and the corona radiata. Beneath the capsule, the claustrum and its associated external capsule fibers were visualized. The dorsal external capsule fibers, or claustrocortical fibers, fanned out toward the corona radiata. The ventral fibers were associated with the inferior fronto-occipital fasciculus (IFOF) and the uncinate fasciculus (UF), with the IFOF crossing the arcuate fasciculus and the UF forming a hook-like turn at the limen of the insula to reach the orbitofrontal and septal regions

The sublenticular section contains temporopontine and temporothalamic fiber bundles, as well as portions of the occipitopontine and occipitothalamic fiber bundles. The optic and auditory radiations are located in the retrolenticular and sublenticular sections of the IC.

AC fibers were observed ventral to the GP. The body of the AC is divided into two parts: anterior (ACa) and posterior (ACp). ACa curled along the nucleus accumbens at the level of the olfactory tract and extended toward the frontal gyrus. The substantia innominata was located between the anterior and posterior limbs of the AC. ACp extended in an anteromedial-to-posterolateral direction within the canal of Gratiolet, located at the base of the LN. The canal of Gratiolet is embedded in the gray matter of the putamen and the caudate nucleus, just anterior and inferior to the border of the GP. Fibers of the ACp extended toward the anterior portions of the temporal and occipital lobes. The number of fibers reaching the occipital lobe was greater than the number reaching the temporal lobe. The fibers extending to the temporal lobe were located inferiorly and merged with fibers of the UF.

Optic radiation fibers emerge from the lateral geniculate nucleus of the thalamus and terminate in the mesial occipital cortex. The dorsal portion of these fibers emerges from the lateral geniculate body and projects directly to the calcarine fissure. However, the ventral fibers initially course anteriorly and then return from the temporal horn to form Meyer's loop. It was observed that the optic radiation fibers, IFOF, and AC fibers destined for the occipital lobe coursed together and merged into the sagittal stratum.

The optic radiation fibers occupy the entire temporal horn, except for the roof and the anterior portion of the lateral wall. When these fibers are removed, the tail of the caudate nucleus, the amygdala, the hippocampus, and the choroid plexus lie together within the temporal horn of the lateral ventricle. When the choroid plexus was removed, the intraventricular hippocampus, located inferomedial to the atrium and temporal horn, was completely visible. The stria terminalis originating from the amygdala and the caudate nucleus tail connecting to the amygdala constituted the deepest layer of the "temporal stem" in the depth of the optic radiation (Figure 6, Table 1).

DISCUSSION

The insula, which is buried deep in the cerebral cortex, has attracted great attention in neuroscience research in recent years (5). The insula has been shown to play roles in memory, mood, autonomic and cardiac regulation, the perception of disgust, and the senses of smell and taste; it is also thought to have many other functions. Changes in the insula have been associated with neuropsychiatric diseases. The anterior insular region, due to its connections with the anterior cingulate gyrus, the dorsolateral prefrontal cortex, the striatum, and the amygdala, mediates autonomic and visceral integration underlying emotional, cognitive, and motivational functions. The posterior insula receives impulses

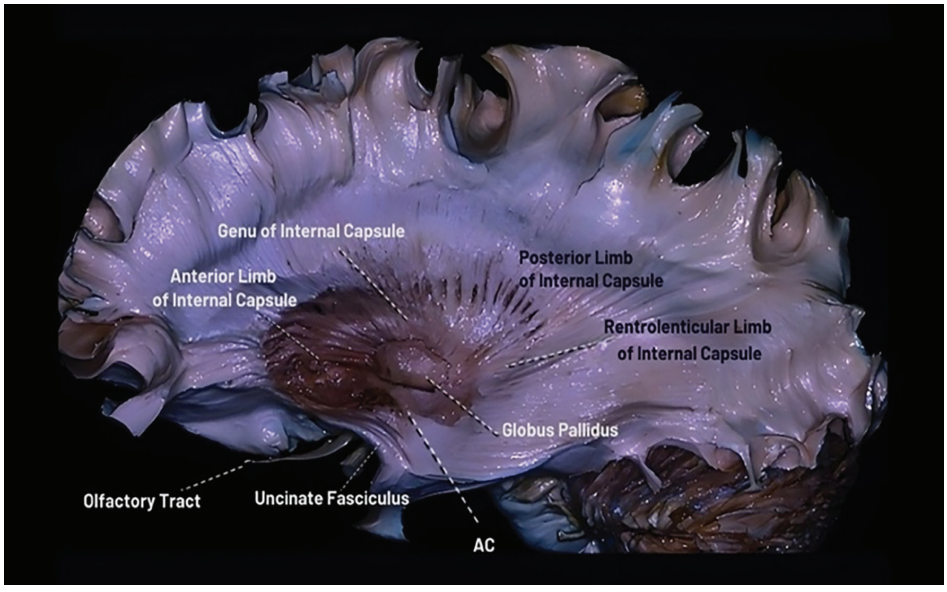


Figure 6. Dissection revealing the lentiform nucleus, internal capsule, anterior commissure (AC), and optic radiation. Removal of claustrocortical fibers exposed the putamen and globus pallidus, components of the lentiform nucleus, and revealed their relationship with segments of the internal capsule. The internal capsule divisions (anterior limb, genu, posterior limb) and associated fiber bundles, including corticospinal, corticobulbar, and thalamic fibers, were demonstrated. The AC was visualized ventral to the globus pallidus, with anterior and posterior components extending toward frontal, temporal, and occipital regions

Table 1. White matter fiber pathways related to the insular cortex: anatomical connections and functional roles

Fiber tract	Origin	Termination	Primary function
U-fibers	Adjacent gyri	Adjacent gyri	Local cortical integration
SLF* II	Angular gyrus	Middle frontal gyrus	Language processing and attentional networks
SLF* III	Supramarginal gyrus	Inferior frontal gyrus	Language production and motor planning
Arcuate fasciculus (ventral)	Posterior superior/middle/ transverse temporal gyri	Pars opercularis and triangularis of the inferior frontal gyrus	Semantic and phonological integration of language
Arcuate fasciculus (dorsal)	Posterior middle and inferior temporal gyri	Middle frontal gyrus (posterior and middle parts)	Repetition and echolalia-related connections
MdLF†	Superior temporal gyrus	Sagittal striatum, corona radiata	Auditory-processing and memory linkage
Extreme capsule fibers	Insula	Opercular areas	Cortex-to-cortex integration
Claustrocortical fibers	Claustrum	Supplementary motor area, posterior parietal cortex	Motor planning and attention networks
IFOF‡	Occipital lobe	Frontal lobe (pars opercularis)	Visual-cognitive and language-orientation pathways
UF	Temporal pole	Orbitofrontal cortex	Emotional memory and semantic association
Anterior commissure	Anterior temporal lobe and olfactory regions and occipital lobe	Contralateral anterior temporal lobe and olfactory regions and occipital lobe	Interhemispheric communication of temporal and olfactory information
Optic radiation	Lateral geniculate body	Calcarine fissure	Visual transmission

*: SLF: Superior longitudinal fasciculus, †: Middle longitudinal fasciculus, ‡: IFOF: Inferior frontal-occipital fasciculus, UF: Uncinate fasciculus

from the brainstem, medulla spinalis, parietal, temporal, and occipital cortices via the thalamus (6,7). In their morphometric study of schizophrenia, Wright et al. (8) found a decrease in gray matter volume in the bilateral temporal region, insula, as well as the left dorsolateral prefrontal cortex and amygdala. Similarly, many morphometric studies have found a relationship between insula volume and schizophrenia (9-11).

The hippocampus, which is closely related to the insula, is an important source of epileptic seizures. It is of great importance for epilepsy surgery, particularly in temporal lobe epilepsy due to hippocampal sclerosis. Microsurgery involving the insula requires a high degree of skill and precision because the insula can be affected by pathologies such as brain tumors, vascular malformations, and epilepsy.

The insula plays an important role in autonomic nervous system function and, owing to its complex structure, regulates emotional, cognitive, and homeostatic processes. Therefore, insular lesions may cause irregularities in cardiovascular function. Additionally, pain signals transmitted through the insula may modulate cardiovascular responses. Somatotopic transmission of painful stimuli via the insula helps us understand potential interactions within the cardiovascular system. Studies have emphasized the role of the insula in pain processing (12). Other studies have reported findings supporting cardiovascular effects associated with insular lesions. Understanding the complex interactions between the insula and the cardiovascular system may represent an important avenue for future research (13).

The insula is a complex structure in the brain that plays a critical role in speech processes and is associated with speech-related white matter fibers. The left insula, in particular, plays a significant role in the production and understanding of language, and in this context, it cooperates closely with Broca's area. The left insula influences speech ability by regulating processes, such as speech planning, motor control, and semantic processing of language (14). AF and SLF, which are white-matter fibers, play an important role in facilitating communication between the insula and other language regions. They connect Wernicke's area with Broca's area and facilitate information exchange between the area that processes language meaning and the area that provides motor control for language. This fiber bundle speech by integrating the cognitive and motor processes of language. The inferior longitudinal fasciculus is associated with language processing of information from the visual and auditory areas and facilitates communication between the insula and other language regions.

Recent advancements in neuroimaging and neuroanatomy have highlighted the insula's pivotal role in integrating somatosensory, gustatory, and visceral inputs, thereby influencing a wide array of bodily functions and emotional responses. The anterior part of the insula, heavily connected with frontal and limbic regions, is crucial for cognitive and emotional processing. In contrast, the posterior insula, linked to sensory pathways, plays a vital role in the perception and contextualization of pain and temperature. The insula's intricate connectivity and multifunctional nature make it a critical consideration in surgical planning, particularly in insular gliomas, where resection can significantly affect the patient's quality of life because of the risk of neuropsychological deficits. Moreover, understanding the insular involvement in autonomic functions underscores its role in disorders like heart disease and dysautonomia, reflecting the need for targeted therapeutic strategies. This discussion extends into neuropsychiatry, exploring correlations between insular abnormalities and mental health disorders such as schizophrenia and depression, highlighting the insula as a potential therapeutic target. Through this comprehensive analysis, we aim to bridge the gap between microsurgical anatomy and clinical practice, providing a foundation for future research and improved surgical interventions.

Study Limitations

Despite the comprehensive anatomical insights provided through meticulous fiber dissections, this study has certain limitations inherent to cadaver-based neuroanatomical research. The number of hemispheres examined was limited, which may reduce the breadth of anatomical variability captured across different specimens. Although the structural relationships and fiber pathways were consistently observed, future studies incorporating a larger number of hemispheres would enhance the generalizability and robustness of these findings. Nevertheless, the detailed layer-by-layer dissections presented here offer a meaningful contribution to the microsurgical understanding of the insular region.

CONCLUSION

The microsurgical anatomy of the insula represents a profoundly intricate and critical domain within neurosurgical practice, essential for the precise diagnosis and management of complex cerebral disorders. Mastery of this region's detailed anatomy, including its vascular and functional components, is paramount to enhancing the safety and efficacy of microsurgical interventions. This comprehensive understanding not only facilitates the minimization of intraoperative risks but also significantly improves postoperative recovery and functional outcomes for patients. The continuous evolution of microsurgical techniques, supported by ongoing research and technological innovations, holds the promise of advancing our capabilities in treating insular pathologies. Thus, as we refine these surgical approaches and deepen our anatomical knowledge, we are poised both to ameliorate the clinical management of insular disorders and to expand the frontiers of our understanding of brain function. This ongoing commitment to research and clinical excellence is crucial for fostering a future where neurosurgical procedures are not only life-preserving but also life-enhancing.

ETHICS

Ethics Committee Approval: Ethics committee approval is not required for studies conducted on commercially sold human cadavers, cadaveric parts, and other biological materials.

Informed Consent: Since no living patients were involved, obtaining informed consent was not applicable.

FOOTNOTES

Authorship Contributions

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

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

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

REFERENCES

1. Terasawa Y, Shibata M, Moriguchi Y, Umeda S. Anterior insular cortex mediates bodily sensibility and social anxiety. Soc Cogn Affect Neurosci. 2013;8:259-66.
2. Fathy YY, Hoogers SE, Berendse HW, van der Werf YD, Visser PJ, de Jong FJ, et al. Differential insular cortex sub-regional atrophy in neurodegenerative diseases: a systematic review and meta-analysis. Brain Imaging Behav. 2020;14:2799-816.
3. Ribas EC, Yagmurlu K, De Oliveira E, Ribas GC, Rhoton A. Microsurgical anatomy of the central core of the brain. J Neurosurg. 2018;129:752-69.
4. Wysiadecki G, Clarke E, Polguj M, Haładaj R, Żytkowski A, Topol M. Klingler's method of brain dissection: review of the technique including its usefulness in practical neuroanatomy teaching, neurosurgery and neuroimaging. Folia Morphol (Warsz). 2019;78:455-66.
5. Uddin LQ, Nomi JS, Hébert-Seropian B, Ghaziri J, Boucher O. Structure and function of the human insula. J Clin Neurophysiol. 2017;34:300-6.

6. Namkung H, Kim SH, Sawa A. The insula: an underestimated brain area in clinical neuroscience, psychiatry, and neurology. Trends Neurosci. 2017;40:200-7.
7. Eckart C, Kaufmann J, Kanowski M, Tempelmann C, Hinrichs H, Elbert T, et al. Magnetic resonance volumetry and spectroscopy of hippocampus and insula in relation to severe exposure of traumatic stress. Psychophysiology. 2012;49:261-70.
8. Wright IC, Ellison ZR, Sharma T, Friston KJ, Murray RM, Mcguire PK. Mapping of grey matter changes in schizophrenia. Schizophr Res. 1999;35:1-14.
9. Kasai K, Yamada H, Kamio S, Nakagome K, Iwanami A, Fukuda M, et al. Neuromagnetic correlates of impaired automatic categorical perception of speech sounds in schizophrenia. Schizophr Res. 2003;59:159-72.
10. Hulshoff Pol HE, Schnack HG, Mandl RC, van Haren NE, Koning H, Collins DL, et al. Focal gray matter density changes in schizophrenia. Arch Gen Psychiatry. 2001;58:1118-25.
11. Ha TH, Youn T, Ha KS, Rho KS, Lee JM, Kim IY, et al. Gray matter abnormalities in paranoid schizophrenia and their clinical correlations. Psychiatry Res. 2004;132:251-60.
12. Leone M, Proietti Cecchini A, Mea E, Tullo V, Curone M, Bussone G. Neuroimaging and pain: a window on the autonomic nervous system. Neurol Sci. 2006;27(Suppl 2):S134-7.
13. Sanchez-Larsen A, Principe A, Ley M, Vaquerizo B, Langohr K, Rocamora R. Insular role in blood pressure and systemic vascular resistance regulation. Neuromodulation. 2024;27:1218-26.
14. Ackermann H, Riecker A. The contribution of the insula to motor aspects of speech production: a review and a hypothesis. Brain Lang. 2004;89:320-8.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):458-467
DOI: 10.4274/BMJ.galenos.2025.2025.10-5

Research

Cross-Cultural Adaptation and Psychometric Validation of the Turkish Version of the American Orthopaedic Foot and Ankle Society Lesser Metatarsophalangeal-Interphalangeal Joint Scale

Amerikan Ortopedik Ayak ve Ayak Bileği Derneği Küçük Parmak Metatarsofalangeal-İnterfalangeal Skalasının Türkçe Uyarlaması ve Psikometrik Geçerlik-Güvenirlik Analizi

Ali Can Koluman[1], Ebru Aloğlu Çiftçi[2,3], Mehmet Utku Çiftçi[4], Yasemin Şahbaz[5], Vedat Öztürk[1], Altuğ Duramaz[1], Emre Baca[1], Nezih Ziroğlu[6]

[1]University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Orthopedics and Traumatology, İstanbul, Türkiye

[2]İstanbul Okan University Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, İstanbul, Türkiye

[3]İstinye University Institute of Graduate Education, Department of Physiotherapy and Rehabilitation, İstanbul, Türkiye

[4]University of Health Sciences Türkiye, Sultan 2. Abdülhamid Han Training and Research Hospital, Clinic of Orthopedics and Traumatology, İstanbul, Türkiye

[5]Beykent University Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, İstanbul, Türkiye

[6]Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, Department of Orthopedics and Traumatology, İstanbul, Türkiye

ABSTRACT

Objective: Lesser toe disorders can cause significant functional impairment and pain, requiring reliable tools for outcome assessment. The American Orthopaedic Foot and Ankle Society (AOFAS) lesser metatarsophalangeal-interphalangeal (MTP-IP) joint scale is a clinician-based instrument frequently used in foot and ankle evaluations, yet no validated Turkish version exists. The aim of this study was to translate, culturally adapt, and evaluate the psychometric properties of the Turkish version of the AOFAS lesser MTP-IP scale.

Methods: The scale was translated following international cross-cultural adaptation guidelines. A total of 43 patients with various lesser-toe pathologies were assessed using the AOFAS lesser MTP-IP, foot and ankle ability measure (FAAM), visual analogue scale, and short form-12 (SF-12). Test-retest reliability was assessed by calculating intraclass correlation coefficients [ICC (2,1)] using a two-way mixed-effects model with absolute agreement; by assessing internal consistency via Cronbach's alpha; and by evaluating agreement using Bland-Altman analysis. Construct validity was tested by correlating AOFAS scores with FAAM and SF-12 subscales. Floor and ceiling effects were also analyzed.

Results: The Turkish version demonstrated excellent test-retest reliability [ICC (2,1)=0.96] and acceptable internal consistency (α=0.76). Bland-Altman plots revealed no systematic bias. Strong correlations were observed with FAAM-activities of daily living (r=0.93) and FAAM-sports (r=0.75), whereas correlations with SF-12 physical component summary (r=0.34) and MCS (r=0.45) were weak but significant, which is consistent with the hypothesized convergent and divergent validity. A notable ceiling effect was identified in the AOFAS function and alignment domains, consistent with the high functional status and low pain levels reported by participants.

Conclusion: The Turkish adaptation of the AOFAS lesser MTP-IP scale is a reliable and valid instrument for evaluating pain, function, and alignment in patients with lesser toe disorders. Its strong psychometric performance supports its use in both clinical and research settings, although the observed ceiling effect should be interpreted in the context of patient characteristics.

Keywords: AOFAS, lesser toe, Turkish validation, reliability, validity, psychometrics, foot and ankle

Address for Correspondence: Ali Can Koluman, MD, University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Orthopedics and Traumatology, İstanbul, Türkiye
E-mail: md.alicankoluman@gmail.com **ORCID ID:** orcid.org/0000-0002-0191-3229

Cite as: Koluman AC, Aloğlu Çiftçi E, Çiftçi MU, Şahbaz Y, Öztürk V, Duramaz A, et al. Cross-cultural adaptation and psychometric validation of the Turkish version of the American Orthopaedic Foot And Ankle Society lesser metatarsophalangeal-interphalangeal joint scale. Med J Bakirkoy. 2025;21(4):458-467

Received: 22.10.2025
Accepted: 11.11.2025
Publication Date: 18.12.2025

ÖZ

Amaç: Küçük parmak patolojileri, belirgin ağrı ve fonksiyonel kısıtlılığa neden olabilir; bu durum, klinik değerlendirme ve tedavi sonuçlarının izlenmesinde güvenilir ölçüm araçlarının kullanılmasını gerektirir. Amerikan Ortopedik Ayak ve Ayak Bileği Derneği (AOFAS) tarafından geliştirilen lesser metatarsofalangeal-interfalangeal (MTP-IP) skalası, ayak ve ayak bileği değerlendirmelerinde sık kullanılan klinisyen temelli bir ölçektir; ancak bu ölçeğin geçerliliği ve güvenirliği kanıtlanmış Türkçe bir versiyonu bulunmamaktadır. Bu çalışmanın amacı, AOFAS lesser MTP-IP skalasının Türkçe uyarlamasını yapmak ve psikometrik özelliklerini değerlendirmektir.

Gereç ve Yöntem: Ölçeğin çeviri ve kültürel adaptasyonu uluslararası standartlara uygun olarak gerçekleştirildi. Farklı küçük parmak patolojileri bulunan toplam 43 hasta, AOFAS lesser MTP-IP, ayak ve ayak bileği yeteneği ölçeği (FAAM), görsel analog skala ve kısa form-12 (SF-12) kullanılarak değerlendirildi. Test-retest güvenirliği sınıf içi korelasyon katsayısı [ICC (2,1)] ile, iç tutarlılık Cronbach alfa katsayısı ile, ölçümler arası uyum ise Bland-Altman analiziyle değerlendirildi. Yapı geçerliği, AOFAS puanlarının FAAM ve SF-12 alt ölçekleriyle olan korelasyonları incelenerek test edildi. Ayrıca tavan ve taban etkileri analiz edildi.

Bulgular: Türkçe versiyon mükemmel test-retest güvenirliği [ICC (2,1)=0,96] ve kabul edilebilir iç tutarlılık (α=0,76) gösterdi. Bland-Altman analizinde sistematik yanlılık saptanmadı. AOFAS toplam puanları FAAM-günlük yaşam aktiviteleri (r=0,93) ve FAAM-spor (r=0,75) ile güçlü, SF-12 fiziksel (r=0,34) ve mental (r=0,45) bileşen puanları ile zayıf ancak anlamlı korelasyonlar gösterdi. Bu sonuçlar, önceden tanımlanan yakınsak ve ıraksak geçerlik hipotezleriyle uyumluydu. Katılımcıların yüksek fonksiyonel durum ve düşük ağrı düzeyleriyle uyumlu olarak, AOFAS fonksiyon ve dizilim alt ölçeklerinde belirgin bir tavan etkisi gözlendi.

Sonuç: AOFAS lesser MTP-IP skalasının Türkçe uyarlaması, küçük parmak patolojilerinde ağrı, fonksiyon ve dizilimi değerlendirmede güvenilir ve geçerli bir ölçektir. Ölçeğin güçlü psikometrik özellikleri, klinik ve araştırma ortamlarında kullanılabilirliğini desteklemektedir. Ancak gözlenen tavan etkisi, örneklemin yüksek fonksiyonel düzeyiyle ilişkilendirilerek yorumlanmalıdır.

Anahtar Kelimeler: AOFAS, küçük parmak, Türkçe geçerlik, güvenirlik, psikometri, ayak ve ayak bileği

INTRODUCTION

Lesser toe disorders, including Freiberg's disease, metatarsalgia, deformities, and sequelae of trauma, are frequently encountered in orthopaedic practice and lead to pain, altered gait mechanics, and functional limitations. Accurate functional assessment is critical not only for guiding clinical decision-making but also for evaluating treatment efficacy in both conservative and surgical contexts (1,2).

Among the tools used to evaluate foot and ankle function, the American Orthopaedic Foot and Ankle Society (AOFAS) clinical rating systems have historically been among the most widely adopted instruments. The AOFAS lesser metatarsophalangeal-interphalangeal (MTP-IP) joint scale, in particular, is designed to assess pain, functional limitations, and alignment in the lesser toes (3). Despite its long-standing utility and simplicity, the scale has been subject to methodological critique because of its categorical structure, ceiling effects, and lack of patient-reported input (4,5). In recognition of these limitations, the AOFAS has officially withdrawn its endorsement of these clinician-based tools, advocating instead for integrating patient-reported outcome measures (PROMs) into both research and practice (6).

Nevertheless, the AOFAS scales remain widely used because of their clinical familiarity and relevance, particularly in reporting postoperative outcomes (7). Recent methodological frameworks—such as COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN)—emphasize the need to evaluate reliability, measurement error, and construct validity through rigorous statistical procedures, including intraclass correlation coefficients (ICC), standard error of measurement (SEM), and minimal detectable change (MDC) (8,9).

In Türkiye, validated versions of the AOFAS hindfoot and hallux scales are available (10-13); however, there is no psychometrically validated Turkish adaptation of the AOFAS lesser MTP-IP scale. This gap limits the comparability of clinical data across international studies and restricts the use of standardized instruments in Turkish-speaking populations. Despite the increasing emphasis on methodologically sound, culturally adapted tools in musculoskeletal outcomes research, a critical unmet need remains.

Therefore, this study aimed to translate and culturally adapt the AOFAS lesser MTP-IP scale into Turkish and to evaluate its psychometric properties in patients with lesser toe disorders. We hypothesized that the Turkish version would demonstrate strong reliability, internal consistency, and construct validity, supporting its use in both clinical practice and multicenter research efforts.

METHODS

Study Design and Participants

This cross-sectional methodological study was conducted at a tertiary-level orthopedic outpatient clinic between January and June 2024. Hospital records were used to identify patients who had been treated for a disorder affecting the lesser toe. These patients were invited to attend follow-up visits. Patients who met the eligibility criteria had received treatment within the last two to four

years. A total of 43 patients diagnosed with pathologies of the lesser toes (e.g., Freiberg disease, deformities, or sequelae of trauma) were enrolled (Figure 1). Although the COSMIN guidelines recommend at least 50 participants for an adequate assessment of reliability and validity, previous methodological reports suggest that 5-10 participants per item may be acceptable for focused validation studies (14). Our participant-to-item ratio of 5.4, combined with an ICC power >0.99, indicates adequate precision despite the modest sample size. For validity studies, it is recommended to include at least five patients per item (14). The AOFAS lesser MTP-IP scale consists of eight items, and our study included 43 patients, corresponding to 5.38 patients per item, thus meeting the recommended sample size criterion. Inclusion criteria included age ≥18 years, literacy in Turkish, and clinical or radiographic evidence of a lesser-toe disorder. Patients with systemic neurological conditions, prior major foot surgery, or cognitive impairment were excluded.

This study was approved by the Clinical Research Ethics Committee of University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital (approval no: 2022-06-05, date: 21.03.2022). Written informed consent was obtained from all participants in accordance with the Declaration of Helsinki.

Translation and Cultural Adaptation

The AOFAS lesser MTP-IP scale was translated into Turkish following the five-step process outlined by Beaton et al. (15): (1) forward translation by two bilingual experts; (2) synthesis of translations; (3) back translation by two native English speakers; (4) expert committee review; and (5) pretesting with 20 patients to assess clarity and comprehension. Minor linguistic adjustments were made based on participant feedback to enhance cultural relevance. The translators' professional backgrounds (medical vs. non-medical) and linguistic expertise were reported to ensure conceptual equivalence and transparency.

Outcome Measurements

AOFAS Lesser MTP-IP Scale (Turkish Version)

The AOFAS lesser MTP-IP scale is a clinician-administered tool designed to evaluate outcomes in patients with disorders of the lesser toes (3). It comprises three domains: pain (40 points), function (45 points), and alignment (15 points), with a total possible score of 100, where higher scores indicate better clinical status. The functional domain includes items related to activity limitations, footwear requirements, and mobility, whereas alignment is rated by the clinician based on joint positioning.

In this study, the original English version of the AOFAS lesser MTP-IP scale was translated and culturally adapted into Turkish using a standardized methodology (Table 1) (15). This Turkish version was administered at two time points, two weeks apart, to assess reliability and construct validity.

Visual Analogue Scale (VAS)

Pain intensity was assessed using a 10-centimeter VAS, which is a validated and widely used unidimensional measure of pain (16). Patients were asked to mark their level of pain on a horizontal line ranging from 0 (no pain) to 10 (worst imaginable pain). VAS scores were recorded in three conditions: at rest, during activity, and at night. This enabled multidimensional pain profiling and facilitated comparison with the AOFAS pain domain. The activity condition refers to pain experienced during daily activities, such as walking, climbing stairs, or performing occupational tasks.

Foot and Ankle Ability Measure

The foot and ankle ability measure (FAAM) is a patient-reported outcome measure (PROMs) developed to assess

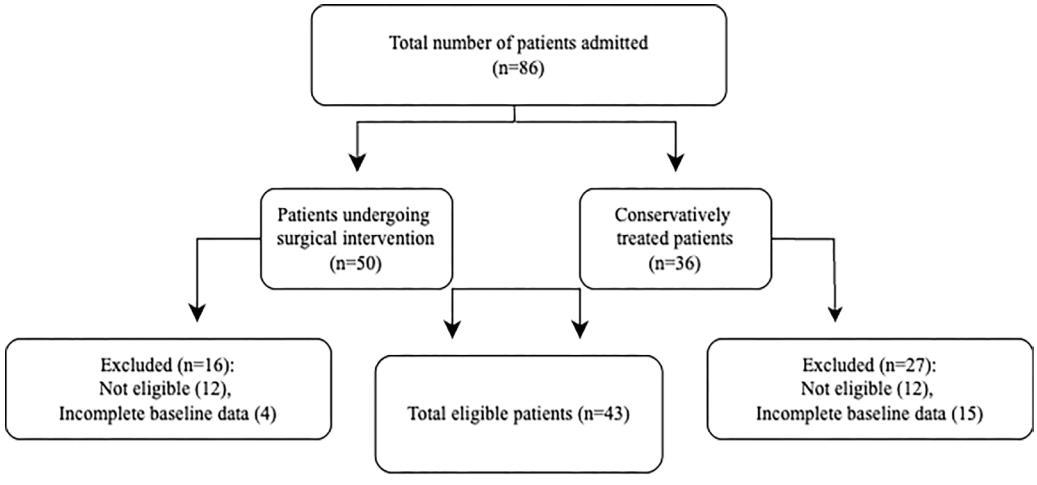


Figure 1. Flowchart. The number of patients in each particular group is shown in brackets

physical function in individuals with musculoskeletal disorderslower-extremity (17). It includes two subscales:

- Activities of daily living (ADL): 21 items,
- Sports: 8 items.

Each item is scored on a 5-point Likert scale (0=unable to do; 4=no difficulty), and raw scores are transformed into percentage scores (0-100), with higher scores reflecting greater functional ability. The Turkish version of the FAAM has been validated and shown to have high internal consistency and construct validity (18).

Short Form-12 (SF-12) Health Survey

SF-12 is a generic measure of health-related quality of lifederived from the original 36-item SF-36 questionnaire (19). It yields two-component summary scores:

- Physical component summary (PCS),
- Mental component summary (MCS).

These scores are standardized [mean=50, standard deviation (SD)=10], with higher scores indicating better self-perceived health. Although not specific to the foot and ankle, the SF-12 provides insight into general health status and permits evaluation of divergent validity when used alongside region-specific measures such as the AOFAS or FAAM (20).

All instruments were administered in Turkish; validated Turkish versions of the FAAM and SF-12 were used to ensure linguistic and conceptual equivalence. The SF-12 was used to test divergent validity, reflecting general health rather than region-specific function. We hypothesized strong correlations with FAAM subscales and weaker correlations with SF-12 PCS and MCS, consistent with COSMIN recommendations for hypothesis-driven construct validation. The Turkish version validated by Soylu and Kütük (21) was used.

Statistical Analysis

All statistical analyses were performed using IBM SPSS statistics, version 27.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were used to summarize demographic and clinical characteristics. Frequencies and percentages were reported for categorical variables; means and standard deviations were reported for continuous variables. The normality of continuous data was evaluated using the Shapiro-Wilk test.

An ICC value of 0.75 was used as the minimum acceptable threshold for reliability in accordance with established guidelines (22). The observed ICC of 0.96 yielded a post-hoc power of approximately 1.00 ($n=43$, $\alpha=0.05$), confirming sufficient statistical power to detect excellent test-retest reliability.

Internal consistency of the Turkish version of the AOFAS lesser MTP-IP scale was assessed using Cronbach's alpha. Values between 0.70 and 0.95 were considered acceptable (9,23).

Test-retest reliability was assessed by calculating ICC using a two-way mixed-effects model with absolute agreement [ICC (2,1)]. ICC values were interpreted as follows: <0.40=poor, 0.40-0.59=fair, 0.60-0.74=good, and ≥0.75=excellent.

Agreement was evaluated using the SEM and the MDC95. The ICC was applied to compute the SEM, which reflects the precision of the measurement. SEM was determined by multiplying the SD of the scores by the square root of (1-ICC). The MDC95 represents the smallest change that exceeds the threshold of measurement error. To estimate the MDC at the 95% confidence interval (CI) level (MDC 95%), the SEM was multiplied by 1.96 and then by the square root of 2 (22).

Construct validity was evaluated using Spearman's rank correlation coefficients between AOFAS total scores and scores on external instruments (FAAM, VAS, SF-12). Correlation strength was interpreted as small (0.10-0.29), moderate (0.30-0.49), strong (0.50-0.69), and very strong (≥0.70), according to Cohen's criteria.

Agreement between test and retest scores was further visualized using Bland-Altman plots, which graphically display the difference between paired measurements against their mean. Systematic bias and 95% limits of agreement (LoA) (LoA; mean±1.96×SD of differences) were examined to determine interchangeability between repeated assessments.

Floor and ceiling effects were evaluated by calculating the proportion of patients scoring the minimum (0-10) or maximum (90-100) possible scores at baseline. Effects were considered present if more than 15% of participants achieved either extreme (24). A p-value of <0.05 was considered statistically significant.

RESULTS

Participant Characteristics

A total of 43 patients (mean age: 38.60±18.70 years; 74.4% female) were included. The most common diagnoses were Freiberg's disease (65.1%), lesser toe fractures (27.9%), and deformities (7.0%). Surgical intervention was performed in 79.1% of cases, and 76.7% of patients reported normal functional status at baseline. Detailed demographic and clinical characteristics are presented in Table 2.

Reliability and Internal Consistency

The Turkish version of the AOFAS lesser MTP-IP scale demonstrated excellent test-retest reliability. The ICC (2,1) for the total score was 0.96 (95% CI: 0.95-0.98), with similarly high values observed for the subscales (pain: 0.96; function: 0.94; alignment: 0.89). Internal consistency, as measured by Cronbach's alpha, was acceptable (α=0.76 for the total score and the function subscale), indicating a coherent item structure (Table 3).

Table 1. AOFAS lesser MTP-IP joint scale-Turkish adaptation *(AOFAS küçük parmak MTP-IP eklem ölçeği Türkçe uyarlaması)*

Ağrı (40 puan)	**Puan**
Hiç yok	40
Hafif ve nadiren	30
Orta derecede ve her gün	20
Ciddi ve neredeyse her zaman	0
Fonksiyon (45 puan)	
Aktivite kısıtlılıkları, destek ihtiyacı	
Kısıtlılık yok	10
Günlük yaşam aktivitelerinde kısıtlılık yok (iş sorumlulukları gibi), eğlence aktivitelerinde kısıtlılık	7
Günlük yaşam aktivitelerinde ve eğlence aktivitelerinde kısıtlılık	4
Günlük yaşam aktivitelerinde ve eğlence aktivitelerinde ciddi kısıtlılık	0
Ayakkabı gereksinimleri	
Modaya uygun ve geleneksel ayakkabı, tabanlık gerekmez	10
Rahat ayakkabı, tabanlık gerekir	5
Modifiye edilmiş ayakkabı veya breys	0
MTP eklem hareketi (dorsifleksiyon ve plantarfleksiyon)	
Normal veya çok hafif kısıtlılık (≥75 °C)	10
Orta derecede kısıtlılık (30 °C-74 °C)	5
Ciddi kısıtlılık (<30 °C)	0
IF eklem hareketi (plantarfleksiyon)	
Kısıtlılık yok	5
Ciddi kısıtlılık (<10 °C)	0
MTP-IP stabilitesi (tüm yönlerde)	
Stabil	5
Tamamen instabil veya disloke	0
Nasır oluşumu (küçük parmak MTP-IP eklemde)	
Nasır yok veya semptom yok	5
Nasır var ve semptomatik	0
Dizilim (15 puan)	
İyi, küçük parmak dizilimi iyi	15
Orta, küçük parmakta bir dereceye kadar kötü dizilim, semptom yok	8
Kötü, belirgin derecede kötü dizilim, semptom var	0
Toplam puan	100

AOFAS: Amerikan Ortopedik Ayak ve Ayak Bileği Derneği, MTP-IP: Metatarsofalangeal-interfalangeal

Table 2. Demographic and clinical characteristics (n=43)

Characteristics	**Mean±SD**
Age, (years)	38.60±18.70
Hospital stay (days)	1.70±1.65
Number of comorbidities	0.21±0.63
Sex, n (%)	
Male	11 (25.6%)
Female	32 (74.4%)
Education, n (%)	
Secondary education	21 (48.8%)
Higher education	22 (51.2%)
Affected side, n (%)	
Right	22 (51.2%)
Left	21 (48.8%)
Surgery on lesser toe, n (%)	
Yes	34 (79.1%)
No	9 (20.9%)
Diagnosis, n (%)	
Freiberg's disease	28 (65.1%)
Lesser-toe fracture	12 (27.9%)
Lesser-toe deformity	3 (7.0%)
Smoking status, n (%)	
Non-smokers	27 (62.8%)
Smokers	16 (37.2)
Functional status	
Normal	33 (76.7%)
Nearly normal	6 (14.4%)
Abnormal	4 (9.3%)
Visual analogue scale	
Rest	1.00±1.27
Activity	1.14±1.40
Night	1.00±1.34
FAAM	
FAAM-ADL	93.05±13.48
FAAM-Sports	88.49±17.64
Short Form-12	
Physical component summary	53.01±8.72
Mental component summary	55.95±6.34

FAAM: Foot and Ankle Ability Measure, ADL: Activities of daily living scale, Sports: Sports scale, SD: Standard deviation

Measurement Error and Agreement

The SEM and MDC95 were calculated as 1.80 and 5.00 points, respectively, for the AOFAS total score. Bland-Altman analysis showed mean differences between test-retest scores (T1, T2) of -0.65 (LoA: -7.65 to 6.35) for AOFAS total; 0.42 (LoA: -1.54 to 2.38) for VAS total; -0.02 (LoA: -7.25 to 5.34) for SF-12 PCS; 0.54 (LoA: -4.82 to 5.90); 0.07 (LoA: -16.05 to 16.19) for FAAM-sports; and -0.36 (LoA: -7.47 to 6.75) for FAAM-ADL. There was no evidence of systematic bias, and the limits of agreement were narrow, supporting the scales' reproducibility (Figure 2). The vast majority of participants fell within ±1.96 SD of the mean.

Internal Consistency

Cronbach's alpha was 0.76 for both the total AOFAS scale and the function subscale, indicating acceptable internal consistency and item coherence within the domains.

Construct Validity

Strong negative correlations were found between the AOFAS total score and VAS scores for pain at rest (r=-0.81), during activity (r=-0.75), and at night (r=-0.82), all statistically significant (p<0.001). AOFAS scores were also strongly correlated with FAAM-ADL (r=0.93) and FAAM-sports (r=0.75), demonstrating convergent validity. Moderate correlations were observed with SF-12 PCS (r=0.34) and MCS (r=0.45), supporting divergent validity (Table 4).

Floor and Ceiling Effects

A ceiling effect was observed in the AOFAS total score, with 60.5% of participants achieving the maximum score. This was most prominent in the alignment and pain domains (90.7% and 76.7%, respectively). Similar ceiling effects were found in the FAAM-ADL (53.5%) and FAAM-sports (41.9%) subscales (Table 3). In contrast, VAS scores exhibited notable floor effects, with approximately half of participants reporting no pain. SF-12 scores did not display significant floor or ceiling effects, though both PCS and MCS subscales were skewed toward higher values, indicating good perceived health.

Importantly, the observed ceiling effect in AOFAS scores aligns with low pain levels and high functional status among participants, suggesting it reflects favorable clinical status rather than a measurement limitation.

Table 3. Test-retest reliability of the Turkish version of the AOFAS lesser MTP-IP scale

		First assessment	Second assessment	Test-retest reliability ICC (95% CI)	SEM	MDC	Floor-ceiling (%)
		Mean±SD (95% CI) (n=43)					
AOFAS lesser MTP-IP-T scale items	Pain	36.97±5.99 (35.13-38.82)	36.74±6.06 (34.88-38.61)	0.96 (0.94-0.98)	0.76	2.11	0-76.7
	Function	41.04±5.59 (39.32-42.76)	40.62±5.85 (38.82-42.42)	0.93 (0.87-0.96)	1.07	2.97	0-60.5
	Alignment	14.34±2.05 (13.71-14.98)	14.35±2.05 (13.72-14.98)	1	0.00	0.00	0-90.7
	Total score	92.21±12.58 (88.34-96.08)	91.56±12.91 (87.58-95.53)	0.96 (0.92-0.97)	1.80	5.00	0-60.5
VAS	Rest	1.00±1.27 (0.61-1.39)	1.58±1.73 (1.04-2.11)	0.60 (0.37-0.76)	0.74	2.06	48.8-0
	Activity	1.13±1.40 (0.71-1.57)	1.58±1.46 (1.12-2.03)	0.65 (0.43-0.79)	0.66	1.83	46.5-0
	Night	1.00±1.34 (0.59-1.41)	1.26±1.63 (0.75-1.76)	0.75 (0.59-0.85)	0.55	1.54	51.2-0
FAAM	ADL	93.05±13.48 (88.90-9.21)	92.69±12.99 (88.69-96.69)	0.96 (0.93-0.97)	1.82	5.06	0-53.5
	Sports	88.48±17.64 (83.06-93.92)	88.56±15.59 (83.76-93.36)	0.87 (0.78-0.93)	4.23	11.73	0-41.9
SF-12	PCS	53.01±8.72 (50.33-55.70)	52.99±7.95 (50.54-55.43)	0.90 (0.82-0.94)	1.88	5.22	0-0
	MCS	55.95±6.34 (54.00-57.90)	56.50±5.80 (54.71-58.28)	0.89 (0.82-0.94)	1.40	3.89	0-0

AOFAS lesser MTP-IP-T: Turkish version of American Orthopaedic Foot and Ankle Society lesser metatarsophalangeal-interphalangeal joint scale, VAS: Visual analogue scale, FAAM-ADL: Foot and Ankle Ability Measure-activities of daily living scale, SF-12 PCS: Short Form-12 physical component summary, SF-12 MCS: Short Form-12 mental component summary, ICC: Intraclass correlation coefficient, CI: Confidence intervals, SEM: Standard error measurement, MDC: Minimal detectable change, SD: Standard deviation

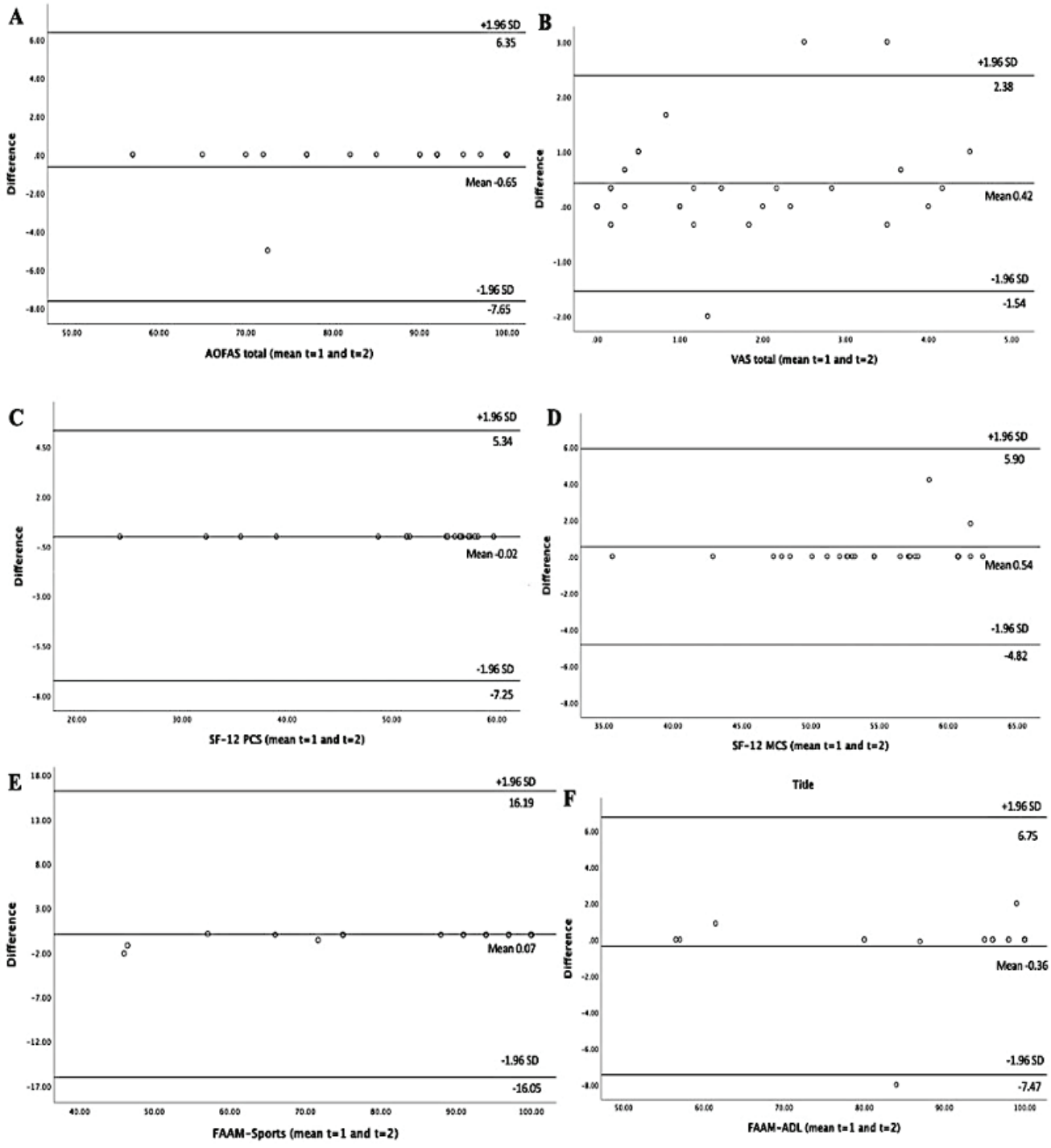


Figure 2. Bland-Altman plots for AOFAS lesser MTP-IP **(A)**, VAS **(B),** SF-12 PCS **(C),** SF-12 MCS **(D),** FAAM-Sports **(E)** and FAAM-ADL **(F)** scores in patients; change scores were calculated from time point 1 (t=1) to time point 2 (t=2). VAS (mean of rest, activity and night values). The solid central line represents the mean difference. The upper and lower solid lines represent the 95% limits of agreement, calculated as the mean difference ±1.96 standard deviation

AOFAS: American Orthopaedic Foot and Ankle Society, MTP-IP: Metatarsophalangeal-interphalangeal, VAS: Visual analogue scale, FAAM: Foot and Ankle Ability Measure, ADL: Activities of daily living scale, Sports: Sports scale, SF-12 PCS: Short Form-12 physical component summary, MCS: Mental component summary

Table 4. Construct and convergent/divergent validity of the Turkish version of the AOFAS lesser MTP-IP scale

	AOFAS lesser MTP-IP-T scale			
	Pain	**Function**	**Alignment**	**Total**
	r (p)			
VAS				
Rest	-0.70 (0.001)**	-0.81 (0.001)**	-0.35 (0.02)*	-0.81 (0.001)**
Activity	-0.66 (0.001)**	-0.73 (0.001)**	-0.33 (0.02)*	-0.75 (0.001)**
Night	-0.68 (0.001)**	-0.82 (0.001)**	-0.32 (0.03)*	-0.82 (0.001)**
FAAM				
ADL	0.76 (0.001)**	0.91 (0.001)**	0.51 (0.001)**	0.93 (0.001)**
Sports	0.65 (0.001)**	0.74 (0.001)**	0.49 (0.001)**	0.75 (0.001)**
Short Form-12				
Physical component summary	0.51 (0.001)**	0.33 (0.02)*	0.26 (0.08)	0.34 (0.02)*
Mental component summary	0.33 (0.02)*	0.42 (0.005)**	0.48 (0.001)**	0.45 (0.002)**

Spearman correlation test: *: p<0.05, **: p<0.01. AOFAS lesser MTP-IP-T scale: Turkish version of American Orthopaedic Foot and Ankle Society lesser metatarsophalangeal-interphalangeal joint scale, VAS: Visual analogue scale, FAAM: Foot and Ankle Ability Measure, ADL: Activities of daily living scale, Sports: Sports scale

DISCUSSION

This study provides compelling evidence that the Turkish adaptation of the AOFAS lesser MTP-IP scale is a reliable and valid instrument for evaluating pain, function, and alignment in patients with lesser toe pathologies. With excellent test-retest reliability [ICC (2,1)>0.9], acceptable internal consistency, and strong correlations with region-specific (FAAM) and general health (SF-12) instruments, the scale demonstrates robust construct validity and temporal stability key criteria for outcome measures used in clinical research and practice.

Despite the AOFAS scales no longer being officially endorsed because of concerns about responsiveness and lack of patient-centeredness (5,6), their use remains widespread in clinical orthopaedics. This paradox underscores the need for well-validated, culturally adapted versions that enable comparability across studies and acknowledge the inherent limitations of the tool. Although the AOFAS has withdrawn endorsement of its legacy clinician-based scores, these instruments remain in extensive clinical use. Their cultural validation enables comparisons across historical datasets and ongoing multicenter studies, ensuring continuity in outcomes research.

One of the novel contributions of this study is the inclusion of MDC95 and SEM, which provide estimates of measurement precision beyond random error; however, they do not indicate clinically meaningful change (minimal clinically important difference). In addition, presenting Bland-Altman analysis, SEM, MDC95, and floor-ceiling effect results for all administered instruments further strengthens the comprehensiveness of our evaluation. Applying Bland-Altman analyses further enhances methodological rigor by demonstrating a low systematic bias and good agreement between repeated assessments an aspect often overlooked in prior validation studies. Mean differences close to zero and the majority of data points falling within the ±1.96 SD limits supported reproducibility. Narrow limits of agreement indicated high reliability for the AOFAS, VAS, SF-12 PCS, SF-12 MCS, and FAAM-ADL scales, whereas wider limits for the FAAM-sports scale suggested greater interindividual variability.

A particularly noteworthy finding was the pronounced ceiling effect observed in the AOFAS pain and alignment domains. The substantial ceiling effect likely reflects the predominantly high-functioning postoperative status of the participants rather than a limitation of the scale itself. Nevertheless, this may restrict the scale's responsiveness in populations with more severe symptoms or functional impairment. Although traditionally viewed as a psychometric limitation, this effect was contextually consistent with the clinical status of the sample: nearly half of the participants reported no pain across VAS domains, and substantial proportions of participants reached maximum FAAM scores. These results parallel those reported by Whittaker et al. (25) and Ponkilainen et al. (26), who noted similar ceiling effects in AOFAS midfoot evaluations effects which could not be fully explained by other outcome metrics. In this light, the ceiling effect in our study appears to reflect favorable clinical outcomes rather than inadequate scale sensitivity.

Nevertheless, some limitations inherent to the AOFAS scoring structure must be acknowledged. Its categorical response format may reduce granularity and responsiveness, particularly in high-functioning or postoperative populations. Future iterations of the scale may benefit from more continuous scaling or the integration of patient input, aligning with broader trends in outcome measurement science.

In this context, the complementary use of PROMs is increasingly essential. Instruments such as FAAM and frameworks like PROMIS offer improved responsiveness, broader biopsychosocial coverage, and in the case of PROMIS, adaptive digital delivery. The integration of PROMIS with clinician-based scales has been shown to enhance the precision and patient-centeredness of musculoskeletal outcome measurement (27,28).

Furthermore, compared with other language versions of the AOFAS scale—including Persian, Italian, and Arabic adaptations—this study distinguishes itself by incorporating SEM, MDC95, and Bland-Altman analyses, which are rarely reported (29-31). These enhancements align with COSMIN standards for outcome measure validation (8,9) and address important methodological gaps in the existing literature.

Study Limitations

Strengths of the study include rigorous adherence to international guidelines for cross-cultural adaptation (15), the use of multiple comparator instruments, and comprehensive statistical analysis. However, the study is limited by its modest sample size, which precluded subgroup analysis by diagnosis or treatment type. Although COSMIN recommendations suggest a minimum sample size of ≥50 for reliability studies, our sample included 43 participants. This limitation should be acknowledged, and future studies with larger, more heterogeneous cohorts are warranted to confirm the stability and generalizability of these results. However, the ICC value of 0.96 corresponded to a post-hoc statistical power of >0.99, indicating that the precision of the reliability estimate was adequate despite the smaller sample size. Additionally, the scale's responsiveness to clinical change over time was not assessed and warrants further longitudinal research. The pronounced ceiling effects, while reflecting high functional recovery, may limit sensitivity in patients with more severe dysfunction and should be considered a psychometric constraint. Additionally, the patient cohort consisted predominantly of postoperative individuals with high functional recovery, which may limit the applicability of the findings to patients with more severe deformity or persistent pain.

CONCLUSION

The Turkish version of the AOFAS lesser MTP-IP scale is a psychometrically robust and clinically practical tool for evaluating lesser toe disorders. The observed ceiling effect, rather than undermining the scale's utility, reflects the functional recovery and low symptom burden of the study population. By providing SEM and MDC95 values and demonstrating strong construct validity, this study enhances the interpretability and applicability of the AOFAS scale in both clinical and research settings. Future studies should investigate the scale's responsiveness and explore its integration with PROMIS-based instruments to support a more holistic and precise approach to musculoskeletal outcome assessment. Future studies in larger, more heterogeneous populations are warranted to confirm the measurement properties and responsiveness of the Turkish AOFAS lesser MTP-IP scale.

ETHICS

Ethics Committee Approval: This study was approved by the Clinical Research Ethics Committee of University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital (approval no: 2022-06-05, date: 21.03.2022).

Informed Consent: Written informed consent was obtained from all participants in accordance with the Declaration of Helsinki.

FOOTNOTES

Authorship Contributions

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

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

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

REFERENCES

1. Coughlin MJ, Mann RA, Saltzman CL. Surgery of the foot and ankle. Elsevier Mosby; 2006.
2. Carmont MR, Rees RJ, Blundell CM. Current concepts review: Freiberg's disease. Foot Ankle Int. 2009;30:167-76.
3. Kitaoka HB, Alexander IJ, Adelaar RS, Nunley JA, Myerson MS, Sanders M. Clinical rating systems for the ankle-hindfoot, midfoot, hallux, and lesser toes. Foot Ankle Int. 1994;15:349-53.

4. Guyton GP. Theoretical limitations of the AOFAS scoring systems: an analysis using Monte Carlo modeling. Foot Ankle Int. 2001;22:779-87.
5. Pinsker E, Daniels TR. AOFAS position statement regarding the future of the AOFAS clinical rating systems. Foot Ankle Int. 2011;32:841-2.
6. Kitaoka HB, Meeker JE, Phisitkul P, Adams SB Jr, Kaplan JR, Wagner E. AOFAS position statement regarding patient-reported outcome measures. Foot Ankle Int. 2018;39:1389-93.
7. Hijji FY, Schneider AD, Pyper M, Laughlin RT. The popularity of outcome measures used in the foot and ankle literature. Foot Ankle Spec. 2020;13:58-68.
8. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19:539-49.
9. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34-42.
10. Analay Akbaba Y, Celik D, Ogut RT. Translation, cross-cultural adaptation, reliability, and validity of Turkish version of the American Orthopaedic Foot and Ankle Society ankle-hindfoot scale. J Foot Ankle Surg. 2016;55:1139-42.
11. Kalaycıoğlu O, Ayanoğlu T, Yılmaz ÖF, Turhan MT. A cross-cultural adaptation and validation of the Turkish version of American Orthopaedic Foot and Ankle Society metatarsophalangeal-interphalangeal scale (AOFAS-MTP-IP) for the hallux. Jt Dis Relat Surg. 2023;34:305-14.
12. Ziroglu N, Birinci T, Koluman A, Şahbaz Y, Çiftçi MU, Baca E, et al. Reliability and validity of the Turkish version of the American Orthopaedic Foot and Ankle Society hallux metatarsophalangeal-interphalangeal joint scale. Foot Ankle Spec. 2025;18:579-88.
13. Ziroğlu N, Şahbaz Y, Koluman A, Birinci T, Çiftçi MU, Baca E. Cross-cultural adaptation, reliability, and validity of the Turkish version of the American Orthopedic Foot and Ankle Society (AOFAS) midfoot scale. Med J Bakirkoy. 2024;20:264-70.
14. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval. 2005;10:7.
15. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976). 2000;25:3186-91.
16. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: visual analog scale for pain (VAS Pain), numeric rating scale for pain (NRS Pain), mcgill pain questionnaire (MPQ), short-form mcgill pain questionnaire (SF-MPQ), chronic pain grade scale (CPGS), short form-36 bodily pain scale (SF-36 BPS), and measure of intermittent and constant osteoarthritis pain (ICOAP). Arthritis Care Res (Hoboken). 2011;63(Suppl 11):S240-52.
17. Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the foot and ankle ability measure (FAAM). Foot Ankle Int. 2005;26:968-83.
18. Çelik D, Malkoç M, Martin R. Evidence for reliability, validity and responsiveness of Turkish foot and ankle ability measure (FAAM). Rheumatol Int. 2016;36:1469-76.
19. Ware J Jr, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996;34:220-33.
20. Ibrahim T, Beiri A, Azzabi M, Best AJ, Taylor GJ, Menon DK. Reliability and validity of the subjective component of the American Orthopaedic Foot and Ankle Society clinical rating scales. J Foot Ankle Surg. 2007;46:65-74.
21. Soylu C, Kütük B. Reliability and validity of the Turkish version of SF-12 health survey. Turk Psikiyatri Derg. 2022;33:108.
22. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15:155-63. Erratum in: J Chiropr Med. 2017;16:346.
23. Nunnally JC, Bernstein IH. Psychometric theory ed., New York; 1994.
24. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995;4:293-307.
25. Whittaker GA, Munteanu SE, Roddy E, Menz HB. Measures of foot pain, foot function, and general foot health. Arthritis Care Res (Hoboken). 2020;72(Suppl 10):294-320.
26. Ponkilainen VT, Uimonen M, Repo JP, Mattila VM, Haapasalo HH. Validity and internal consistency of the American Orthopaedic Foot and Ankle Society midfoot scale in patients with Lisfranc injury. Foot Ankle Surg. 2020;26:523-9.
27. Bernstein DN, Kelly M, Houck JR, Ketz JP, Flemister AS, DiGiovanni BF, et al. PROMIS pain interference is superior vs numeric pain rating scale for pain assessment in foot and ankle patients. Foot Ankle Int. 2019;40:139-44.
28. Hung M, Baumhauer JF, Licari FW, Bounsanga J, Voss MW, Saltzman CL. Responsiveness of the PROMIS and FAAM instruments in foot and ankle orthopedic population. Foot Ankle Int. 2019;40:56-64.
29. Alhadhoud M, Alsiri N, Alsaffar M, Glazebrook M. Cross-cultural adaptation and validation of an Arabic version of the American Orthopedics Foot and Ankle score (AOFAS). Foot Ankle Surg. 2020;26:876-82.
30. Vosoughi AR, Yoosefinejad AK, Dehbarez YS, Kargarshouraki Z, Mahdaviazad H. Evaluating the validity and reliability of the Persian version of American Orthopedic Foot and Ankle Society midfoot scale. Foot Ankle Spec. 2024;17:442-50.
31. Leigheb M, Vaiuso D, Rava E, Pogliacomi F, Samaila EM, Grassi FA, et al. Translation, cross-cultural adaptation, reliability, and validation of the Italian version of the American Orthopaedic Foot and Ankle Society-metatarsophalangeal-interphalangeal scale (AOFAS-MTP-IP) for the hallux. Acta Biomed. 2019;90(12-S):118-26.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):468-473
DOI: 10.4274/BMJ.galenos.2025.2023.11-3

Case Report

Magnetic Resonance Imaging of Benign Ovarian Masses

İyi Huylu Yumurtalık Kanserlerinin Manyetik Rezonans Görüntülemesi

Ahmet Kerem İmrek[1], Safiye Gürel[2], Bircan Alan[2]

[1]Niğde Training and Research Hospital, Clinic of Radiology, Niğde, Türkiye
[2]Bolu Abant İzzet Baysal University Faculty of Medicine, Department of Radiology, Bolu, Türkiye

ABSTRACT

Ovarian tumors have a wide spectrum and are pathologically classified in four main groups as epithelial, germ cell, sex cord-stromal, and metastatic tumors. Our aim is to demonstrate magnetic resonance imaging findings of benign ovarian masses with illustrative cases.

Epithelial Tumors

• Serous cystadenoma is the most common epithelial tumor. Case 1 is serous cystadenoma.

• Mucinous cystadenomas are often unilateral and tend to be larger cystic lesions. Case 2 is a mucinous cystadenoma.

Germ Cell Tumors

• Mature teratomas are unilocular masses filled with oily material and their walls contain hair follicles, skin glands, etc. Case 3 is mature teratoma.

Sex Cord-Stromal Tumors

• Fibrothecoma is the most common sex cord tumor. Case 4 is fibrothecoma.

• Cystadeofibroma is an uncommon benign tumor in which fibrous stroma is the dominant component. Case 5 is cystadenofibroma.

• Sclerosing stromal tumor is a rare benign tumor. Case 6 is a sclerosing stromal tumor.

Keywords: Adnexial benign lesions, magnetic resonance imaging, epithelial tumors

ÖZ

Yumurtalık tümörleri geniş bir spektruma sahiptir ve patolojik olarak epitelyal, germ hücreli, seks kord-stromal ve metastatik tümörler olmak üzere 4 ana grupta sınıflandırılır. Amacımız benign over kitlelerinin manyetik rezonans görüntüleme bulgularını örnek olgularla ortaya koymaktır.

Epitelyal Tümörler

• Seröz kistadenom en sık görülen epitelyal tümördür. Olgu 1 seröz kistadenomdur.

• Müsinöz kistadenomlar sıklıkla tek taraflıdır ve daha büyük kistik lezyonlar olma eğilimindedir. Olgu 2 müsinöz kistadenomdur.

Germ Hücreli Tümörler

• Matür teratomlar içi yağlı madde ile dolu, tek gözlü kitlelerdir ve duvarları kıl follikülleri, deri bezleri vb. içerir. Olgu 3 matür teratomdur.

Seks Kord-Stromal Tümörler

• Fibrotekom en sık görülen seks kord tümörüdür. Olgu 4 fibrotekomdur.

• Kistadeofibroma, fibröz stromanın baskın bileşen olduğu nadir görülen iyi huylu bir tümördür. Olgu 5 kistadenofibromdur.

• Sklerozan stromal tümör nadir görülen benign bir tümördür. Olgu 6 sklerozan stromal tümördür.

Anahtar Kelimeler: Yumurtalık iyi huylu tümörleri, manyetik rezonans görüntüleme, epitelyal tümörler

Address for Correspondence: Ahmet Kerem İmrek, MD, Niğde Training and Research Hospital, Clinic of Radiology, Niğde, Türkiye
E-mail: dr.ahmetimrek@gmail.com **ORCID ID:** orcid.org/0000-0003-2576-4683

Cite as: İmrek AK, Gürel S, Alan B. Magnetic resonance imaging of benign ovarian masses. Med J Bakirkoy. 2025;21(4):468-473

Received: 26.11.2023
Accepted: 30.06.2025
Publication Date: 18.12.2025

INTRODUCTION

Ovarian tumors have a wide spectrum and are pathologically classified into four main groups: epithelial, germ cell, sex cord-stromal tumors, and metastatic tumors (1).

The most common group is epithelial cancer. Epithelial benign cancers include serous cystadenoma, mucinous cystadenoma, cystadenofibroma, and Brenner tumors. Benign germ cell tumors include mature cystic teratomas. Benign sex cord-stromal tumors include thecoma, fibroma, fibrothecoma, and sclerosing stromal tumor (1).

The goal of our presentation is to demonstrate magnetic resonance imaging (MRI) findings of benign ovarian masses with illustrative cases.

CASE REPORT

It was stated that the names of the cases presented in case reports would not appear, and that permission was obtained for publication. Written and verbal consent was obtained from all patients/legal representatives.

Patients whose ultrasonographic (US) examinations revealed adnexal masses were referred to our department for lesion characterization with MRI.

Our MRI protocole consisted of big field of view (FOV) axial and coronal T1 weighted turbo spin-echo, small FOV and high resolution T2-weighted (T2W) turbo spin-echo sequences in 3 different planes, axial T2W echoplanar imaging, sagittal pre- and dynamic post-contrast gradient-echo T1-weighted (T1W) sequence and subtraction images (2,3).

Epithelial Tumors

• **Serous Cystadenoma:** It is the most common epithelial ovarian tumor, and 15-20% of serous cystadenomas are bilateral. In MRI, it has low-medium intensity on T1W and high-intensity on T2W images as a pure cyst and sometimes has small millimetric papillary projections. Often, thin <3 mm septa can be seen (4-6). Their average diameter is 10 cm, while 30 cm in diameter cystadenomas are reported in the literature. Case 1 was pathologically proven to be a serous cystadenoma (Figure 1).

• **Mucinous Cystadenoma:** They are often unilateral, multiloculated, and tend to be larger cystic lesions. They often contain only thin walls and septa. Depending on the density of mucinous content, lesions' T1W and T2W signals may differ. The presence of a thick septum (>5 mm) or a wall or a solid component with contrast enhancement should raise the suspicion of malignancy (7).

Case 2 was pathologically proven to be a mucinous cystadenoma (Figure 2).

Germ Cell Tumors

• **Mature Teratoma:** Mature cystic teratomas are unilocular masses filled with oily material and their walls contain hair follicles, skin glands, muscles, etc. Imaging findings, may range from a pure cyst to a mixed mass containing components of 3 germ sheets or a solid non-cystic mass containing mostly fat.

In MRI, hyperintensity T1W and suppression hyperintensity in fat-suppressed T1W sequence and variable fat-induced hyperintensity in the T2W sequence is typical (4).

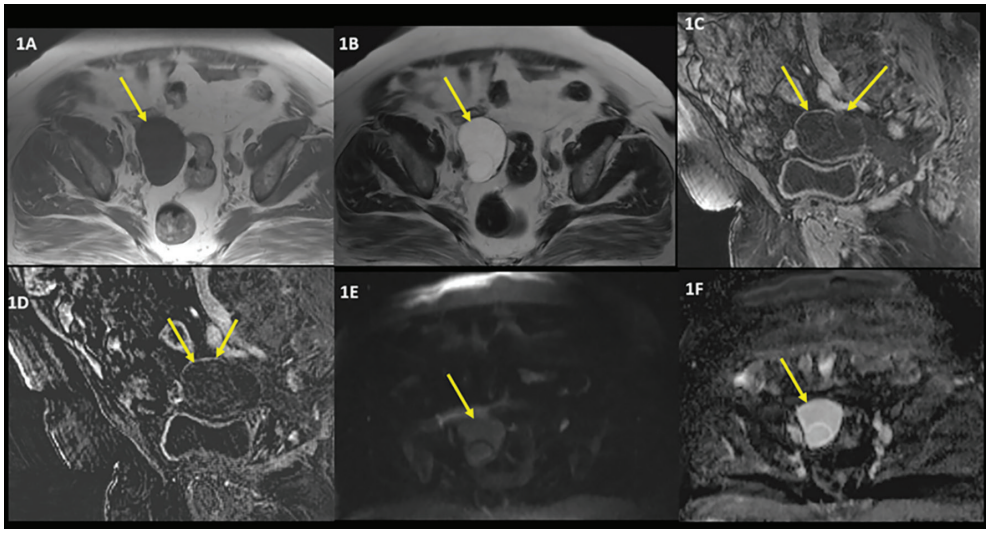


Figure 1. A 78-year-old woman with incidental ovarian mass. A 61x47x47 mm lesion was observed in the right ovary compared to the striated muscle; homogeneously hypointense in T1-weighted (Figure 1A arrow); homogeneously hyperintense in T2-weighted (Figure 1B arrow). In the postcontrast series, a triloculated cystic lesion with slight contrast enhancement in its wall and septa (Figure 1C arrow and subtraction Figure 1D arrow) and increased diffusion on diffusion weighted imagings (Figures 1E and 1F arrows) was observed, biopsy proven serous cystadenoma

Fat-suppressed T1W sequence is the key point in MRI diagnosis (5). Teeth and calcifications can be recognized as hypointense in all sequences. Ovarian teratomas can be associated with various complications including torsion (16% of ovarian teratomas), rupture (1-4%), malignant transformation (1-2%), infection (1%). Due to rupture, granulomatous peritonitis may occur.

Case 3 was pathologically proven to be a mature teratoma (Figure 3).

Sex Cord-Stromal Tumors

• **Fibrothecoma:** It is the most common sex cord tumor seen in the pre-post menopausal period. On MRI, the low signal in T1W and very low signal in T2W sequences are typical diagnostic features (5). Scattered high-signal areas within the mass may be observed due to edema or degeneration. Case 4 was pathologically proven to be a fibrothecoma (Figure 4).

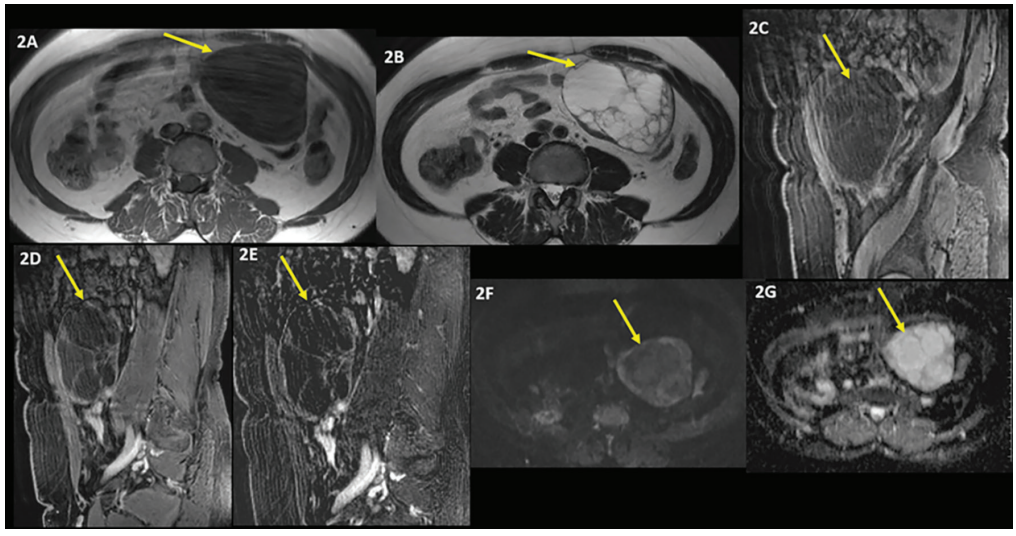


Figure 2. A 57-year-old woman with acne consulted from dermatology. A complicated cystic lesion measuring 102x67x117 mm in the left ovary was observed, exhibiting T1-weighted (T1W) hypointensity (Figure 2A arrow); T2-weighted hyperintensity (Figure 2B arrow); fatsat T1W (Figure 2C arrow) and with multiple septa, measuring 8 mm in the thickest part of the wall. In the postcontrast series, a complicated cystic lesion with minimal enhancement in its wall and septa (Figure 2D arrow and subtraction Figure 2E arrow) and increased diffusion in diffusion-weighted imagings (Figures 2F and 2G arrows) was observed, proven by biopsy mucinous cystadenoma

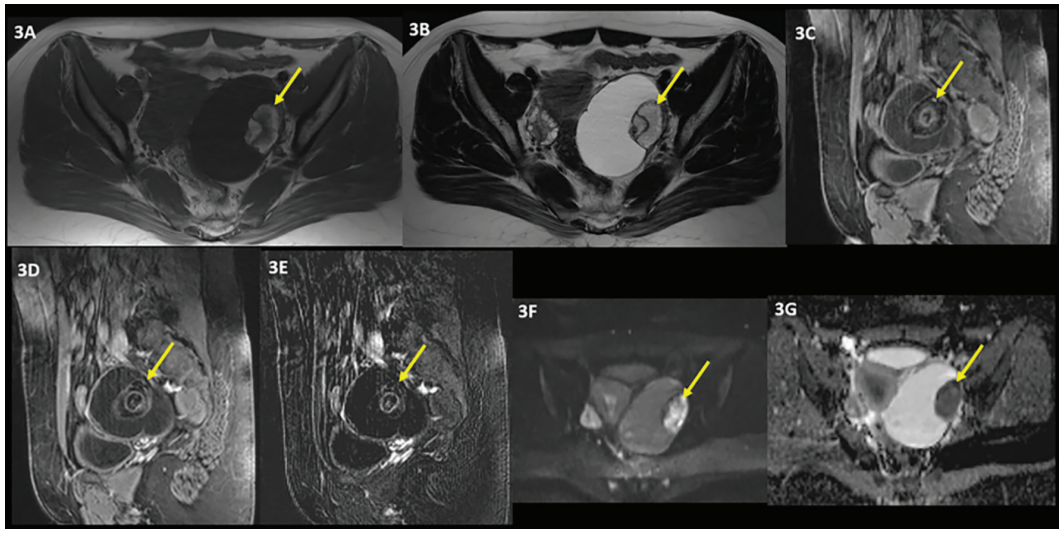


Figure 3. A 22-year-old woman with irregular menses. In the left ovary, an 88x62x71 mm lesion with 25x38 mm dimensions, containing fat, calcification, and hair-like signals (Figures 3A and B arrows), and suppressed fat in fat-suppressed sequence (Figure 3C arrow) was observed. In the postcontrast series, a lesion compatible with teratoma was observed that did not have contrast (Figure 3D arrow and subtraction Figure E arrow), and showed increased diffusion in the cystic component (Figures 3F and G arrows), proven by biopsy as a mature teratoma

• **Cystadeofibroma:** Cystadeofibroma is an uncommon benign tumor in which fibrous stroma is the dominant component of the lesion. They have either a pure cystic or complex cystic pattern with nodular or trabecular pattern (4).

On MRI, in T2W sequences fibrous stromal components such as septa have low signal intensity while cystic components have high signal and this composition creates the "black sponge" like appearance. Septa may show moderate contrast enhancement.

Case 5 was pathologically proven to be a cystadenofibroma (Figure 5).

• **Sclerosing Stromal Tumor:** Sclerosing stromal tumor is a rare benign ovarian tumor that occurs in young women in the second and third decades. On MRI, on T2W sequences, it is a large mass with a solid heterogeneous component containing hyperintense cystic areas and a medium-high signal (8). Contrast-enhanced MRI shows early peripheral and progressive centripetal enhancement

Case 6 was pathologically proven to be a sclerosing stromal tumor (Figure 6).

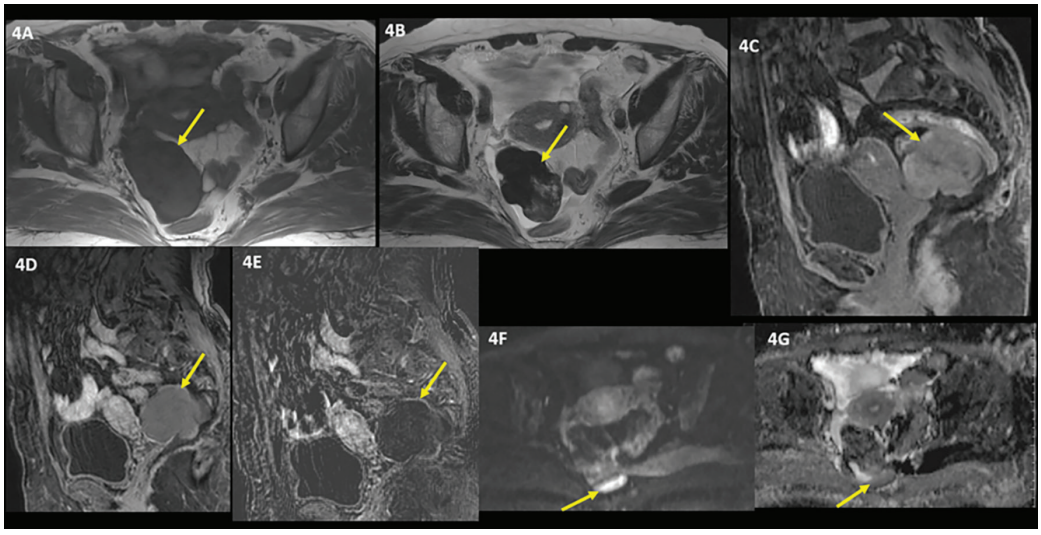


Figure 4. A 77-year-old woman with postmenopausal complaints. In the right ovary, a solid lesion measuring 70x47x60 mm, heterogeneous, iso-to mildly hypointense in T1-weighted (T1W)s compared to the striated muscle (Figure 4A arrow) and fat-saturated T1W (Figure 4C arrow), markedly hypointense in T2-weighted (Figure 4B arrow), with millimetric cystic areas was observed. In the post-contrast series, mild heterogeneous contrast enhancement was detected in the late phase (Figure 4D arrow and subtraction Figure E arrow), and a lobulated-contour solid lesion containing partially diffusion-restricted areas (Figures 4F and G arrows) was observed on diffusion-weighted imagings (proven by biopsy fibrothecoma)

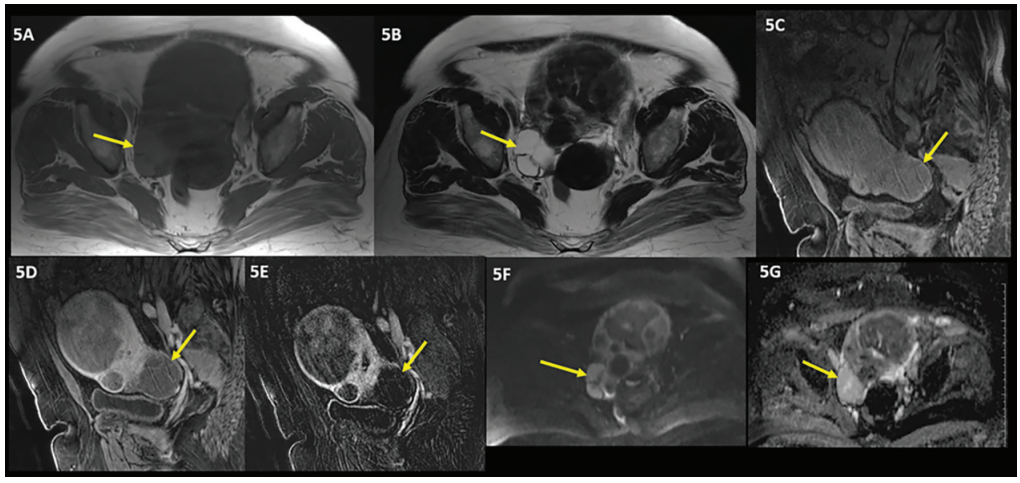


Figure 5. A 67-year-old woman referring for cervical smear. In the right ovary, a lesion sized 46x40x45 mm, slightly hyperintense compared to the striated muscle in T1-weighted (T1W) axial (Figure 5A), fatsat T1W (Figure 5C arrow), markedly hyperintense in T2-weighted axial (Figure 5B), and multiloculated cystic, was observed. In the post-contrast series, a slightly contrasting cystic lesion was observed on only one wall (Figure 5D arrow and subtraction Figure 5E arrow), and a non-diffusion-restricting cystic lesion was observed in diffusion-weighted imagings (Figures 5F and 5G) (proven by biopsy cystadenofibroma)

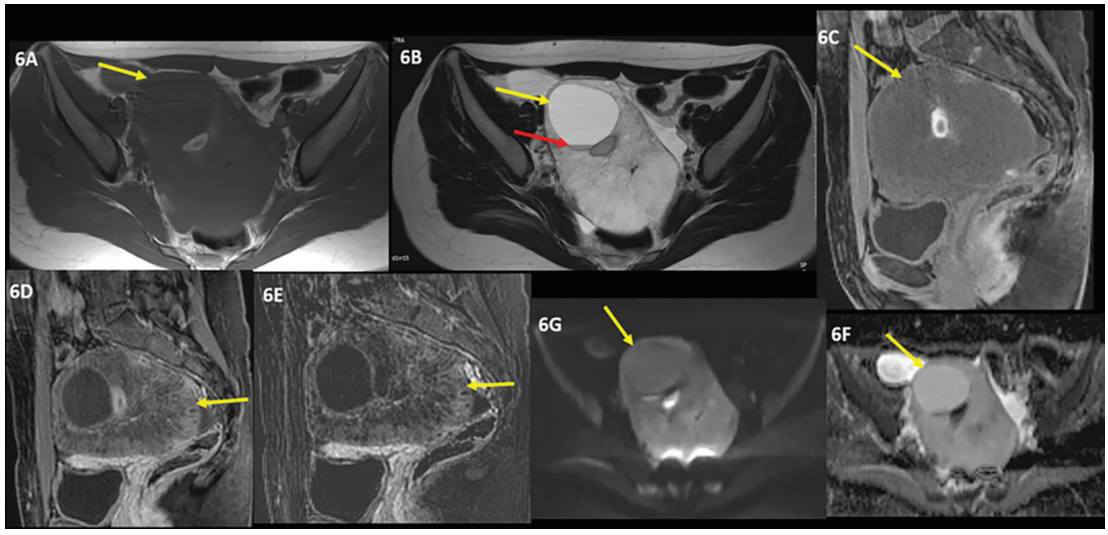


Figure 6. A 25-year-old woman with acute pelvic pain. The left ovary is 131x74x78 mm, and its size has increased compared to previous measurements. The normal stroma of the left ovary can only be seen as a crescent in a section of 30x12 mm (yellow arrows in Figures 6A and 6B). Apart from this the parenchyma signal is greatly increased, and numerous peripheral, follicle-like, millimetric, diffuse cystic areas can be observed. In the central part of the enlarged ovary, 2 hemorrhagic cystic-necrotic areas with a fluid-blood interface (Figure 6B red arrow) are observed. Additionally, there are signal changes consistent with subacute hemorrhage in the form of vaguely circumscribed areas. Slight enhancement was observed in the ovarian stroma in the postcontrast series (Figure 6D arrow and subtraction Figure 6E arrow). In diffusion-weighted imagings, increased diffusion was observed (arrows in Figures 6F and 6G). (biopsy proved sclerosing stromal tumor and ovarian torsion)

DISCUSSION

MRI is regarded as a problem-solving method for evaluating adnexal masses. When evaluating adnexal masses, gadolinium-enhanced MRI tends to be more accurate than US.

Even if solid components (such as Rokitansky nodules) are also present, the presence of fat in a cystic adnexal lesion is indicative of a cystic teratoma. Both conventional and fat-suppressed T1W imaging are necessary to demonstrate fat since the latter helps distinguish fat from blood products as the source of the high T1 signal intensity (6). Nonetheless, ovarian fibromas may be better characterized by T2W imaging. To evaluate a clinically suspected adnexal mass, US is still the predominant imaging modality, according to established practice and a review of the literature. If the outcome of the US evaluation is uncertain, MRI represents a financially advantageous course of action. Complex adnexal masses can be detected and characterized with great detection and characterization capabilities using gadolinium-enhanced MRI, and it also exhibits strong inter- and intraobserver agreement.

In case of a suspected adnexal mass, pelvic or transvaginal US is the first-step imaging modality. However, MRI's high potential for tissue characterization makes it the second step and a decisive imaging modality. MRI diagnosis of benign ovarian masses is critical in the management of patients affected by them.

ETHICS

Informed Consent: Written and verbal consent was obtained from all patients/legal representatives.

Acknowledgment

This article was previously published as a poster at the European Congress of Radiology 2023.

FOOTNOTES

Authorship Contributions

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

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

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

REFERENCES

1. Pekindil G. Over tümörlerinde görüntüleme. Trd Sem. 2015;3:108-26.
2. Jeong YY, Outwater EK, Kang HK. Imaging evaluation of ovarian masses. Radiographics. 2000;20:1445-70.
3. Kurtz AB, Tsimikas JV, Tempany CM, Hamper UM, Arger PH, Bree RL, et al. Diagnosis and staging of ovarian cancer: comparative values of doppler and conventional US, CT, and MR imaging correlated with surgery and histopathologic analysis report of the Radiology Diagnostic Oncology Group. Radiology. 1999;212:19-27.

4. Siegelmann ES, Outwater EK. Tissue characterization in the female pelvis by means of MR imaging. Radiology. 1999;212:5-18.
5. Imaoka I, Wada A, Kaji Y, Hayashi T, Hayashi M, Matsuo M, et al. Developing an MR imaging strategy for diagnosis of ovarian masses. Radiographics. 2006;26:1431-48.
6. Hricak H, Chen M, Coakley FV, Kinkel K, Yu KK, Sica G, et al. Complex adnexal masses: detection and characterization with MR imaging. Radiology. 2000;214:39-46.
7. Thomassin-Naggara I, Aubert E, Rockall A, Jalaguier-Coudray A, Rouzier R, Daraï E, et al. Adnexal masses: development and preliminary validation of an MR imaging scoring system. Radiology. 2013;267:432-43.
8. Atram M, Anshu, Sharma S, Gangane N. Sclerosing stromal tumor of the ovary. Obstet Gynecol Sci. 2014;57:405-8.

Medical Journal of Bakırköy

Med J Bakirkoy 2025;21(4):474
DOI: 10.4274/BMJ.galenos.2025.e001

Erratum

DOI: 10.4274/BMJ.galenos.2024.2024.8-12

Med J Bakirkoy 2025;21(2):172-180

The article titled "*Yılmaz D, Şahin E, Ezirmik E, Işıksaçan N, Karandere F, Arslan F. Association of IL6 with laboratory parameters, obesity, and osteoporosis in Turkish geriatric patients. Med J Bakirkoy. 2025;21(2):172-180. doi: 10.4274/BMJ.galenos.2024.2024.8-12*" published in Medical Journal of Bakırköy has been corrected due to an error in the Received and Accepted dates.

Published on page 172;

Received: 08.25.2024

Accepted: 10.05.2024

Corrected page 172;

(The corrected parts are given in bold)

Received: **25.08.2024**

Accepted: **31.10.2024**

2025 Referee Index - 2025 Hakem İndeksi

Abbas Aras
Abdullah Hızır Yavuzsan
Adnan Burak Bilgiç
Ahmet Akgül
Ahmet Aybar
Ahmet Öz
Ahmet Cem Dural
Ahmet Cem Esmer
Ahmet Ufuk Kömüroğlu
Ali Can Koluman
Aliye Çayır
Alkan Bayrak
Alpen Yahya Gümüşoğlu
Aysegul İnci Sezen
Aziz Arı
Bahar Beker
Bahar Kural
Başar Burak Çakmur
Begüm Şirin Koç
Beray Gelmez Taş
Berhan Bayram
Berkay Eren Pehlivanoğlu
Betül Güven
Buket Rende Kılıç
Bulent Tanrıverdi
Burak Bayraktar
Burak Buğday
Buruç Erkan
Celal Bozkurt
Coşkun Doğan
Çetin Murat Altay
Deniz Güzey
Derya Ece İliman
Diğdem Bezen
Dilek Altun
Ecem Eren
Eda Çelebi Bitkin
Elif Hocaoğlu
Elif Torun Parmaksız
Elif Yıldız
Emine Çalışkan
Emine Elif Altuntaş
Emine Ergül Sarı
Emrah Egemen
Emre Durdağ
Erdal Kaya
Ergül Demirçivi
Erkan Şimşek
Ersan Atahan
Esra Demirel
Esra Döğer
Evin Bozkur
Evren Karaali
Eyüp Gemici
Fatih Aydemir
Fatma Gül Demirkan
Fatma Ketenci Gencer
Fatma Nihan Turhan Çağlar
Ferit Böyük
Figen Palabıyık
Fırat Fidan
Furkan Tontu
Fusun Yıldız
Gökhan Sertcakacılar
Gökçe Çelep
Gökhan Albayrak
Gökhan Aygün
Gökhan Sayer
Gökhan Yılmaz
Gökmen Kahiloğulları
Gülşen Pınar Soydemir
Güray Koç
Gülay Koçak
Güneş Özlem Yıldız
Gürkan Mustafa Erdoğan
Güzelali Özdemir
Habip Gedik
Hakan Başar
Hakan Çakın
Hakan Yiğitbaş
Halime Ezgi Türksan
Hamide Pişkinpaşa
Hasan Hüseyin Mutlu
Hatice Toprak
Hatice Kutbay Özçelik
Hayriye Esra Ataoğlu
Hilal Özkaya
Hilal Telli
Hüseyin Aslan
Hüseyin Kılavuz
Hüseyin Hayri Kertmen
Işıl Kibar Akıllı
İbrahim Sayın
İbrahim Ethem Cakcak
İbrahim Faruk Aktürk
İsmail Yürekli
İhsan Yılmaz
İlker Solmaz
İsmet Alkış
Kadriye Kart Yaşar
Kamil Hakan Kaya
Leyla Aydın
Lütfi Şinasi Postalcı
Mahir Topaloğlu
Malik Çelik
Mediha Gönenç Ortaköylü
Medine Ayşin Taşar
Mehmet Seçer
Mehmet Soy
Mehmet Abdussamet Bozkurt
Mehmet Fatih Özdağ

2025 Referee Index - 2025 Hakem İndeksi

Melih Bestel
Melih Üçer
Melike Ersoy
Meltem Erol
Meral Mert
Merve Aldıkaçtıoğlu Talmaç
Mesut Ayer
Mesut Yılmaz
Metin Demirkaya
Mevlüt Özgür Taşkapılıoğlu
Mithat Ekşi
Müge Bilge
Muhammet Akyüz
Muharrem Kanar
Muhterem Duyu
Murat Aslan
Murat Ekin
Mustafa Yalın
Mustafa Alper Karalök
Mustafa Orhan Nalbant
Mürvet Yılmaz
Nagehan Özdemir Barışık
Necati Doğan
Nevin Hatipoğlu
Nezih Ziroğlu
Onur Çetin
Onur Kardeş
Ömer Şevgin
Özlem Altuntaş Aydın
Özlem Bekem
Özlem Harmankaya
Özlem Persil Özkan
Özlem Sancaklı
Özlem Terzi
Pelin İsmailoğlu
Peri Sohrabi Molla Yousefi
Polat Göktaş
Ramazan Atiç
Regayip Zehir
Rüstem Mustafaoğlu
Rüstü Türkay
Sabri Cavkaytar
Seçil Arıca
Seda Turgut
Seher Yıldız Tacar
Selçuk Köse
Selenay Furat Rençber
Selma Şengiz Erhan
Sema Bilgiç Gazioğlu
Sema Karakaş
Semra Işık
Semra Yüksel
Serhat Akçaalan
Şermin Dinç
Sibel Oba
Sibel Çağlar Okur
Sina Ferahman
Sinem Oral Cebeci
Süleyman Çağan Efe
Şerife Gül Karadağ
Şevki Serhat Baydın
Tevfik Çatal
Tuğba Ekiz Yılmaz
Uğur Topal
Uzay Erdoğan
Ülkem Koçoğlu Barlas
Ümit Yasar Sinan
Vedat Bakuy
Yasemin Akkoyunlu
Yasemin Çoban
Yasemin Denkboy Öngen
Yasemin Tekdöş Şeker
Yavuz Onur Danacıoğlu
Yener Özel
Yusuf Emre Özdemir
Zahide Mine Yazıcı
Zeynal Yasacı
Zeynep Çizmeci