E-ISSN: 2587-0262

NKMJ
NAMIK KEMAL MEDICAL JOURNAL

Volume: 13
Issue: 3
September 2025

ORIGINAL ARTICLES

CASE REPORTS

galenos yayınevi

www.namikkemalmedj.com

NAMIK KEMAL MEDICAL JOURNAL

EDITORIAL BOARD

Publisher Contact
Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1 34093 İstanbul, Türkiye
Phone: +90 530 177 30 975
E-mail: info@galenos.com.tr/yayin@galenos.com.tr **Web:** www.galenos.com.tr **Publisher Certificate Number:** 14521
Online Publication Date: October 2025
E-ISSN: 2587-0262
International scientific journal published quarterly.

Associate Editors

ADVISORY BOARD

ADVISORY BOARD

Please refer to the journal's webpage (**https://namikkemalmedj.com/**) 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). Namık Kemal Medical Journal (NKMJ) is indexed by **Emerging Sources Citation Index, Ulakbim Tr Index (TR DIZIN/ULAKBIM), EBSCO: CINAHL Complete, Türk Medline, Ideal Online, J-Gate, CAB International (CABI), Gale Academic OneFile, DOAJ, Embase, Chemical Abstracts Services** and **Türkiye Citation Index**.

The journal is published online.

Owner: Erdoğan Gültekin on behalf of Tekirdağ Namık Kemal University Faculty of Medicine Dean

Responsible Manager: Burçin Nalbantoğlu

NAMIK KEMAL MEDICAL JOURNAL

CONTENTS

ORIGINAL ARTICLES

NAMIK KEMAL MEDICAL JOURNAL

CONTENTS

CASE REPORTS

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.07108
Nam Kem Med J 2025;13(3):200-208

The Impact of Higher Disease Activity on Sleep Quality in Patients with Rheumatoid Arthritis

Romatoid Artrit Hastalarında Yüksek Hastalık Aktivitesinin Uyku Kalitesi Üzerine Etkisi

Elif ALTUNEL KILINÇ, İpek TÜRK

Çukurova University Faculty of Medicine, Department of Rheumatology, Adana, Türkiye

ABSTRACT

Aim: Rheumatoid arthritis (RA) is an autoimmune disease that affects joints and surrounding tissues, significantly reducing the quality of life of individuals. This study aims to evaluate RA patients' sleep quality and reveal the relationship between this impairment and disease activity.

Materials and Methods: The study was cross-sectional and involved 66 RA patients and 30 healthy controls. RA patients underwent assessments using the disease activity score 28 (DAS28), visual analogue scale (VAS), health assessment questionnaire (HAQ), and Pittsburgh sleep quality index (PSQI). The RA group was divided into two subgroups based on PSQI scores: PSQI <5 and PSQI >5.

Results: PSQI scores of RA patients were significantly higher compared to the healthy control group ($p<0.001$). Correlation analyses showed positive relationships between PSQI scores and DAS28 ($r=0.444$, $p<0.001$), VAS ($r=0.277$, $p=0.024$), and HAQ ($r=0.244$, $p=0.048$). When RA patients were divided into PSQI <5 and PSQI >5 groups, significant differences were observed in age, DAS28, and VAS scores. DAS28 and age were identified as independent risk factors for the increase in PSQI scores.

Conclusion: This study demonstrated that the sleep quality of RA patients was significantly impaired compared to healthy individuals, with worse sleep quality associated with higher disease activity and older age among RA patients. Identifying DAS28 as an independent risk factor for increased PSQI underscores the importance of disease activation in negatively impacting sleep quality. It highlights the necessity of integrating sleep quality improvement into these patients' evaluation and treatment processes.

Keywords: Rheumatoid arthritis, sleep quality, visual analog scale

ÖZ

Amaç: Romatoid artrit (RA), otoimmün bir hastalık olarak, eklemler ve çevresindeki dokuları etkileyerek bireylerin yaşam kalitesini önemli ölçüde düşürmektedir. Bu çalışma, RA hastalarının uyku kalitesini değerlendirmeyi ve bu bozulmanın hastalık aktivitesi ile ilişkisini ortaya koymayı amaçlamaktadır.

Gereç ve Yöntem: Çalışma; 66 RA hastası ve 30 sağlıklı kontrol dahil edilerek yapılmış ve kesitsel bir çalışma olarak tasarlanmıştır. RA hastalarına; hastalık aktivite skoru 28 (DAS28), vizüel analog skala (VAS), sağlık değerlendirme anketi (HAQ) ve Pittsburgh uyku kalitesi indeksi ölçeği (PUKİ) ölçekleri uygulanmıştır. RA grubundaki hastalar PUKİ <5 ve PUKİ >5 olacak şekilde iki gruba ayrılarak incelenmiştir.

Bulgular: RA hastalarının PUKİ skorları, sağlıklı kontrol grubuna göre anlamlı derecede yüksek bulunmuştur ($p<0,001$). Korelasyon analizleri, PUKİ skorları ile DAS28 ($r=0,444$, $p<0,001$), VAS ($r=0,277$, $p=0,024$) ve HAQ ($r=0,244$, $p=0,048$) arasında pozitif ilişkiler göstermiştir. RA hastaları PUKİ <5 ve PUKİ >5 olarak ayrıldığında; yaş, DAS28, ve VAS skorlarında anlamlı fark olduğu görülmüştür. DAS28 ve yaş, PUKİ skorlarındaki artış için bağımsız risk faktörüdür.

Sonuç: Bu çalışma, RA hastalarının uyku kalitesinin sağlıklı bireylere göre belirgin şekilde bozulduğunu, uyku kalitesi kötü olan RA hastalarında, hastalık aktivitesinin daha yüksek olduğunu ve yaş ortalamasının daha fazla olduğunu ortaya koymuştur. DAS28'in PUKİ artışı için bağımsız risk faktörü olması; RA hastalarında hastalık aktivasyonunun uyku kalitesini olumsuz yönde etkileyen önemli bir faktör olduğunu, bu hastaları değerlendirme ve tedavi sürescinde uyku kalitesini iyileştirmeyi tedavi süreçlerine entegrasyonunun önemini göstermektedir.

Anahtar Kelimeler: Romatoid artrit, uyku kalitesi, vizüel analog skala

Address for Correspondence: Elif ALTUNEL KILINÇ MD, Çukurova University Faculty of Medicine, Department of Rheumatology, Adana, Türkiye
E-mail: elifaltunel@hotmail.com **ORCID ID:** orcid.org/0000-0003-2501-2473
Received: 04.02.2025 **Accepted:** 23.04.2025 **Publication Date:** 07.10.2025

Cite this article as: Altunel Kılınç E, Türk İ. The impact of higher disease activity on sleep quality in patients with rheumatoid arthritis. Nam Kem Med J. 2025;13(3):200-208

INTRODUCTION

Rheumatoid arthritis (RA) is a systemic, autoimmune, and inflammatory condition that can affect various tissues and organs, including synovial joints and surrounding structures[1]. RA affects not only physical health but also the psychological and social lives of individuals. This condition leads to a decrease in patients' quality of life and difficulties in their daily activities[2].

More than half of RA patients report sleep disturbances[3]. Comprehensive phenotyping of sleep disturbances is important because associations with clinical outcomes (e.g., pain) may vary depending on the type of sleep disturbance. An experimental study showed that sleep restriction led to increased pain intensity and the number of painful joints the next day in patients with RA[4]. There is a well-established relationship between sleep disorders and overall well-being in RA. The presence of one can exacerbate the symptoms of the other. For example, inadequate disease management leading to joint pain can negatively affect sleep. Conversely, sleep disorders can reduce pain tolerance and increase arthritis-related symptoms[5-8]. Additionally, RA can also lead to psychological issues such as stress and depression, which can further negatively impact an individual's sleep quality. As a result, RA patients are more likely to experience sleep disorders than healthy individuals.

Sleep is critical to overall health and well-being. Sleep quality is a multidimensional phenomenon that affects both physical and mental health. Studies have shown that RA is closely related to sleep disorders[9-11]. These disorders affect not only the duration of sleep, but also its depth and restorative quality.

Sleep physiology is a complex process involving the regulation and quality of sleep. Sleep consists of rapid eye movement (REM) and non-REM (non-REM) stages; non-REM sleep contributes to physical repair and immune system strengthening, while REM sleep is critical for learning and emotional balance. Insufficient sleep can negatively affect pain perception, exacerbating existing pain and leading to chronicity. Additionally, sleep disorders can induce stress on the nervous system, thereby enhancing inflammatory responses. Therefore, maintaining a healthy sleep pattern plays a crucial role in pain management and improving the overall health of RA patients.

The aim of our study was to investigate the effects of high disease activity on sleep quality in RA patients and to provide a perspective on the importance of considering sleep factors alongside disease activity assessment when optimizing individualized treatment approaches to improve patients' overall quality of life.

MATERIALS AND METHODS

This cross-sectional study included 66 patients diagnosed with RA according to the 2010 American College of Rheumatology criteria who presented to the Rheumatology Outpatient Clinic of Rheumatology University Çukurova Balcalı Hospital between September and October 2024, along with a control group of 30 healthy individuals matched for sex and age[12]. The study was approved by the Ethics Committee of Çukurova University (decision no: 6/9/2024-147, date: 06.09.2024). All patients included in the study signed an informed consent form.

Technical Information

Case Selection and Description

The study included RA patients over the age of 18 with no fibromyalgia or serious neuropsychiatric disease and individuals over the age of 18 with no rheumatological disease and no diagnosis of fibromyalgia.

Demographic data (age, sex) of the patients, disease duration, smoking status, presence of extra-articular involvement, autoantibody status (rheumatoid factor and anti-cyclic citrullinated peptide antibody), disease activity score 28 (DAS28), visual analogue scale (VAS), health assessment questionnaire (HAQ), and Pittsburgh sleep quality index scale (PSQI) results were recorded. The medications used by RA patients were recorded as non-biological (conventional synthetic disease-modifying drugs-methotrexate, leflunomide, colchicine, hydroxychloroquine, steroids) and biological [anti-tumor necrosis factor (anti-TNF) alpha inhibitors, interleukin-6 (IL-6) inhibitors, IL-1 inhibitors, rituximab, JAK inhibitors, abatacept] were recorded and analyzed to evaluate the effects of treatment methods on sleep quality.

DAS28

It is a tool that evaluates 28 joints to measure disease activity in individuals diagnosed with RA. It is derived from a validated formula that considers multiple clinical parameters, including the number of swollen joints, the number of tender joints, VAS scores, and C-reactive protein levels. A DAS28 score above 5.1 indicates high disease activity, between 3.2 and 5.1 indicates moderate disease activity, and between 2.6 and 3.2 indicates low disease activity. A score below 2.6 indicates disease remission[13].

VAS

On a pain assessment scale, the patient is asked to mark numbers ranging from 0 to 10 at equal intervals on a page according to their pain sensation. 0 indicates no pain, while 10 indicates the most severe pain they have ever experienced[14,15].

HAQ

A widely used assessment for daily living activities and disability, with proven validity and reliability in musculoskeletal disorders, including regional pain syndromes and rheumatic disorders. The test consists of 20 questions. Daily living activities are divided into eight sections: dressing, standing, eating, walking, hygiene, reaching, grasping, and daily tasks. The questions are answered using a Likert scale (0=I can do it easily, 1=I can do it with some difficulty, 2=I can do it with considerable difficulty, 3=I cannot do it). A minimum of 0 and a maximum of 60 points are assigned in increments of 20, resulting in scores ranging from 0 to 3. In this assessment, a lower score indicates superior functional status[16]. The Turkish validation of the questionnaire was performed by Küçükdeveci et al.[17]

PSQI

This scale, which assesses sleep quality, consists of 18 questions and 7 components. The answers to the questions refer to sleep quality over the past month. The first 4 questions are open-ended and include temporal assessments.

The remaining questions are designed for participants to make categorical evaluations. Responses to questions evaluated in 4 separate categories are scored on a scale of 0-3. The component scores are summed to obtain the total PSQI score. The total PSQI score can range from 0 to 21. As the scores increase, sleep quality deteriorates. Those who score 5 or more out of a total of 7 components are considered to have poor sleep quality, while those with scores less than 5 are considered to have good sleep quality[18].

PSQI components are grouped as follows.

1. Subjective Sleep Quality: The participant's score for their own sleep quality. Answers range from very good (0 points) to very poor (3 points).

2. Sleep Latency: This is assessed using an open-ended question about the time it takes to fall asleep at night and a categorical question about whether this time is more than 30 minutes. The longer it takes to fall asleep, the poorer the sleep quality.

3. Sleep Duration: Measured by evaluating the participant's response to an open-ended question about their nighttime sleep duration. A score of 0 points is given for sleep duration of 7 hours or more, and 3 points for sleep duration of less than 5 hours.

4. Usual Sleep Activity: The value obtained by dividing the difference between the time of getting out of bed in the morning and the time of going to bed at night by the time spent asleep in bed is used. A score of 0 is given for values above 85.00%, and 3 points for values below 65%.

5. Sleep Disorder: Responses to 9 different questions such as waking up at night to use the restroom, feeling pain or cold during sleep are evaluated as none (0 points) during the week, less than 1 time per week (1 point), 1-2 times per week (2 points), and 3 or more times per week (3 points). The scores obtained from the 9 questions are added together to obtain the sleep disorder score.

6. Use of Sleep Medication: The use of sleep medication to help sleep in the last month is asked. It is evaluated on a scale from never used (0 points) to used 3 times or more per week (3 points).

7. Daytime Dysfunction: This is evaluated based on the total score obtained from the answers to 2 questions asking whether the respondent felt sleepy during a daytime activity in the past month and whether this feeling interfered with their work.

Statistical Analysis

Statistical analyses were performed using SPSS for Windows version 25.0. The normality of variables was assessed using the Kolmogorov-Smirnov and Shapiro-Wilk tests.

A p-value greater than 0.05 in the Kolmogorov-Smirnov test indicates that the data are normally distributed. In our study, continuous quantitative parameters that showed normal distribution were evaluated using the t-test for two-group comparisons, the Mann-Whitney U test when not normally distributed, and the chi-square test for categorical variables. When there were three subgroups and the data did not follow a normal distribution, the Kruskal-Wallis test was used. Descriptive statistics included mean, standard deviation, median, min, and max values. Spearman's correlation test was used for correlation analysis, and the correlation coefficient was denoted as rho. A correlation coefficient less than 0.25 indicates no relationship or a very weak relationship; a coefficient between 0.25 and 0.5 indicates a weak to moderate relationship; a coefficient between 0.5 and 0.75 indicates a strong relationship; and a coefficient greater than 0.75 indicates a very strong relationship. A logistic regression analysis was conducted to assess the diagnostic power of the measurement parameters. A p-value of less than 0.05 was considered statistically significant.

RESULTS

Our study included 66 RA patients (87.8% female, mean age: 55.6±11.12) and 30 healthy controls (83.3% female, mean age: 51.4±9.3). There was no difference between the groups in terms of age and sex distribution (p=0.865 and p=0.778). According to the Kolmogorov-Smirnov and Shapiro-Wilk tests, all continuous quantitative data showed a normal distribution, while the other parameters did not show a normal distribution (p=0.200, p<0.001 for all). The sample size was

determined based on the study by Taylor-Gjevre et al.[3], which recommended n=19 cases with a 5% margin of error, 80% power, and a standard effect size of 0.64[19].

Median disease duration of RA patients was 11.5 (min-max: 0-30 years). Fifty-one (77%) of the patients were non-smokers, while 15 (23%) were active smokers or had a history of smoking. Fifty-nine (89%) patients had no extra-articular involvement. Forty-five (68%) patients were seropositive. Mean DAS28 scores were 3.8±1.8, mean VAS scores were 1.61±2.4, median was 0 (range: 0-9), and median HAQ score was 0 (range: 0-2.6). Demographic data of RA patients are presented in Table 1. Twenty percent (n=6) of the healthy control group had a history of active smoking or smoking.

Of the 66 RA patients included in the study, 50 (76%) were receiving nonbiologic therapy, while 16 (24%) were receiving biologic therapy. Five (7.6%) of the 66 RA patients were using psychiatric medications. Of the 5 patients, 3 (60%) had a diagnosis of generalized anxiety disorder, and 2 (40%) had a diagnosis of depressive disorder. One patient was on essitalopram + mirtazapine + trifluperazine, one patient was on mirtazapine, one patient was on essitalopram, one patient was on vortioxetine, and one patient was on agomelatin.

When the PSQI questionnaires of RA patients were analyzed; total sleep duration in hours was on average: 6.5±1.8, median: 7 (min-max: 2-9), sleep onset latency in minutes was on average: 31.3±38, median: 20 (min-max: 0-180), sleep efficiency in percentage, average: 81.2±19, median: 88 (min-max: 31-100).

Table 1. Demographic data of patients with rheumatoid arthritis

Number of patients, n	66
Age; year, mean ± SD	55.6±11.12
Female, n (%)	58 (87.8)
Disease duration, year, median (min-max)	11.5 (0-30)
Smoking*, n (%)	15 (23)
Extraarticular involvement, n (%)	7 (11)
Seropositive, n (%)	45 (68)
Treatment, n (%)	
csDMARD	50 (75.8)
Biologic	16 (24.2)
DAS28, median (min-max)	0.9 (0-7.2)
VAS, median (min-max)	0 (0-9)
PSQI, median (min-max)	4.75 (0-18)
HAQ, median (min-max)	0 (0-2.6)

*: Active smoker or smoking history, csDMARD: Conventional synthetic disease-modifying drug, DAS28: Disease activity score 28, PSQI: Pittsburgh sleep quality index, SD: Standard deviation, VAS: Visual analogue scale, HAQ: Health assessment questionnaire

The question about difficulty falling asleep within 30 minutes was answered as "never" by 24 (66%) patients, "less than once a week" by 13 (19.6%) patients, "once or twice a week" by 11 (16.6%) patients, and "three or more times a week" by 18 (27%) patients. The question about sleep medication use was answered as follows: 60 (89.5%) patients reported "none in the past month", 2 (3%) patients reported "once a week or less", 1 (1.5%) patient reported "once or twice a week", and 4 (6%) patients reported "three or more times a week". Subjective sleep quality (SSQ) was reported as "very good" by 6 (9.5%) patients, "fairly good" by 36 (54.5%) patients, "fairly poor" by 16 (24%) patients, and "very poor" by 8 (12%) patients (Table 2).

The average PSQI score for RA patients was 6.1±4.7, median: 4.75 (min-max: 0-18), while the average PSQI score for the healthy control group was 3.2±1.8, median: 3 (min-max: 0-9). There was a significant difference between the two groups (p<0.001). A comparative evaluation of the total sleep duration, sleep onset latency, sleep efficiency, inability to fall asleep within 30 minutes, use of sleep medication, and SSQ results between RA patients and the healthy control group is presented in Table 3.

In the correlation analysis conducted in RA patients, PSQI scores were positively correlated with DAS28 (r=0.444, p<0.001), VAS (r=0.277, p=0.024), age (r=0.262, p=0.035), and HAQ (r=0.244, p=0.048), while no correlation was found with disease duration (p=0.241). When DAS28 scores were categorized as low disease activity (Group 1) for scores <3.2, moderate disease activity (Group 2) for scores between 3.2 and 5.1, and high disease activity (Group 3) for scores >5.1, 15 patients were assigned to Group 1, 15 to Group 2, and 36 to Group 3. Analysis between subgroups revealed significant differences in PSQI scores (p=0.008). PSQI scores were higher in Group 3 than in Group 2 (p=0.049) and higher in Group 2 than in Group 1 (p=0.002) (Table 4).

RA patients were divided into two groups: those with PSQI <5 (good sleep quality) and those with PSQI >5 (poor sleep quality). There were 33 patients in each group. The mean age of RA patients with good sleep quality was 52±2 years, while the mean age of RA patients with poor sleep quality was 60±1.7 years, which was statistically significant (p=0.044). The sex distribution was equal in both groups, with 29 (87.8%) women and 4 (12.2%) men. There were no differences between the two groups in terms of year of diagnosis, medication used (biological or conventional synthetic disease-modifying drug), autoantibody status (seropositive/seronegative), extra-articular involvement, and smoking status. (p=0.690, p=0.566, p=0.792, p=0.230, p=0.314, respectively). The DAS28 score was 2.2±0.8 in RA patients with good sleep quality, while it

was 5.3±1.1 in the RA group with poor sleep quality, and this difference was statistically significant (p<0.001). Similarly, VAS scores were 0.97±0.2 (median: 0, min-max: 0-8) in the RA group with good sleep quality and 2.1±0.5 (median: 0, min-max: 0-9) in the RA group with poor sleep quality, and this difference was statistically significant (p=0.042). There was no difference in HAQ scores (p=0.162).

Logistic regression analysis of RA patients evaluated according to PSQI scores is presented in Table 5.

Table 2. Comparative evaluation of PSQI components in rheumatoid arthritis and healthy control groups

	Rheumatoid arthritis, n=66	Healthy control group, n=30	p-value
Total sleep duration, hours, median (min-max)	7 (2-9)	8 (5-12)	0.784
Sleep latency, minutes	20 (0-180)	15 (0-30)	<0.001*
Sleep activity, %	55 (31-100)	90 (50-100)	<0.001*
Inability to fall asleep within 30 minutes, n (%)			
None	24 (66)	22 (73.3)	0.654
Less than once a week	13 (19.6)	4 (13.3)	0.048*
Once or twice a week	11 (16.6)	2 (6.7)	<0.001*
Thrice or more in a week	18 (27)	2 (6.7)	<0.001*
Use of sleep medication, n (%)			
None	59 (89.5)	28 (93)	0.723
Less than once a week	2 (3)	1 (3.5)	0.882
Once or twice a week	1 (1.5)	1 (3.5)	0.098
Thrice or more in a week	4 (6)	0	0.03*
Subjective sleep quality, n (%)			
Very good	6 (9.5)	24 (79.9)	<0.001*
Quite good	16 (24)	2 (6.7)	<0.001*
Quite bad	36 (54.5)	2 (6.7)	<0.001*
Very bad	8 (12)	2 (6.7)	<0.001*

*p<0.05, PSQI: Pittsburgh sleep quality index

Table 3. Intergroup evaluation of PSQI score distributions

	Rheumathoid arthritis, n=66	Healthy control group, n=30	p-value
PSQI score, median (min-max)	4.75 (0-18)	3 (0-9)	<0.001*
PSQI components, median (min-max)			
Sleep quality	2 (0-3)	0 (0-3)	<0.001*
Sleep latency	0 (0-3)	0 (0-3)	<0.001*
Sleep duratipn	1 (0-3)	1 (0-3)	0.784
Normal sleep activity	2 (0-3)	0 (0-3)	<0.001*
Sleep disorder	1 (0-3)	0 (0-2)	<0.001*
Use of sleep medication	0 (0-3)	0 (0-2)	0.428
Daytime dysfunction	0 (0-3)	0 (0-2)	<0.001*

*: p<0.05, PSQI: Pittsburgh sleep quality index

Table 4. Evaluation of PSQI scores in patients with rheumatoid arthritis grouped according to disease activity

	n (%)	PSQI score mean ± SD	p-value*
Group 1	15 (28)	4.5±5	0.002[a]
Group 2	15 (28)	6.198±3	0.049[c]
Group 3	36 (46)	7.4±6.6	0.001[b]

*: The Kruskal-Wallis test, Group 1: Low disease activity, Group 2: Medium disease activity, Group 3: High disease activity, [a]: Group 1-Gruop 2, [b]: Group 1-Group 3, [c]: Group 2-Group 3, SD: Standard deviation, PSQI: pittsburgh sleep quality index

Table 5. Regression analysis results based on PSQI scores

	Age	DAS28	VAS	Psychiatric diagnosis/ treatment	RA drug use (biologic/non-biologic)
B	0.080	0.701	0.053	1.885	0.499
SE	0.034	0.243	0.123	1.160	0.587
OR (95% CI)	1.083 (1.014-1.158)	2 (1.252-3.243)	1.055 (0.828-1.343)	6.5 (0.679-64)	1.647 (0.521-5.2)
p-value	0.018*	0.004*	0.666	0.104	0.395

*p<0,05, B: Regression coefficient, DAS28: Disease activitty score 28, SE: Standard error, OR: Odds ratio, CI: Confidence interval, PSQI: Pittsburgh sleep quality index, VAS: Visual analogue scale, RA: Rheumatoid arthritis

DISCUSSION

In this study, we evaluated the sleep quality of RA patients by comparing them with a healthy control group. The results show that PSQI scores are significantly higher in RA patients than in healthy individuals. PSQI scores of RA patients correlated with factors such as DAS28, VAS, HAQ, and age. This correlation highlights the complex relationship between RA patients' quality of life, pain, disease activity, and sleep quality. When RA patients were grouped based on their PSQI scores into those with good and poor sleep quality, the group with poor sleep quality had higher DAS28 and VAS scores, as well as a higher average age. When looking specifically at the factors affecting PSQI scores, the recording of DAS28 and age as independent risk factors highlights the effects of these factors on sleep quality.

A review of the literature shows that the first studies on this subject date back to 2009. Kiper and Sunal[10] conducted a comprehensive study involving 150 RA patients and 150 healthy control subjects. In this study, it was determined that PSQI scores were associated with social and demographic factors such as female sex, marital status, education level, monthly income, exercise status, and smoking. However, they did not find any relationship with certain parameters such as age, disease duration, and medications used for treatment[10]. Similarly, in our study, no relationship was found between disease duration and medications used for treatment and PSQI scores. In contrast, while this study found a relationship with female sex and smoking, our study found a relationship with age and disease activity, which was not evaluated in this study. In Kiper and Sunal'study[10], the mean age of the RA group was determined as 49.5±12.44, while in our study, this mean was found to be 55.6±11.12. Considering that sleep quality may deteriorate in older individuals, the lack of association with age can be explained by the younger age of the patients included in the study. This supports the idea that age-related factors in RA patients need to be further investigated in terms of their effect on sleep quality[19]. In conclusion, while Kiper and Sunal' study[10] emphasized the relationship between demographic factors and PSQI scores, our study draws attention to the importance of different parameters such as disease activity. These differences may stem from demographic and clinical differences (seropositivity, extra-articular involvement, smoking, etc.) in the populations used in the studies. Therefore, future research conducting in-depth analyses on larger and more diverse patient groups could contribute to personalized treatment strategies aimed at improving the quality of life of RA patients.

In their study involving 55 RA patients and 20 healthy controls, Yiğit et al.[9] reported that PSQI scores were higher in RA patients than in the healthy control group, and that PSQI was correlated with DAS28. They also reported that sleep quality was better in patients using anti-TNF alpha inhibitors than in those not using them, although this difference was not statistically significant. HAQ scores were not evaluated in this study. Similarly, İnanır et al.[11] also found higher PSQI scores in the RA group in a study involving 20 RA patients and 20 healthy controls. In our study, similar results were obtained in terms of higher PSQI scores in RA patients compared to healthy individuals and correlation with DAS28. In contrast, Yiğit et al.[9] found that sleep quality was higher in patients using anti-TNF alpha inhibitors. When comparing the two studies, our study had a higher female sex ratio (87.8% vs. 72%), longer disease duration (11.5 years vs. 6.3 years), and higher anti-TNF inhibitor use (24% vs. 12%). These differences highlight the complexity of factors affecting sleep quality in RA patients and the importance of personalized treatment strategies. Future research may contribute to a better understanding and management of sleep disorders associated with RA, which could provide important steps toward improving patient quality of life. It is also important to highlight the association between an increase in DAS28 scores and worsening sleep quality, as revealed by this study. Pain, inflammation, and sleep are interconnected mechanisms that influence one another. High levels of pain and inflammation in RA patients can directly affect sleep quality. Pain, in turn, can negatively affect both the physical and psychological aspects of sleep, making it difficult for patients to achieve deep sleep. Insufficient sleep shortens sleep duration while also reducing sleep depth and restorative quality, creating a vicious cycle; because without adequate sleep, pain and inflammation

may increase[20]. On the other hand, the relationship between sleep disorders and inflammation has also been an important area of research. Sleep deprivation can lead to increased levels of pro-inflammatory cytokines [e.g., (TNF-alpha, IL-6)]. These cytokines can promote inflammation and increase disease activity. Many studies have shown that poor sleep quality has negative effects on inflammation markers and can increase the severity of inflammation and pain[21,22]. While the relationship between sleep and pain is a vicious cycle, a trend in the literature suggests that the temporal effect of sleep on pain may be stronger than the effect of pain on sleep. Therefore, sleep disorder is considered the starting point. From this perspective, in diseases accompanied by chronic pain, sleep disorders should be investigated and necessary measures should be taken, as they are expected to increase pain.

In terms of sleep duration, no differences were found between RA patients and healthy control groups in the studies by Kiper and Sunal[10], Yiğit et al.[9], and İnanır et al.[11]. Similarly, no differences were found in our study. This suggests that, despite adequate sleep duration in RA patients, other factors that negatively affect sleep quality, such as sleep fragmentation, difficulty falling asleep, or feeling unrested after waking, may be at play. Therefore, adequate sleep duration alone does not imply good sleep quality, and it is necessary to evaluate broader factors that affect sleep quality.

In the study conducted by Doğan et al.[22], 92 RA patients were divided into two groups based on low and high DAS28 scores. This study is important in terms of examining the relationship between DAS28 scores and PSQI scores, which assess sleep quality with higher VAS scores. The results show that VAS scores are also higher in the group with high DAS28 scores; in addition, higher values were found in this group for PSQI scores and sleep disorder subgroup assessment. However, it is noteworthy that the total PSQI scores were similar between the two groups, suggesting that many factors other than DAS28 may play a role in the assessment of sleep quality. When the demographic characteristics of the patients were examined, the female sex ratio was reported as 73.9% in the study. In addition, the rate of patients with a disease duration of more than 5 years was found to be 79.3%[22]. These results are important in understanding the effects of sex and disease duration on sleep quality. Although the study differs from ours in terms of methodology, it is interesting that the rate was lower than in our study, whereas female sex was expected to increase PSQI scores. Additionally, the high median disease duration of 11.2 years (range: 0-30) in our study is considered a potential factor contributing to the differences in results.

In a recent study by Yaseen et al.[23], 385 RA patients were examined for sleep disorders, and it was reported that only 9% of all patients had no sleep disorders, 38.9% had insomnia, and other sleep disorders such as sleep apnea and hypersomnia were present in lower proportions. When the PSQI scores in the study were examined, poor sleep quality was found in 75% of patients, and all sleep disorders were shown to be associated with DAS28[23]. Compared to our study, our demographic data were similar, but PSQI scores were higher than in our study. We believe that the fact that only 20% of the patients included in the study by Yaseen et al.[23] were in remission, the exclusion of fibromyalgia patients in our study, and the lower DAS28 scores may account for the difference in PSQI results.

The association between increased disease activity in RA patients, i.e., high DAS28 scores, and poor sleep quality, as well as the identification of poor sleep quality as an independent risk factor, are the most important findings of our study. There are studies in the literature that support this result[24,25]. When we look at the mechanism, we see that pro-inflammatory cytokines such as TNF-alpha and IL-6, which contribute to sleep regulation, are elevated in patients with RA. These cytokines increase pain sensitivity and contribute to nighttime joint pain, making it difficult for patients to fall asleep or stay asleep. This interaction between pain and inflammation probably explains the higher prevalence of insomnia in patients with high disease activity[18]. This situation emphasizes the importance of evaluating RA patients for sleep disorders and of a multidisciplinary approach.

The results of this study show that there is a strong relationship between sleep quality and disease activity in RA patients. The association between high DAS28 scores and poor sleep quality in RA patients can be explained by inflammation and pain mechanisms. Increased levels of pro-inflammatory cytokines such as TNF-alpha and IL-6 increase pain sensitivity, leading to fragmented sleep throughout the night and difficulty transitioning to deep sleep. Therefore, effective inflammation control and pain management strategies are necessary to improve sleep quality in RA patients. It is crucial to consider not only physical symptoms but also psychological and social factors in RA patients. Additionally, other studies in the literature indicate that poor sleep quality in RA patients has significant negative effects on quality of life and may be associated with other comorbidities. For example, insufficient sleep can increase the risk of psychiatric disorders such as depression and anxiety, which can further worsen patients' overall health[26]. In the treatment of RA, considering patients' psychosocial status as well as their physical health will increase the success of the treatment process and provide patients with more comprehensive care.

Study Limitations

This cross-sectional study has some limitations. First, the number of participants is limited. The assessments are based on the participants' own reports, which carries a risk of subjective bias; participants' ways of assessing factors such as pain and sleep quality may vary depending on their individual perceptions.

The wide confidence interval identified in the analysis evaluating the psychiatric medications used by the patients included in the study suggests that these medications are not reliable predictors of sleep quality. In addition, the number of patients using psychiatric medications (5 patients) is quite small, which makes it difficult to interpret the potential effects of psychiatric medications on sleep quality. In future studies, more comprehensive analyses of the effects of such drugs should be conducted with larger and more representative samples. An important limitation of this study is that potential confounding variables (e.g., medication use, comorbidities, and socioeconomic factors) that may affect sleep quality and disease activity were not sufficiently included in the analysis. In future studies, taking these variables into account will enable more comprehensive and reliable results to be obtained. Finally, the lack of evaluation of other biopsychosocial factors in addition to disease activity and treatment process variables limits the achievement of a broader perspective on the relationship between sleep quality and RA patients. It is important to take these limitations into account in the design of future studies and in the interpretation of results. The strengths of our study are that it is a rare recent study on this topic, that it includes scores such as VAS and HAQ to address sleep disorders from different perspectives, and that it aims to achieve more objective results by including a healthy control group.

CONCLUSION

Our study reveals that sleep quality is an important factor that negatively affects RA and that RA patients have significantly poorer sleep quality than healthy individuals. The identification of DAS28 and age as independent risk factors for the PSQI score strengthens the complex relationship between pain, disease activity, and sleep quality.

Therefore, we emphasize the necessity of considering sleep health in the treatment process of patients. This study contributes to identifying the negative effects of sleep quality on RA patients and suggests the need to develop strategies to improve sleep health. Future studies could enhance our understanding in this area by conducting more detailed investigations into the factors influencing sleep quality in larger sample groups and individuals with different demographic characteristics.

Ethics

Ethics Committee Approval: The study was approved by the Ethics Committee of Çukurova University (decision no: 6/9/2024-147, date: 06.09.2024).

Informed Consent: All patients included in the study signed an informed consent form.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Briggs AM, Fary RE, Slater H, Ranelli S, Chan M. Physiotherapy co-management of rheumatoid arthritis: identification of red flags, significance to clinical practice and management pathways. Man Ther. 2013;18:583-7.
2. Martinec R, Pinjatela R, Balen D. Guality of life in patients with rheumatoid arthrıtıs - a preliminary study. Acta Clin Croat. 2019;58:157-66.
3. Taylor-Gjevre RM, Gjevre JA, Nair B, Skomro R, Lim HJ. Components of sleep quality and sleep fragmentation in rheumatoid arthritis and osteoarthritis. Musculoskeletal Care. 2011;9:152-9.
4. Irwin MR, Olmstead R, Carrillo C, Sadeghi N, Fitzgerald JD, Ranganath VK, et al. Sleep loss exacerbates fatigue, depression, and pain in rheumatoid arthritis. Sleep. 2012;35:537-43.
5. Finan PH, Goodin BR, Smith MT. The association of sleep and pain: an update and a path forward. J Pain. 2013;14:1539-52.
6. Wong ITY, Chandran V, Li S, Gladman DD. Sleep disturbance in psoriatic disease: prevalence and associated factors. J Rheumatol. 2017;44:1369-74.
7. Westhovens R, Van der Elst K, Matthys A, Tran M, Gilloteau I. Sleep problems in patients with rheumatoid arthritis. J Rheumatol. 2014;41:31-40.
8. Wolfe F, Michaud K, Li T. Sleep disturbance in patients with rheumatoid arthritis: evaluation by medical outcomes study and visual analog sleep scales. J Rheumatol. 2006;33:1942-51.
9. Yiğit Gökçe İ, Ünlü M, Solak Ö, Dündar Ü, Demirdal Ü S, Fidan F. Romatoid artrit hastalarının pittsburgh uyku kalite indeksi ile değerlendirilmesi. Turk J Phys Med Rehab. 2009;55.
10. Kiper S, Sunal N. Artritli hastalarda uyku kalitesinin değerlendirilmesi. KTD. 2009;10:33-9.
11. İnanır A, Okan S, İnanır S, Erkorkmaz Ü. Romatoid artritli hastalarda uyku kalitesinin değerlendirilmesi (The evaluation of the quality of sleep in patients with rheumatoid arthritis). Gaziosmanpaşa Tıp Dergisi. 2010;2:43-50.
12. Kay J, Upchurch KS. ACR/EULAR 2010 rheumatoid arthritis classification criteria. Rheumatology (Oxford). 2012;51 (Suppl 6):vi5-vi9.
13. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatol. 2003;42:244-57.
14. Morrey BF. The elbow and its disorders. St Louis: Elsevier Health Sciences. 2009;6:626-56.

15. Fedorczyk JM. Elbow tendinopathies: clinical presentation and therapist's management of tennis elbow. In: Rehabilitation of the Hand and Upper Extremity. 2011;6:1098-1108.

16. Bruce B, Fries JF. The health assessment questionnaire (HAQ). Clin Exp Rheumatol. 2005;23:S14.

17. Küçükdeveci AA, Sahin H, Ataman S, Griffiths B, Tennant A. Issues in cross-cultural validity: example from the adaptation, reliability, and validity testing of a Turkish version of the Stanford Health Assessment Questionnaire. Arthritis Rheum. 2004;51:14-9.

18. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality ındex: a new instrument for psychiatric practice and research. Psychiatry Res. 1989;28:193-213.

19. Miyata S, Noda A, Iwamoto K, Kawano N, Okuda M, Ozaki N. Poor sleep quality impairs cognitive performance in older adults. Journal of Sleep Research. 2013;22:535-41

20. Tański W, Świątoniowska-Lonc N, Tomasiewicz A, Dudek K, Jankowska-Polańska B. The impact of sleep disorders on the daily activity and quality of life in rheumatoid arthritis patients - a systematic review and meta-analysis. Eur Rev Med Pharmacol Sci. 2022;26:3212-29.

21. Gouda W, Mokhtar M, Elazab SA, Alreefi R, Alomar T, Kushk F, et al. Sleep disorders in patients with rheumatoid arthritis: association with quality of life, fatigue, depression levels, functional disability, disease duration, and activity: a multicentre cross-sectional study. J Int Med Res. 2023;51:3000605231204477.

22. Doğan İ, Gök K, Apaydın H, Armağan B. Evaluation of quality of life scales according to disease activity in rheumatoid arthritis. Acıbadem Üniversitesi Sağlık Bilimleri Dergisi. 2022;13:374-8.

23. Yaseen A, Yasin S, Ul Haq Z, Ahmad I, Hussain M, Pervez N. Frequency of sleep disorders in rheumatoid arthritis patients. Cureus. 2024;16:e72417.

24. Sariyildiz MA, Batmaz I, Bozkurt M, Bez Y, Cetincakmak MG, Yazmalar L, et al. Sleep quality in rheumatoid arthritis: Relationship between the disease severity, depression, functional status and the quality of life. J Clin Med Res. 2014;6:44-52.

25. Austad C, Kvien TK, Olsen IC, Uhlig T. Sleep disturbance in patients with rheumatoid arthritis is related to fatigue, disease activity, and other patient-reported outcomes. Scand J Rheumatol. 2017;46:95-103.

26. Druce KL, Cordingley L, Short V, Moore S, Hellman B, James B, et al. Quality of life, sleep and rheumatoid arthritis (QUASAR): a protocol for a prospective UK mHealth study to investigate the relationship between sleep and quality of life in adults with rheumatoid arthritis. BMJ Open. 2018;8:e018752.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.57431
Nam Kem Med J 2025;13(3):209-221

Naltrexone Induced Agitation Management: Employing a Hybrid Artificial Neural Network Model to Determine the Appropriate Dosage of Intravenous Diazepam

Naltrekson Kaynaklı Ajitasyon Yönetimi: İntravenöz Diazepamın Uygun Dozajını Belirlemek için Hibrit Yapay Sinir Ağı Modeli Kullanılması

Seyed Ali MOHTARAMI[1], Sayed Masoud HOSSEINI[2], Shahin SHADNIA[2], Babak MOSTAFAZADEH[2], Peyman Erfan Talab EVINI[2], Omid MEHRPOUR[3], Shakiba FARHADPUR[4], Mitra RAHIMI[2]

[1]Payame Noor University (PNU) Faculty of Engineering, Department of Computer Engineering and Information Technology, Tehran, Iran
[2]Toxicological Research Center, Excellence Center of Clinical Toxicology, Department of Clinical Toxicology, Loghman Hakim Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[3]Michigan Poison & Drug Information Center, Wayne State University Faculty of Medicine, Michigan, USA
[4]Department of Pharmacology, School of Pharmacy, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ABSTRACT

Aim: Agitation is a prevalent symptom of opioid withdrawal caused by naltrexone. Managing agitation induced by naltrexone poses challenges as current drug interventions are either ineffective or require intensive care. This study sought to determine the most effective diazepam dosage for alleviating naltrexone-induced agitation.

Materials and Methods: This research examined a dataset comprising 615 patient medical records from Loghman Hakim Hospital in Tehran, Iran, focusing on cases of patients experiencing naltrexone-induced agitation. The dataset included individuals who were administered low-dose diazepam (<10 mg; 383 cases) and high-dose diazepam (>10 mg; 232 cases). The predictive performance of the developed models was assessed based on metrics such as accuracy, specificity, sensitivity, F1-score, and ROC curve analysis.

Results: The bat algorithm demonstrated the highest performance among meta-heuristic algorithms, achieving a score of 89.5% (0.895) at iteration 128. A comparative evaluation of five decision tree classifiers revealed that the Extra Trees Classifier surpassed others, attaining an accuracy of 0.8649, sensitivity of 0.8649, precision of 0.8645, F1-score of 0.8649, and area under the curve (AUC) of 0.9343. Following the determination of feature importance and training of a multilayer perceptron neural network with weighted features, the model exhibited superior performance with an accuracy of 0.91, sensitivity of 0.9, precision of 0.92, F1-score of 0.91, and AUC of 0.94.

Conclusion: Features for predicting the appropriate dose of diazepam in patients with naltrexone-induced agitation included recent opioid use, Richmond Agitation-Sedation scale, amount of ingested naltrexone, pulse rate, systolic blood pressure, level of consciousness, serum levels of sodium, creatinine, and lactate dehydrogenase. Our research findings indicate that a weighted multilayer perceptron neural network shows promise in accurately forecasting the necessity of increased doses of diazepam for patients experiencing naltrexone-induced agitation. This is particularly evident when utilizing meta-heuristic techniques for feature selection and assigning importance of selected features based on the classifier with the highest AUC. This model could guide clinicians in tailoring diazepam doses to manage naltrexone induced agitation safely.

Keywords: Naltrexone, agitation, diazepam, meta-heuristic algorithm, multilayer perceptron, machine learning

Address for Correspondence: Mitra RAHIMI MD, Toxicological Research Center, Excellence Center of Clinical Toxicology, Department of Clinical Toxicology, Loghman Hakim Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
E-mail: mrahimi744@gmail.com **ORCID ID:** orcid.org/0000-0002-8785-8207
Received: 07.12.2024 **Accepted:** 02.05.2025 **Publication Date:** 07.10.2025

Cite this article as: Rahimi M, Mohtarami SA, Hosseini SM, Shadnia S, Farhadpour S, Mehrpour O, et al. Naltrexone induced agitation management: employing a hybrid artificial neural network model to determine the appropriate dosage of intravenous diazepam. Nam Kem Med J. 2025;13(3):209-221

ÖZ

Amaç: Ajitasyon, naltreksonun neden olduğu opioid yoksunluğunun yaygın bir semptomudur. Naltreksonun neden olduğu ajitasyonu yönetmek, mevcut ilaç müdahalelerinin ya etkisiz olması ya da yoğun bakım gerektirmesi nedeniyle zorluklar ortaya çıkarmaktadır. Bu çalışma, naltreksonun neden olduğu ajitasyonu hafifletmek için en etkili diazepam dozajını belirlemeyi amaçlamaktadır.

Gereç ve Yöntem: Bu araştırma, İran'ın Tahran kentindeki Loghman Hakim Hastanesi'nden alınan 615 hasta tıbbi kayıtlarından oluşan bir veri setini incelemiş ve naltreksonun neden olduğu ajitasyon olgularına odaklanmıştır. Veri seti, düşük doz diazepam (<10 mg; 383 olgu) ve yüksek doz diazepam (>10 mg; 232 olgu) verilen bireyleri içermektedir. Geliştirilen modellerin tahmin performansı, doğruluk, özgüllük, duyarlılık, F1 skoru ve ROC eğrisi analizi gibi metriklere göre değerlendirildi.

Bulgular: Yarasa algoritması, meta-sezgisel algoritmalar arasında en yüksek performansı göstererek, 128. yinelemede %89,5 (0,895) puan aldı. Beş karar ağacı sınıflandırıcısının karşılaştırmalı değerlendirmesi, Ekstra Ağaç Sınıflandırıcı diğerlerini geride bırakarak 0,8649 doğruluk, 0,8649 duyarlılık, 0,8645 kesinlik, 0,8649 F1 skoru ve 0,9343 eğri altındaki alan (AUC) elde ettiğini ortaya koymuştur. Özelliklerin öneminin belirlenmesi ve ağırlıklı özelliklere sahip çok katmanlı bir algılayıcı sinir ağının eğitilmesinin ardından, model 0,91 doğruluk, 0,9 duyarlılık, 0,92 kesinlik, 0,91 F1 skoru ve 0,94 AUC ile üstün performans sergilemiştir. Naltrekson kaynaklı ajitasyon hastalarında uygun diazepam dozunu tahmin etmek için kullanılan özellikler arasında son zamanlarda opioid kullanımı, Richmond Ajitasyon-Sedasyon ölçeği, alınan naltrekson miktarı, nabız hızı, sistolik kan basıncı, bilinç düzeyi, serum sodyum, kreatinin ve laktat dehidrojenaz düzeyleri yer almaktadır.

Sonuç: Araştırma bulgularımız, ağırlıklı çok katmanlı algılayıcı sinir ağının, naltrekson kaynaklı ajitasyon yaşayan hastalar için diazepam dozunun artırılması gerekliliğini doğru bir şekilde tahmin etmede umut vaat ettiğini göstermektedir. Bu, özellikle özellik seçimi için meta-sezgisel teknikler kullanıldığında ve en yüksek AUC'ye sahip sınıflandırıcıya göre seçilen özelliklerin önemine göre atandığında belirgindir. Bu model, klinisyenlere naltrekson kaynaklı ajitasyonu güvenli bir şekilde yönetmek için diazepam dozlarını ayarlamada rehberlik edebilir.

Anahtar Kelimeler: Naltrekson, ajitasyon, diazepam, meta-sezgisel algoritma, çok katmanlı algılayıcı, makine öğrenimi

INTRODUCTION

Naltrexone, an opioid antagonist, is increasingly being prescribed to address opioid use disorder, alcohol use disorder, and chronic pain[1]. This non-selective opioid antagonist is commonly used for maintenance therapy in opioid dependency due to its long-acting nature and high affinity to μ receptors. The primary long-acting metabolite of naltrexone, six-beta-naltrexone, extends the antagonistic effects of naltrexone on narcotic receptors[2].

There are three potential scenarios in which a naltrexone-related fatality may occur: opioid overdose during oral naltrexone treatment as patients may consume high doses of opioids in an attempt to bypass the blockade; opioid overdose following the cessation of naltrexone treatment as individuals resume opioid use after treatment and lose their tolerance to opioids; and toxicity from naltrexone itself, particularly harmful to the liver in doses exceeding five times the recommended safe dose[3,4].

In preclinical investigations involving opioid-naïve animal subjects, naltrexone demonstrated a relatively low level of acute toxicity[5]. It was observed that naltrexone led to a non-significant reduction in respiratory rate and pupillary size, as well as a significant decrease in body temperature among five individuals with a history of addiction. Clinical trials conducted on opioid use disorder revealed that naltrexone effectively counteracted the effects of heroin for a duration of up to 72 hours and that it exhibited no signs of toxicity at doses of up to 200 mg per day. Research on naltrexone in opiate-naive healthy individuals suggested that the substance may possess certain opiate-like characteristics. Following the administration of 50 mg of naltrexone, participants reported experiencing drowsiness, dysphoria, sexual thoughts, penile erection, and an elevation in luteinizing hormone levels[6,7].

The literature suggests that the mortality risk associated with naltrexone is significantly increased in cases of relapse following opioid abstinence, primarily due to opioid toxicity. Patients undergoing naltrexone treatment experience a reduced tolerance to agonist opioids compared to their pre-treatment levels, making them susceptible to potentially fatal overdoses at the end of a dosing interval, after missing a dose, or upon discontinuation of treatment. Attempts to circumvent the opioid blockade can also result in fatal overdoses[8]. Additionally, the use of naltrexone in opioid-dependent individuals can trigger acute and severe withdrawal symptoms, characterized by heightened agitation compared to withdrawal from abstinence[9].

Diazepam has been commonly used for symptomatic relief, though case reports suggest variable effectiveness, often requiring high doses for sedation. One study reported initial limited relief with 10 mg intravenous diazepam, necessitating escalation to 60 mg for effective symptom control[10]. A comparative trial showed that midazolam had a faster onset of action (67 minutes) compared to diazepam (81 minutes), though neither was deemed ideal for rapid agitation management[11]. Furthermore, a review of primary studies reported that the doses of diazepam administered to treat agitation induced by naltrexone ranged from 5 mg to over 40 mg, depending the different naltrexone formulation and severity of symptoms[12].

There is limited empirical evidence supporting treatment recommendations, and consensus among experts is lacking. Traditional symptomatic therapies like antiemetics, clonidine,

benzodiazepines, and titrated doses of an opioid agonist are generally effective in managing opioid withdrawal symptoms[12]. Various medications have been utilized to address agitation induced by naltrexone in addicted patients, but none have demonstrated satisfactory efficacy in controlling agitation upon admission. Benzodiazepines are considered safe sedatives with documented efficacy in similar scenarios, although their use for managing agitation resulting from inappropriate naltrexone use is uncertain. Prompt management of agitation is crucial during severe episodes, as patients may pose a risk to themselves, companions, or medical staff[13]. Therefore, this study utilized a hybrid artificial neural network model to determine the optimal intravenous diazepam dosage for maximizing the efficacy of benzodiazepines in controlling naltrexone-induced agitation.

MATERIALS AND METHODS

Study Design and Setting

This study is a retrospective cross-sectional analysis of medical records of patients experiencing naltrexone-induced withdrawal symptoms at Loghman Hakim Hospital from April 2002 to March 2016. Trained clinical toxicologists documented patients' medical history, treatment trends, and vital signs. Figure 1 summarize the methodology of study. The study received approval from the Ethics Committee of Shahid Beheshti University of Medical Sciences (decision no: IR.SBMU.RETECH.REC.1402.626, date: 01.07.2024). Patient data were de-identified using file numbers to protect confidentiality.

Data Set Description and Participants

The dataset consists of 615 patient records from Loghman Hakim Hospital, focusing on individuals experiencing naltrexone-induced withdrawal symptoms. This hospital serves as a primary referral center for individuals affected by poisoning cases. Among the dataset entries, 232 cases were treated with high-dose diazepam (>10 mg) while 383 cases received low-dose diazepam (≤10 mg). The study included all patients presenting with naltrexone-induced withdrawal symptoms, whether due to intentional or accidental poisoning, at Loghman Hakim Hospital. The exclusion criteria consisted of cases with multiple drug toxicity, severe chronic comorbidities (e.g., cardiovascular diseases, neurological disorders, psychiatric conditions, or seizure disorders), as well as patients with incomplete medical records related to demographic information, vital signs, or paraclinical data at admission.

Benzodiazepines are considered safe sedatives with documented efficacy in similar scenarios, although their use for managing agitation resulting from inappropriate naltrexone use is uncertain[11].

Data Gathering

A comprehensive examination of patient medical records was conducted by a team of six researchers. Data were extracted from the electronic databases of Loghman Hakim Hospital (Sabara and Shafa databases) using a pre-made checklist. The collected information included demographic details such as age, gender, last opioid intake, and the purpose of naltrexone usage. Additionally, vital signs and withdrawal symptoms upon admission were documented. Furthermore, details on the administration of diazepam, blood glucose levels, electrocardiograms (ECGs), venous blood gases (VBG), blood electrolytes, liver and kidney function tests, and the Richmond Agitation-Sedation scale (RASS) were also recorded.

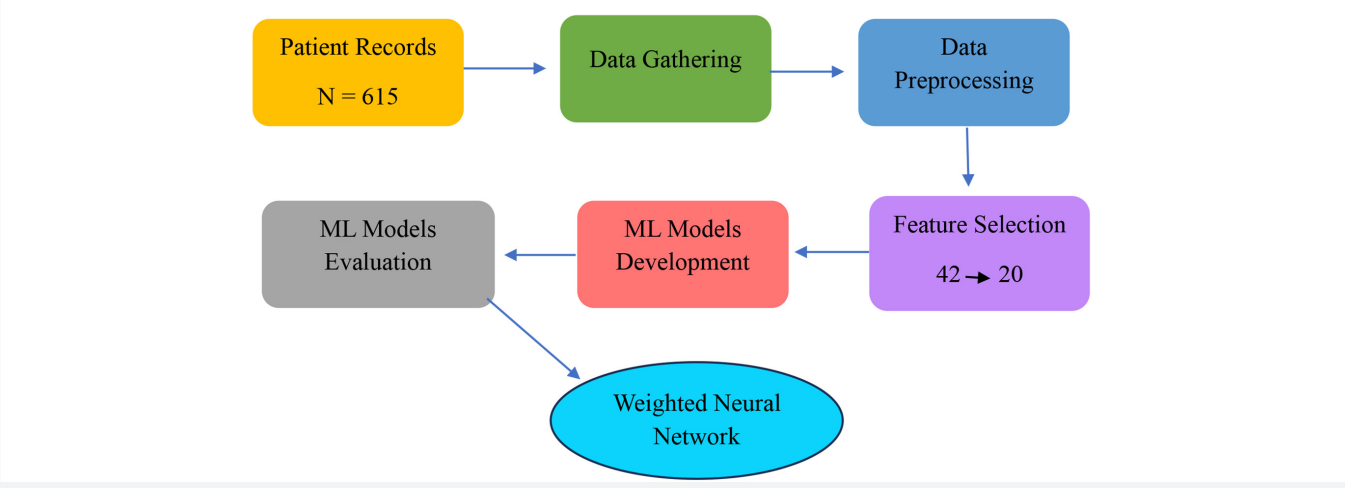


Figure 1. The flowchart visualizing the methodology of study

ML: Machine learning

Pre-processing of the Data

Various preprocessing techniques were used in this study to optimize classification algorithms after data collection. Methods included removing variables with over 70% missing data, identifying and eliminating noisy data and outliers, and addressing missing data in other variables with Mean Imputation and Stochastic Regression Imputation. To mitigate bias from mean imputation, we restricted its use to variables with minimal missingness (<5%). Data points falling outside the normal range were excluded in consultation with clinical committees. The dataset was split into train and test sets, then further divided into Class A for patients receiving diazepam doses of 10 mg or less, and Class B for doses over 10 mg.

Feature Selection

The process of feature selection, which involves eliminating redundant variables from the initial dataset while retaining essential information, is crucial in mitigating the risk of overfitting. In the initial phase of the research, a total of 42 features were identified, encompassing a diverse array of clinical and paraclinical symptoms, reflecting the high complexity and dimensionality of the characteristics under investigation. A crucial aspect of the study involved the selection of an optimal subset of features. To achieve this, the formulation of a loss function was imperative. The loss function was established through the utilization of the following equation in the research process:

$$\text{Loss Function} = \frac{1}{Accuracy}$$

To assess accuracy in the context of medical sciences for predicting treatment methods, where accuracy is crucial, the decision tree family of machine learning (ML) algorithms was employed. This family of algorithms is adept at addressing complex cases and offers valuable insights into feature importance for decision-making processes. Moreover, these algorithms demonstrate resilience when faced with imbalanced datasets. The evaluation of ML models from the decision tree family was conducted based on metrics such as accuracy, precision, sensitivity, area under the curve (AUC), and F1-score.

The decision tree algorithms utilize conditional statements to establish predictive criteria, with each algorithm possessing distinct architectural characteristics beyond this fundamental condition. The dataset was trained using a ten-fold cross-validation technique, involving the partitioning of the data into ten subsets and iteratively applying the holdout method. Meta-parameters were fine-tuned according to the training dataset through the utilization of the cross-validation methodology. The loss function was established based on the predictive accuracy of the model exhibiting the highest AUC metric.

The process of feature selection was conducted through the utilization of metaheuristic algorithms. The bat algorithm is a metaheuristic algorithm that operates on a population-based approach and is designed for addressing continuous optimization problems. This algorithm has demonstrated efficacy in optimizing solutions across various domains such as cloud computing, feature selection, image processing, and control engineering challenges[14].

When employing a search method based on 42 features, the computational complexity increases significantly, as the number of potential states grows exponentially at a rate of 42 squared. Furthermore, this search approach lacks convergence due to its random nature, in contrast to meta-heuristic algorithms which exhibit convergence and reduce the number of potential states as the search space is optimized.

The meta-heuristic algorithm has a time complexity of O (N^2), which is the same as a quadratic polynomial. In order to efficiently address the NP problem, we suggest using binary feature algorithms that have minimal time complexity and cost for organizing and distributing tasks in our feature selection issue. We outline an objective function and provide a table displaying the average duration for each iteration.

In this study, seven meta-heuristic algorithms were employed to select features based on the loss function including Binary Genetic Algorithm, Binary Particle Swarm Optimization, Binary Cuckoo Search, Binary Firefly Algorithm, Binary Bat Algorithm, Binary Gravitational Search Algorithm, and Binary Dragon Fly Algorithm. The meta-heuristic algorithm was chosen for its proven efficiency in medical feature selection[15,16]. The algorithm aims to minimize the loss function and systematically searches for a binary list. The binary list indicates feature selection, with a selection indicator represented by the number one and non-selection indicated by zero. Algorithm with the best score and the least loss function was utilized for feature selection[17].

Statistical Analysis

The Kolmogorov-Smirnov and the Shapiro-Wilk tests results revealed that all continuous variables were distributed non-normally. Consequently, the continuous variables were represented by their median values and interquartile ranges and were analyzed using the Mann-Whitney U test. Categorical variables were reported as absolute frequencies and respective percentage and were analyzed using the chi-square test. The performance of classification models was assessed through the receiver-operating curve. Additionally, other performance metrics including the accuracy, sensitivity, and specificity were computed. In this research, the Python Programming Language (version 13.1) and associated libraries were used. Libraries such as Matplotlib, NumPy, Seaborn, and Pandas were used for data analysis and visualization purposes. The scikit-learn library was employed to develop algorithms and evaluate ML models

performance. For descriptive analyses, the SPSS version 26 was utilized.

Model Evaluation (Stage 2)

After feature selection, five classifiers from the ML realms were employed to construct a predictive model for the appropriate diazepam dosage in patients experiencing naltrexone induced withdrawal. Among the ML models employed were the Light Gradient Boosting Machine, Random Forest Classifier, Gradient Boosting Classifier, Extreme Gradient Boosting and Extra Trees Classifier. The aim of utilizing this array of classifiers was to improve prediction accuracy and gain insights into the intricate factors influencing the optimal diazepam dosage for managing agitation induced by naltrexone. The dataset was divided randomly into training (70%) and testing (30%) sets to develop and validate the ML algorithms. A ten-fold cross-validation technique was applied to train the dataset with 20 selected features, involving the division of the dataset into ten sections and conducting the holdout method iteratively. Hyperparameters were adjusted based on the training dataset using the cross-validation approach. Subsequently, the classification algorithms were tested on the testing dataset to evaluate their performance. The performance of the classifiers in predicting appropriate dosage of diazepam in patients with naltrexone induced agitation was assessed using underfitting and overfitting evaluation methods, along with five standard efficiency testing metrics such as accuracy, specificity, sensitivity, precision, and F1-score according to the following equations:

1) $\text{classification accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} * 100$

2) $\text{classification sensitivity} = \frac{\text{TP}}{\text{TP+FN}} * 100$

3) $\text{classification specificity } = \frac{\text{TN}}{\text{TN+FP}} * 100$

4) $\text{classification percision } = \frac{\text{TP}}{\text{TP + FP}} * 100$

5) $\text{F1} - \text{Score} = \frac{2\times(\text{precision} \times \text{Sensitivity})}{\text{precision+Sensitivity}}$

The performance of each classifier was compared against other ML algorithms using these metrics. The best-performing model was selected based on the efficiency results to proceed with further data analysis and to determine the significance of features for neural network weighting through model tuning.

Feature Weight Calculation Using Decision Trees

In order to ascertain the importance coefficient of 20 selected features, the researchers enlisted the assistance of the most effective classifier based on the AUC metric in the realm of ML. Leveraging the unique characteristic of decision trees in calculating the Gini index, these trees were employed to evaluate feature importance. The efficacy of decision tree algorithms within this family is underscored by the root node, which encompasses all initial data pertinent to the issue at hand, in this instance, the 20 selected features. Subsequently, the attribute selection measure was utilized to identify the optimal features based on their level of importance. The feature that yields the most substantial decrease in impurity within a node is deemed the most valuable. Both Gini and Entropy methodologies can be applied to assess the impurity associated with each attribute. The research utilized the Gini index technique to assess the feature importance. This method involves favoring and choosing features with a lower Gini index over those with a higher Gini index in the decision tree. The Gini index is determined through the following mathematical formula:

$$G(t) = 1 - \sum_{i}^{c} p_i^2$$

The Gini impurity at a given node "t" is denoted as "G (t)", where "pi" represents the proportion of observations belonging to Class C at node "t". The Gini index is determined by subtracting the sum of the squared probabilities of each class from one. The information pertaining to the 20 selected features underwent initial processing through MinMax Scaler. This technique involves scaling the data, ensuring that the minimum feature is set to zero and the maximum feature is set to one. Notably, this approach maintains the original distribution shape of the data.

$$FX_i' = \frac{x_i - x_{min}}{x_{man} - x_{min}}$$

Next, the features intended for incorporation into the neural network were assigned weights based on the subsequent formula:

$$FX_i^{new} = FX_i' * w_i^{\alpha}$$

In the aforementioned relationship, "w" denotes the weight assigned to individual features derived from the decision tree. The parameter α is indicative of the hyperparameter utilized in this context to ascertain the impact of feature weights prior to their integration into the neural network. For the purposes of this research, α was set at a value of one. Finally weighted features were integrated in a multilayer perceptron (MLP) neural network. In this study, we used three layers including the input and output layers and the hidden layer. The weighted neural network's performance was evaluated against the ML models exhibiting the most accurate predictive capabilities[17].

RESULTS

Patient's Characteristics

The medical documentation of 907 individuals presenting with naltrexone toxicity was examined, with 292 patients being excluded based on predetermined exclusion criteria. The patient

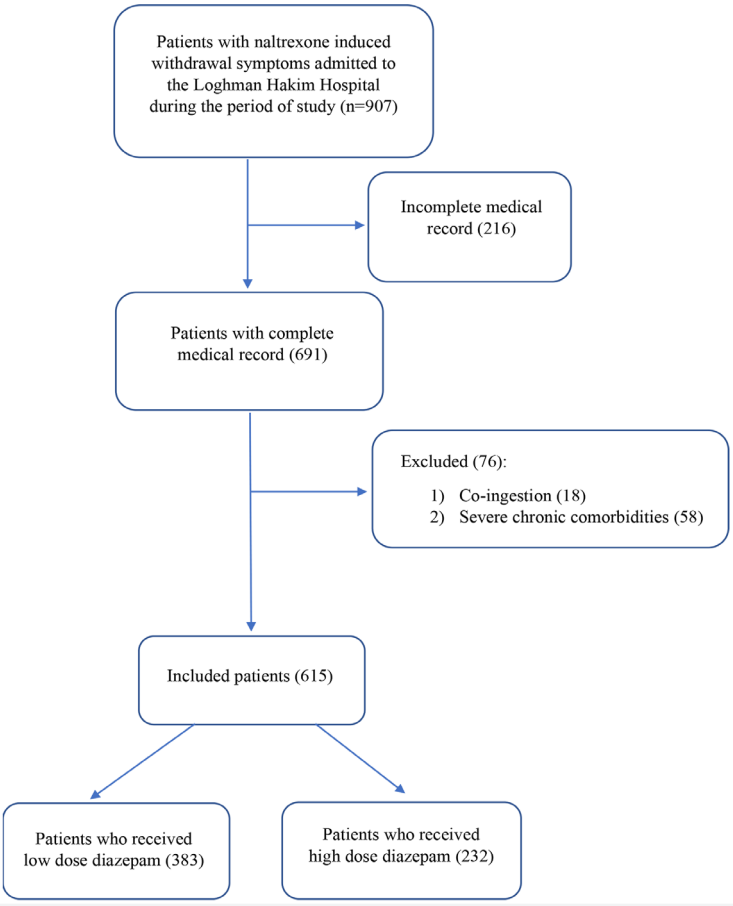


Figure 2. The patient selection flowchart

selection methodology is depicted in Figure 2 for clarity and transparency. The remaining cohort of 615 patients exhibited symptoms of naltrexone-induced agitation, with a mean age of 37.27 years and a standard deviation of 11.52 years, falling within the age range of 14 to 70 years. The study comprised 589 male participants with an average age of 37.27±11.50 and 26 female participants with an average age of 37.23±12.27. Statistical analysis revealed no significant difference in average age between the sexes (p=0.149). Among the 615 cases of poisoning examined, agitation in 232 cases (37.7%) was managed using a high dose of diazepam (exceeding 10 mg), while agitation in 383 cases (62.3%) was controlled with a low dose of diazepam (equal to 10 mg). The descriptive and analytical statistical outcomes pertaining to these two dosage categories are presented in Table 1.

Feature Selection

The initial step in feature selection involved the determination of a loss function to identify the most suitable features. This function was established by evaluating the performance of decision tree models using 42 initial features within the initial phase of model evaluation. Performance metrics of ML algorithms during this stage are detailed in Table 2.

The findings indicated that the Extra Trees Classifier model, exhibiting the highest AUC, outperformed other models. Consequently, the loss function was defined based on the accuracy of this particular model. Meta-heuristic algorithms were employed for feature selection based on the Loss function. These algorithms are designed to minimize the loss function and systematically search for a binary list with the lowest value. The results of using a list of the most important meta-heuristic algorithms are shown in supplementary information (Table 1). Furthermore, the evaluation of four meta-heuristic algorithms based on their performance revealed that the bat algorithm outperformed the others by achieving the highest score of 89.5% (0.895) in iteration 128 (Figure 3). The population data is refreshed, and the primary iteration concludes, repeating until the specified termination criterion is satisfied. In this particular scenario, the stopping criterion was defined as reaching a total of 200 iterations. The bat algorithm identified a total of 20 features for selection. The feature selection process is illustrated in supplementary information (Figure 1).

Table 1. Patient's characteristics

Variables	Scale	Class dose diazepam				
			Patients who received >10 mg diazepam (232)	Patients who received >10 mg diazepam (383)	Total (615)	p-value
Gender (n, %)	Nominal	Male	224 (38%)	365 (62%)	589	0.455
		Female	8 (31%)	18 (69%)	26	
Age (median, IQR)	Interval	------	34 (18)	37 (16)		0.149
Purpose of naltrexone use	Nominal	Opioid addiction treatment	3 (16%)	15 (84%)	18	0.019
		Suicide	52 (48%)	57 (52%)	109	
		Accident	5 (23%)	16 (77%)	21	
		Unknown	172 (36%)	295 (62%)	476	
Temperature	Interval	----	37 (0.1)	36.9 (0.1)	----	0.14
Systolic blood pressure	Interval	----	120 (20)	120 (15)	----	0.742
Diastolic blood pressure	Interval	----	75 (10)	75 (10)	----	0.902
$SpaO_2$	Interval	----	97 (1)	97 (0)	----	0.582
Vomiting	Nominal	Yes	35 (30%)	80 (70%)	115	0.07
		No	197 (39%)	303 (61%)	500	
Nausea	Nominal	Yes	37 (34%)	70 (66%)	107	0.46
		No	195 (38%)	313 (62%)	508	
Diarrhea	Nominal	Yes	19 (37%)	32 (63%)	51	0.94
		No	213 (38%)	351 (62%)	564	
Yawning	Nominal	Yes	5 (38%)	8 (625)	13	0.95
		No	277 (46%)	375 (54%)	602	
Lacrimation	Nominal	Yes	2 (18%)	9 (82%)	11	0.17
		No	230 (38%)	374 (62%)	604	
Respitatory	Nominal	Yes	3 (27%)	8 (73%)	11	0.47
		No	229 (38%)	375 (62%)	604	
ECG_Rate	Interval	---	80 (15)	80 (8)	---	0.95
Rhythm	Nominal	Normal sinus	182 (36%)	317 (64%)	499	0.324
		Not-sinus-AF	23 (44%)	29 (56%)	52	
		AF	24 (44%)	31 (56%)	55	
		Sinus tachycardia	1 (25%)	3 (75%)	4	
		Sinus bradycardia	2 (40%)	3 (60%)	5	
Axis deviation	Nominal	Normal	213 (38%)	348 (62%)	561	0.445
		Right axis deviation	12 (43%)	16 (57%)	28	
		Left axis deviation	7 (27%)	19 (73%)	26	
ST elevation	Nominal	Yes	2 (33%)	4 (67%)	6	0.82
		No	230 (38%)	379 (62%)	609	
T invert	Nominal	Yes	5 (28%)	13 (72%)	18	0.377
		No	227 (46%)	370 (54%)	597	
T flat	Nominal	Yes	5 (55%)	4 (45%)	9	0.266
		No	227 (37%)	379 (63%)	606	
VBG_PH	Interval	----	7.41 (0)	7.41 (0)	----	0.34
VBG_Pco_2	Interval	----	40 (0)	40 (0)	----	0.36
VBG_BE	Interval	----	0.7 (0)	0.7 (0)	----	0.05
p	Interval	----	80 (13)	80 (12)	----	0.54
K	Interval	----	4.1 (0.4)	4.1 (0.5)	----	0.722

Table 1. Continued

Variables	Scale	Class dose diazepam				
			Patients who received >10 mg diazepam (232)	Patients who received <10 mg diazepam (383)	Total (615)	p-value
BS	Interval	----	107 (24)	108 (23)	----	0.44
VBG_Po_2	Interval	----	38.4 (0)	38.4 (0)	----	0.629
VBG_Hco_3	Interval	----	24.5 (0)	24.5 (0)	----	0.145
Hgb	Interval	----	13.9 (1.45)	13.9 (1.3)	----	0.756
Na	Interval	----	141 (4)	141 (4)	----	0.759
Cr	Interval	----	1.04 (0.2)	1.04 (0.2)	----	0.691
AIT	Interval	----	38 (0)	38 (0)	----	0.004
BUN	Interval	----	30 (13)	30 (10)	----	0.311
AST	Interval	----	34 (0)	34 (0)	----	0.075
ALK	Interval	----	214 (0)	214 (0)	----	0.654
CK	Interval	----	666 (0)	666 (0)	----	0.064
LDH	Interval	----	662 (0)	662 (0)	----	0.192
Seizure	Nominal	Yes	2 (13%)	13 (87%)	15	0.04
		No	230 (38%)	370 (62%)	600	
Recent opioid use, <1 week, (n, %)	Nominal	Yes	148 (91%)	13 (9%)	161	0.000
		No	84 (18%)	370 (82%)	454	
Time elapsed before hospital admission	Interval	----	2 (4)	3 (3)		0.000
Naltrexone intake quantity (mg)	Interval	----	50 (50)	50 (0)		0.000
Level of consciousness	Ordinal	Conscious	183 (38%)	297 (62%)	480	0.86
		Grade 1	40 (37%)	69 (73%)	109	
		Grade 2	7 (37%)	12 (73%)	19	
		Grade 3	2 (40%)	3 (60%)	5	
		Grade 4	0	2 (100%)	2	
RASS	Ordinal	---	2 (4)	0 (2)	---	0.00

IQR: Interquartile range, ECG: Electrocardiogram, ST: Segment, AF: Atrial fibrillation, K: Potassium, BS: Blood sugar, Hgb: Hemoglobin, Na: Sodium, Cr: Creatinine, AIT: Autoimmune thyroiditis, BUN: Blood urea nitrogen, AST: Aspartate aminotransferase, ALK: Alkaline phosphatase, CK: Creatine kinase, LDH: Lactate dehydrogenase, RASS: Richmond agitation-sedation scale

Table 2. Ten-fold cross-validation for classifiers performance on selected predictors in classification decision tree family (42 features) in train datasets

Model		Accuracy	AUC	Sensitivity	Precision	F1
Light gradient boosting machine	Train	0.9186	0.9290	0.9186	0.9225	0.9169
	Test	0.8919	0.9210	0.8919	0.8919	0.8919
Random Forest Classifier	Train	0.9047	0.9241	0.9047	0.9079	0.9031
	Test	0.8757	0.9201	0.8757	0.8761	0.8758
Gradient Boosting Classifier	Train	0.9023	0.9241	0.9023	0.9054	0.9006
	Test	0.8865	0.9098	0.8865	0.8862	0.8863
Extreme Gradient Boosting	Train	0.9000	0.9252	0.9000	0.9030	0.8983
	Test	0.8919	0.9241	0.8919	0.8919	0.8919
Extra Trees Classifier	Train	0.8977	0.9226	0.8977	0.9017	0.8955
	Test	0.8703	**0.9306***	0.8703	0.8695	0.8695

AUC: Area under the curve, *: Maximum value

Model Evaluation (Stage 2)

The outcomes of ten-fold cross-validation for the performance of classifiers on a set of 20 features in the classification decision tree family are presented in Table 3. The analysis indicates that Extreme Gradient Boosting (XGBoost) achieved the highest levels of accuracy (0.8811), sensitivity (0.8811), precision (0.8819), and F1-score (0.8814) on the test dataset. Furthermore, the Extra Trees Classifier exhibited the highest AUC compared to other decision tree classifiers (AUC: 0.9343). Based on the findings in Table 3, the Extra Trees Classifier model emerged as the most suitable choice for further investigation into feature importance for weighting neural network performance through model tuning. The observed discrepancy between values of model's performance metrics was due to feature selection. Models trained on a reduced set of 20 selected features demonstrated superior performance compared to models using the original 42 features, highlighting the impact of feature selection on predictive accuracy.

Feature Weight Calculation Using Decision Trees

To determine the importance coefficient of 20 chosen features, researchers utilized the assistance of the Extra Trees Classifier, selected based on the optimal AUC metric obtained during model evaluation. The top 20 features and their respective importance values are detailed in Figure 4. The figure illustrates the ranking of feature importance in a descending order on the y-axis, with the x-axis representing the corresponding importance values. The researchers identified and ranked the most significant features in the following sequence: recent opioid use, RASS score, naltrexone intake quantity, heart rate, systolic blood pressure, time elapsed before hospital admission, sodium level, creatinine level, level of consciousness, VBG-HCO_3, purpose of naltrexone use, nausea, lactate dehydrogenase, AXIS, $SpaO_2$, seizure occurrence, gender, T-flat wave, respiratory rate, and lacrimation (Figure 4).

Neural Network Results

The outcomes of incorporating weighted features into a MLP neural network were detailed in Table 4. The MLP model demonstrated notable performance metrics, including an AUC of 0.94, accuracy of 0.91, precision of 0.92, sensitivity of 0.9, and an F1-score of 0.91. Additionally, Figures 5 to 7 visually depict the enhanced effectiveness of the weighted MLP neural network when compared with ML techniques like Extreme Gradient Boosting and Extra Trees Classifier. The values reported in the Tables 3 and 4 report the average values of evaluation metrics calculated for the two study groups. The minor discrepancies between the values shown in Figures 6 and 7 and those in Tables 3 and 4 is because the figures report numbers rounded to two decimal places.

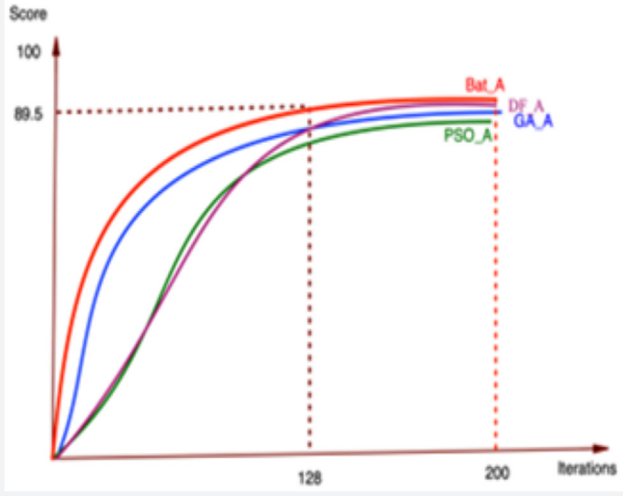


Figure 3. Performance of four meta-heuristic algorithms

Table 3. Ten-fold cross-validation for classifiers performance on selected predictors in classification decision tree family (20 features) in train datasets

Model		Accuracy	AUC	Sensitivity	Precision	F1
Gradient Boosting Classifier	Train	0.9256	-	0.9256	0.9303	0.9237
	Test	0.8703	0.9227	0.8703	0.8697	0.8699
Extreme Gradient Boosting	Train	0.9209	-	0.9209	0.9254	0.9201
	Test	**0.8811***	0.9175	**0.8811***	**0.8819***	**0.8814***
Light Gradient Boosting Machine	Train	0.9140	-	0.9140	0.9148	0.9129
	Test	0.8703	0.9252	0.8703	0.8703	0.8703
Random Forest Classifier	Train	0.9116	-	0.9116	0.9128	0.9107
	Test	0.8703	0.9302	0.8703	0.8697	0.8699
Extra Trees Classifier	Train	0.9070	-	0.9070	0.9093	0.9056
	Test	0.8649	**0.9343***	0.8649	0.8645	0.8649

*: Maximum value, AUC: Area under the curve

DISCUSSION

This research employed metaheuristic-based algorithms for the purpose of feature selection. Metaheuristic algorithms are a type of optimization methods used to tackle complex optimization problems that traditional approaches may have difficulty solving effectively. These algorithms are known for their adaptability in various optimization fields, including engineering, logistics, finance, and artificial intelligence[18]. Metaheuristic algorithms excel at identifying the most suitable subset of features in a dataset while maintaining model accuracy. Considering their effectiveness, this research focuses on leveraging metaheuristic algorithms to address feature selection complexities[19]. These algorithms efficiently solve complex optimization problems. Their significance is particularly notable in complex medical scenarios, such as diagnosis and treatment, especially in domains like drug dosage determination where predictive variables may be scarce. The utilization of metaheuristic algorithms to address these optimization challenges has emerged as a promising approach for handling NP-hard problems.

Modern challenges require quick solutions, making classical approaches inadequate. This has led to the rise of meta-heuristic algorithms, which explore spaces efficiently using a single fitness function, often with swarm intelligence. These algorithms can be population-based, like Genetic Algorithm, or path-based, like bat algorithm, which excels in complex biomedical scenarios. In this study, the bat algorithm outperformed other methods in feature selection, showing promise for solving challenging optimization problems in various fields[20,21].

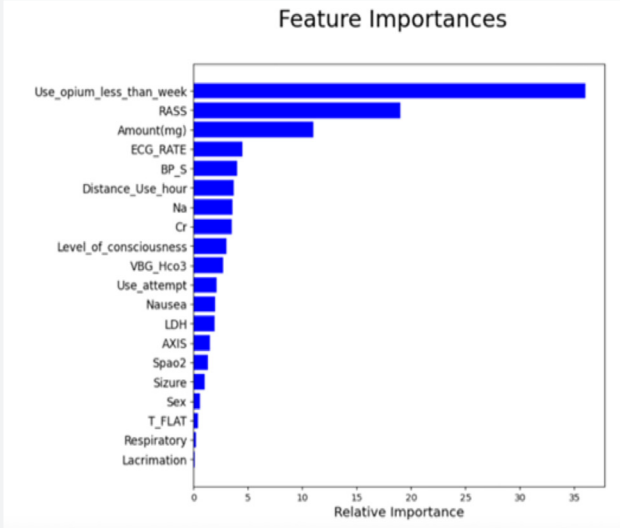


Figure 4. Feature importance

ECG: Electrocardiogram, Na: Sodium, Cr: Creatinine, LDH: Lactate dehydrogenase, RASS: Richmond Agitation-sedation scale, VBG: Venous blood gases

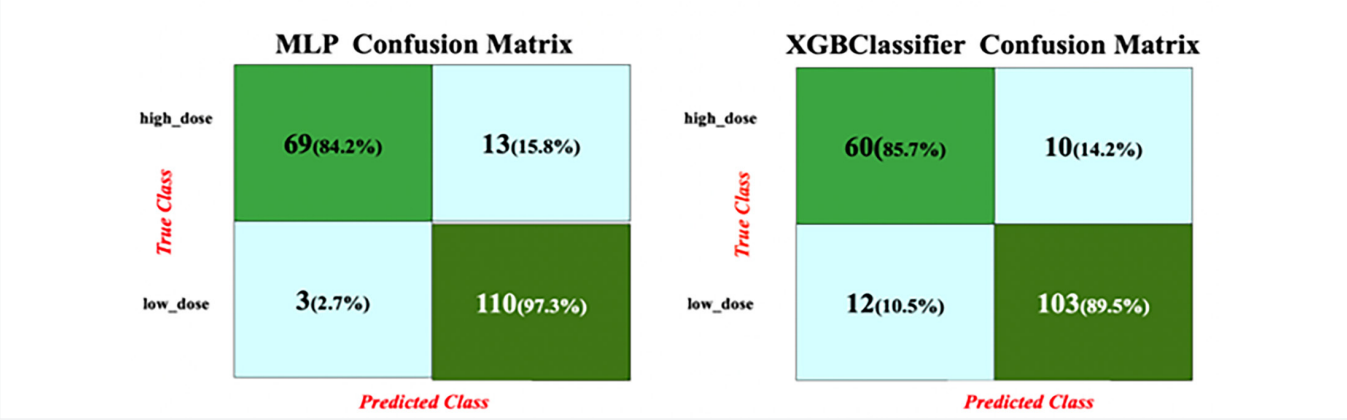


Figure 5. Confusion matrix comparison between the optimal machine learning classifier model, XGBoost, and the weighted MLP neural network model

MLP: Multilayer perceptron

Table 4. The outcomes of incorporating weighted features into a multilayer perceptron neural network

Model	Dataset	Accuracy	AUC	Sensitivity	Precision	F1
MLP	Train	**0.93**	-----	**0.92**	**0.94**	**0.93**
	Test	**0.91**	**0.94**	**0.90**	**0.92**	**0.91**

AUC: Area under the curve, MLP: Multilayer perceptron

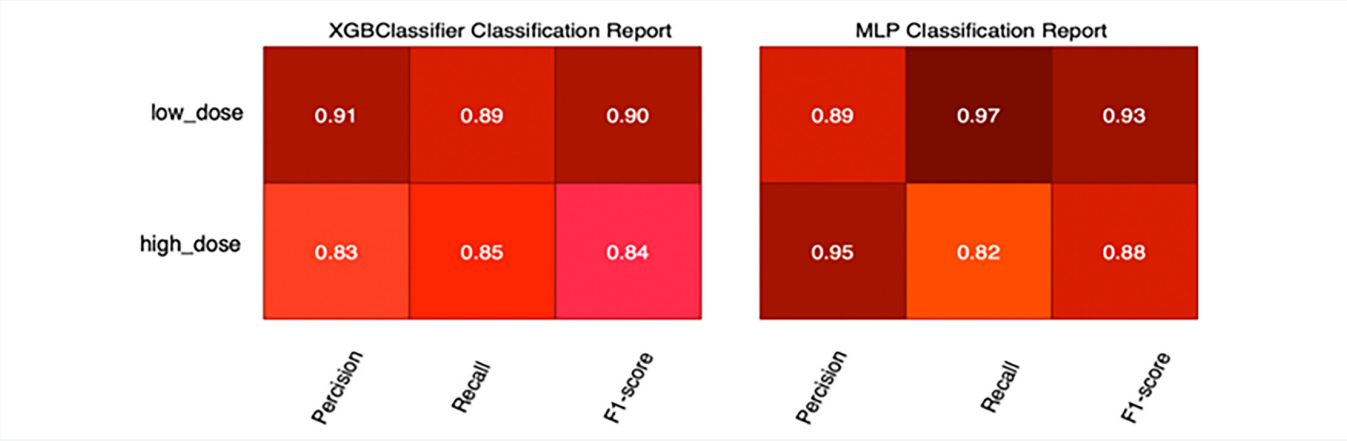


Figure 6. Comparison of the precision, sensitivity, and F1-score metrics in the optimal machine learning model (XGBoost) and weighted MLP neural network model

MLP: Multilayer perceptron

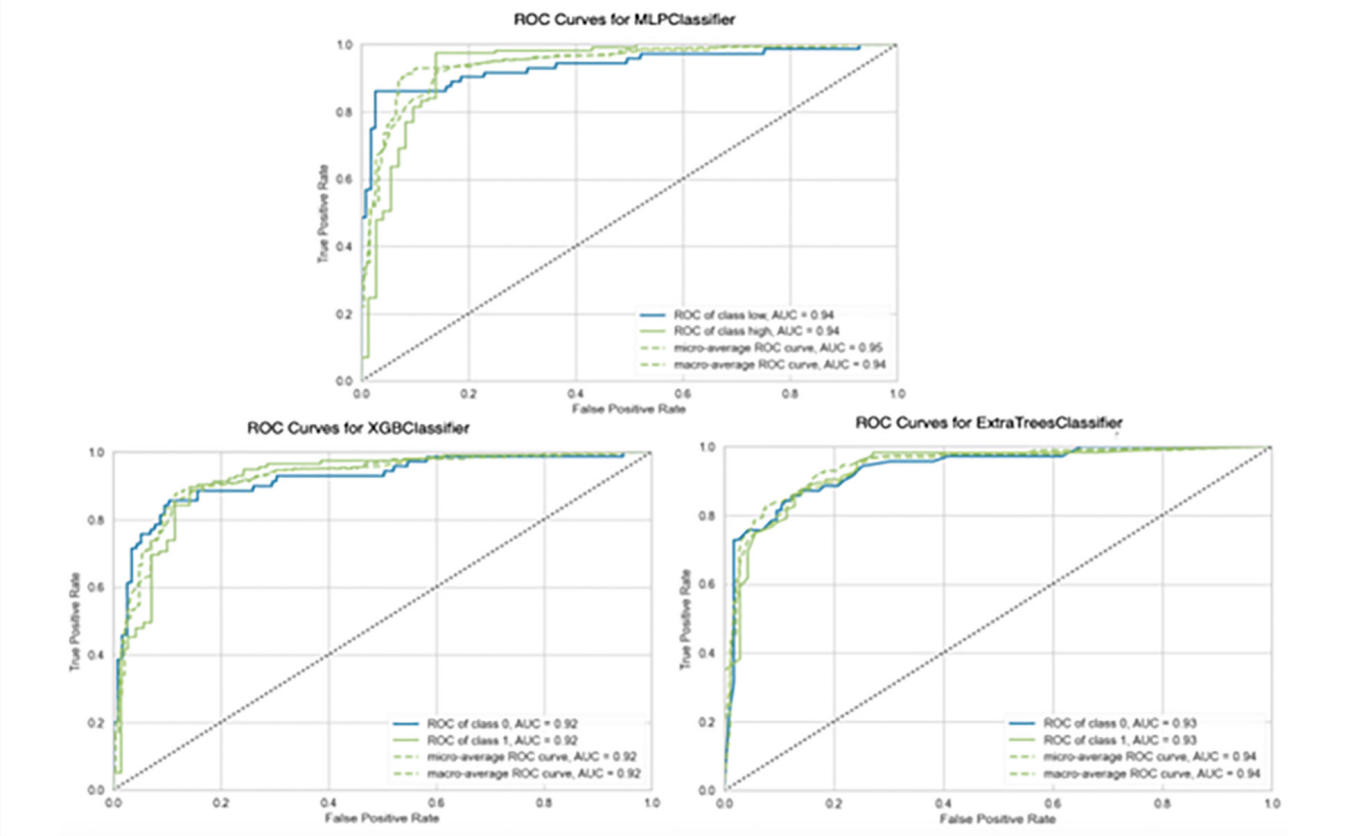


Figure 7. Comparison of the ROC curve between two machine learning models (XGBoost and Extra Trees classifier) and weighted MLP neural network model

MLP: Multilayer perceptron, AUC: Area under the curve

Among developed ML models, the Extra Trees Classifier yield the strongest performance based on its AUC score. The Extra Trees algorithm excels at selecting informative features via gradient-based techniques, enhancing predictive modeling outcomes. For example, Mathpal et al.[22] reported how Extra Trees played a central role in identifying compounds targeting mutant PBP4 in *Staphylococcus aureus*, achieving an 81% predictive accuracy. This highlights the model's effectiveness in not only making predictions but also elucidating which features contribute most significantly, thus aiding in clinical decision-making and research explorations. Numerous studies demonstrate that predictions using Extra Trees Classifiers have been successfully applied across various critical healthcare settings, from early detection of diabetes[23] to the prediction of abdominal aortic aneurysms[24]. The ability to predict outcomes reliably emphasizes the model's critical role in improving patient management and therapeutic strategies. A study compared decision tree classifiers with a weighted MLP neural network, finding that the MLP network showed superior performance in AUC and other evaluation criteria. Despite typically requiring large training samples, MLP models can be accurate with limited data by restricting input parameters and utilizing feature weight calculation from the highest AUC classifier. The ROC curve, unaffected by imbalanced data, makes AUC a preferred metric for evaluating ML models[25,26].

Recent opioid use was found to be the primary factor influencing the necessary dosage of diazepam to manage agitation triggered by naltrexone. The absorption of naltrexone through oral intake occurs swiftly, reaching peak levels in the bloodstream within 3 hours. It is advised that individuals be free of opioids for 7 to 10 days before receiving naltrexone. Consequently, patients who have used opioids within a week prior to naltrexone administration are at a higher risk of encountering acute opioid receptor blockade and intense opioid withdrawal symptoms, necessitating increased diazepam dosages[9]. A higher RASS score was found to be a strong indicator of a high dosage of diazepam, consistent with prior research that has demonstrated a relationship between RASS score and the amount of sedative and analgesic drugs administered[27,28]. A higher RASS score may signal a need for >10 mg diazepam to achieve sedation. However, the exact RASS cut-off value needs to be determined in future studies. Increased consumption of naltrexone was linked to an increased need for diazepam to alleviate agitation, likely attributable to its competitive antagonistic properties and the dose-dependent manner in which it blocks opioid receptors[29]. Although, in the study conducted by Sabzghabaee et al.[11] it has been reported that diazepam cannot effectively reduce agitation until 120 minutes post-administration, and its mean onset of action is lower than midazolam. In a systematic review by Kunzler et al.[12], reported diazepam dosing regimens varied by treatment scenario: 5-10 mg for oral naltrexone, 10 mg for extended-release injectable naltrexone (Vivitrol®), up to 30 mg for naltrexone implants or nalmefene (18 mg), and 10-40 mg for high-dose naltrexone (50 mg).

Several factors such as elevated pulse rate, increased systolic blood pressure, heightened level of consciousness, presence of nausea, reduced lacrimation, and decreased need for mechanical intubation are more prevalent in severely agitated patients, potentially necessitating higher doses of diazepam for resolution. Gender was identified as a feature; further research is needed to explore the potential relationship between gender, agitation intensity, and diazepam dosage requirements. It is important to note that suicide is a complex issue with multiple contributing factors. Additional predictors warranting further investigation for their potential association with increased diazepam dosages in these patients include ECG abnormalities (such as T-wave flattening and axis deviation), levels of sodium and creatinine, arterial oxygen saturation, and bicarbonate levels in VBG analysis.

Study Limitations

There are several limitations that need to be acknowledged. Primarily, the study was constrained by the challenges associated with data collection from multiple medical facilities in Iran, resulting in the utilization of data from a single hospital. As regional differences in opioid use patterns may affect generalizability, future investigations should consider expanding the sample size or incorporating data from multi-center datasets. Furthermore, the study was restricted to a limited selection of five ML models. To gain a more comprehensive understanding, it is advisable for subsequent research to explore a broader array of models. Moreover, the method to handle the missing data may introduce bias. For instance, mean imputation, may underestimate variability for variables with non-random missingness and stochastic regression imputation relies on correctly specified regression models. Future work could explore advanced methods (e.g., Bayesian imputation) for complex missingness patterns. Lastly, the retrospective nature of the dataset may introduce biases and uncertainties stemming from missing data.

CONCLUSION

Our research findings indicate that the utilization of a weighted MLP neural network can be an effective tool in developing a prediction model for the necessity of higher doses of diazepam for patients experiencing naltrexone-induced agitation. Particularly noteworthy is the enhanced predictive capability achieved when employing meta-heuristic techniques for feature selection and subsequently weighting these selected features using a classifier with the highest AUC value. In conclusion, our study emphasizes the value

of employing a weighted MLP neural network to improve predictive accuracy and facilitate the tailored management of patients experiencing naltrexone-induced agitation and could reduce trial-and-error dosing, improving patient safety and staff efficiency.

Ethics

Ethics Committee Approval: The study received approval from the Ethics Committee of Shahid Beheshti University of Medical Sciences (decision no: IR.SBMU.RETECH.REC.1402.626, date: 01.07.2024).

Informed Consent: This study is a retrospective cross-sectional analysis of medical records of patients experiencing naltrexone-induced withdrawal symptoms at Loghman Hakim Hospital from April 2002 to March 2016.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Goel A, Kapoor B, Wu M, Iyayi M, Englesakis M, Kohan L, et al. Perioperative naltrexone management: a scoping review by the perioperative pain and addiction interdisciplinary Network. Anesthesiology. 2024;141:388-99.
2. Zangiabadian M, Golmohammadi S, Nejadghaderi SA, Zahmatkesh MM, Nasiri MJ, Sadeghian M. The effects of naltrexone on retention in treatment and being opioid-free in opioid-dependent people: a systematic review and meta-analysis. Front Psychiatry. 2022;13:1003257.
3. Gibson A, Degenhardt L. Mortality related to naltrexone in the treatment of opioid dependence: a comparative analysis NDARC Technical Report No. 229.
4. Rahimi M, Kargar A, Hazegh Fetratjoo D, Hosseini SM, Mahdavinejad A, Shadnia S. Demographic and clinical characteristics of 907 cases with naltrexone intoxication; a 14-year cross-sectional study. Arch Acad Emerg Med. 2022;10:e34.
5. Saadabadi DSA. Naltrexone StatPearls Treasure Island (FL): StatPearls Publishing. Available from: https://www.ncbi.nlm.nih.gov/books/NBK534811/.
6. Erlich Y. Modulators, mediators, and specifiers in brain function: interactions of neuropeptides, cyclic nucleotides, and phosphoproteins in mechanisms underlying neuronal activity, behavior, and neuropsychiatric disorders. Springer Science & Business Media. 2012.
7. Volavka J, Mallya A, Bauman J, Pevnick J, Cho D, Reker D, et al. Hormonal and other effects of naltrexone in normal men. Adv Exp Med Biol. 1979;116:291-305.
8. Darke S, Farrell M, Duflou J, Larance B, Lappin J. Circumstances of death of opioid users being treated with naltrexone. Addiction. 2019;114:2000-7.
9. Hassanian-Moghaddam H, Afzali S, Pooya A. Withdrawal syndrome caused by naltrexone in opioid abusers. Hum Exp Toxicol. 2014;33:561-7.
10. Aggarwal K, Ranjan R, Mallik S, Salian HH, Shekhar S. Delirium secondary to inadvertent administration of naltrexone in patient with opioid dependence syndrome: a case report. East Asian Arch Psychiatry. 2024;34:87-8.
11. Sabzghabaee AM, Eizadi-Mood N, Gheshlaghi F, Javani A, Shirani S, Aghaabdollahian S. Role of benzodiazepines in the management of agitation due to inappropriate use of naltrexone. Iran J Nurs Midwifery Res. 2012;17:365-9.
12. Kunzler NM, Wightman RS, Nelson LS. Opioid withdrawal precipitated by long-acting antagonists. J Emerg Med. 2020;58:245-53.
13. Rund DA, Ewing JD, Mitzel K, Votolato N. The use of intramuscular benzodiazepines and antipsychotic agents in the treatment of acute agitation or violence in the emergency department. J Emerg Med. 2006;31:317-24.
14. Abdel-Basset M, Abdel-Fatah L, Sangaiah AK. Metaheuristic algorithms: a comprehensive review. Computational intelligence for multimedia big data on the cloud with engineering applications. 2018:185-231.
15. Singh J, Sandhu JK, Kumar Y. Metaheuristic-based hyperparameter optimization for multi-disease detection and diagnosis in machine learning. Service Oriented Computing and Applications. 2024;18:163-82.
16. Parizi MK. Binary modified cat and mouse based optimizer for medical feature selection: a COVID-19 case study. 2024.
17. Mohtarami SA, Mostafazadeh B, Shadnia S, Rahimi M, Evini PET, Ramezani M, et al. Prediction of naloxone dose in opioids toxicity based on machine learning techniques (artificial intelligence). Daru. 2024;32:495-513.
18. Watgule U. Metaheuristics Algorithms 2023. Available from: https://medium.com/@ujwalwatgule/metaheuristics-algorithms-889d2b0f9652.
19. Akinola OO, Ezugwu AE, Agushaka JO, Zitar RA, Abualigah L. Multiclass feature selection with metaheuristic optimization algorithms: a review. Neural Comput Appl. 2022;34:19751-90.
20. Sadeghian Z, Akbari E, Nematzadeh H, Motameni H. A review of feature selection methods based on meta-heuristic algorithms. Journal of Experimental & Theoretical Artificial Intelligence. 2023:1-51.
21. Kar AK. Bio inspired computing - a review of algorithms and scope of applications. Expert Systems with Applications. 2016;59:20-32.
22. Mathpal S, Joshi T, Priyamvada P, Ramaiah S, Anbarasu A. Machine learning and cheminformatics-based Identification of lichen-derived compounds targeting mutant $PBP4^{R200L}$ in Staphylococcus aureus. Mol Divers. 2025;29:3345-70.
23. Arya M, Sastry G H, Motwani A, Kumar S, Zaguia A. A novel extra tree ensemble optimized DL framework (ETEODL) for early detection of diabetes. Front Public Health. 2022;9:797877.
24. Forneris A, Beddoes R, Benovoy M, Faris P, Moore RD, Di Martino ES. AI-powered assessment of biomarkers for growth prediction of abdominal aortic aneurysms. JVS Vasc Sci. 2023;4:100119.
25. Fujita T, Sato A, Narita A, Sone T, Iokawa K, Tsuchiya K, et al. Use of a multilayer perceptron to create a prediction model for dressing independence in a small sample at a single facility. J Phys Ther Sci. 2019;31:69-74.
26. Siciarz P, Alfaifi S, Uytven EV, Rathod S, Koul R, McCurdy B. Machine learning for dose-volume histogram based clinical decision-making support system in radiation therapy plans for brain tumors. Clin Transl Radiat Oncol. 2021;31:50-7.
27. Lee JA, Duby JJ, Cocanour CS. Effect of early and focused benzodiazepine therapy on length of stay in severe alcohol withdrawal syndrome. Clin Toxicol (Phila). 2019;57:624-7.
28. Ely EW, Truman B, Shintani A, Thomason JW, Wheeler AP, Gordon S, et al. Monitoring sedation status over time in ICU patients: reliability and validity of the Richmond Agitation-Sedation Scale (RASS). JAMA. 2003;289:2983-91.
29. Sirohi S, Dighe SV, Madia PA, Yoburn BC. The relative potency of inverse opioid agonists and a neutral opioid antagonist in precipitated withdrawal and antagonism of analgesia and toxicity. J Pharmacol Exp Ther. 2009;330:513-9.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.56933
Nam Kem Med J 2025;13(3):222-230

Evaluating the Quality of YouTube Educational Videos on Kangaroo Care for Newborns: An Observational Study

Yenidoğanlarda Kanguru Bakımı ile İlgili YouTube Eğitim Videolarının Kalitesinin Değerlendirilmesi: Gözlemsel Bir Çalışma

Mustafa Törehan ASLAN[1], Nedim SAMANCI[2]

[1]*Koç University Hospital, Department of Pediatrics, Division of Neonatology, İstanbul, Türkiye*
[2]*Tekirdağ Namık Kemal University Faculty of Medicine, Department of Pediatrics, Division of Neonatology and Pediatric Intensive Care, Tekirdağ, Türkiye*

ABSTRACT

Aim: Although YouTube is now a popular health education platform, the quality and veracity of its content is highly variable. Videos on kangaroo care (KC) for newborns, for example, are created both by medical professionals and non-professionals, eliciting the concern that the provided information might be not trustworthy and effective. This study aimed to evaluate the quality, engagement, and content reliability of YouTube videos on KC, comparing videos produced by medical professionals and non-professionals.

Materials and Methods: The study analyzed 100 eligible YouTube videos. Modified DISCERN (mDISCERN) instrument name and global quality scale (GQS) scores were used to assess video quality, and presenter types (e.g., medical professionals, parents, non-governmental organization representatives) and video features (video length, use of animation, etc.) were recorded. Views, likes, and comments were also examined for viewer engagement metrics.

Results: Video content uploaded by medical professionals had higher mDISCERN and GQS scores than non-professionals ($p<0.05$). Views and likes were similar in both groups. Non-professional videos were found to have similar viewer engagement although less detailed and shorter. The use of animations and graphics was found to increase the understandability of the videos, and approximately half of the videos were found to have no existing annotations.

Conclusion: The study found that videos created by medical professionals were more credible and effective at conveying accurate health information. Despite this, demand for content-rich videos in the healthcare industry was notable. Future efforts to improve video quality and digital literacy will be critical to increase the credibility of health-related content on YouTube.

Keywords: YouTube video quality, educational videos, kangaroo care, newborns

ÖZ

Amaç: YouTube artık popüler bir sağlık eğitimi platformu olmasına rağmen, içeriğinin kalitesi ve doğruluğu oldukça değişkendir. Örneğin, yenidoğanlar için kanguru bakımı (KB) videoları hem tıp uzmanları hem de profesyonel olmayanlar tarafından oluşturulmaktadır ve bu da sağlanan bilgilerin güvenilir ve etkili olmayabileceği endişesini uyandırmaktadır. Bu çalışma, tıp uzmanları ve profesyonel olmayanlar tarafından üretilen videoları karşılaştırarak KB hakkındaki YouTube videolarının kalitesini, etkileşimini ve içerik güvenilirliğini değerlendirmeyi amaçlamaktadır.

Gereç ve Yöntem: Çalışmada, uygun olan 100 YouTube videosu analiz edilmiştir. Modifiye DISCERN (mDISCERN) enstrüman adı ve küresel Kalite ölçeği (GQS) puanları video kalitesini değerlendirmek için kullanıldı ve sunucu tipleri (örneğin, tıp uzmanları, ebeveynler, sivil toplum kuruluşları temsilcileri) ve video özellikleri (video uzunluğu, animasyon kullanımı vb.) kaydedildi. Görüntülemeler, beğeniler ve yorumlar da izleyici etkileşimi metrikleri açısından incelendi.

Address for Correspondence: Mustafa Törehan ASLAN MD, Koç University Hospital, Department of Pediatrics, Division of Neonatology, İstanbul, Türkiye
E-mail: torehanaslan@yahoo.com **ORCID ID:** orcid.org/0000-0002-3966-4635
Received: 03.03.2025 **Accepted:** 07.05.2025 **Publication Date:** 07.10.2025

Cite this article as: Aslan MT, Samancı N. Evaluating the quality of Youtube educational videos on kangaroo care for newborns: an observational study. Nam Kem Med J. 2025;13(3):222-230

Bulgular: Tıbbi profesyoneller tarafından yüklenen video içerikleri, profesyonel olmayanlara göre daha yüksek mDISCERN ve GQS puanlarına sahipti ($p<0,05$). Videoların görüntülenme ve beğeni sayıları iki grupta benzerdi. Daha az ayrıntılı ve daha kısa süreli olmasına rağmen, profesyonel olmayan videoların da benzer izleyici etkileşimi olduğu saptandı. Animasyon ve grafiklerin kullanılmasının videoların anlaşılabilirliğini artırdığı ve videoların yaklaşık yarısında mevcut eklemelerin olmadığı bulundu.

Sonuç: Çalışma, tıp uzmanları tarafından oluşturulan videoların doğru sağlık bilgilerini iletmede daha güvenilir ve etkili olduğunu gösterdi. Buna rağmen sağlık sektöründe zengin içerikli videolara olan talep dikkat çekiciydi. YouTube'da olan sağlıkla ilgili içeriklerin güvenilirliğini artırmak için video kalitesini ve dijital okuryazarlığını iyileştirmeye yönelik gelecekteki çabalar kritik öneme sahip olacaktır.

Anahtar Kelimeler: YouTube video kalitesi, eğitim videoları, kanguru bakımı, yenidoğanlar

INTRODUCTION

The scientific practice of kangaroo care (KC) demonstrates strong evidence as an essential method to support newborns especially during neonatal and preterm health care. KC applies when an infant joins skin with their caregiver (typically the mother) for direct contact, enabling multiple advantages to develop in both players[1]. Medical literature provides significant documentation about KC, which shows it helps control temperature levels while enhancing breastfeeding success, lowering infant stress, and establishing premature parent-child bonds from the start[2].

The usage of online platforms, especially YouTube, has experienced rapid growth in spreading health and medical-related information during previous years. YouTube is the world's leading video-sharing platform because people access its educational content about KC an is a convenient platform[3]. YouTube videos allow information to reach many viewers who can obtain crucial guidance for parents, health practitioners, and caregivers about KC practices and advantages. YouTube videos emerge as suitable instructional tools for delivering important health practices such as KC since viewers typically find video content more engaging than written materials[4].

The simple accessibility to educational videos represents a major benefit, but it creates crucial doubts concerning the quality and reliability of the information provided. Each day, millions of uploaded videos make it harder for users to identify accurate, evidence-based, or credible information from the pool of available content[5]. The educational materials available on YouTube demonstrate wide-ranging quality differences since they depend on creators' professionalism, production skills, and adherence to best health communication methods[6]. This study aimed to assess the quality of educational videos on YouTube regarding KC because stakeholders must acquire correct and beneficial information that affects newborn health outcomes. This evaluation determines the accuracy of reliable, evidence-based data in the videos while examining their content clarity, source credibility, and success rate in displaying KC's benefits and implementation strategies.

Public health is greatly impacted by the quality standards established for educational videos concerning healthcare education, particularly in areas related to newborn care. Parents and caregivers use Internet-based information to understand newborn care methods, particularly when they need assistance with premature babies or other newborn medical issues[7]. The infant and the parent experience better well-being when they receive correct practical guidance during these circumstances. Faulty information on newborn care can cause detrimental outcomes for families, along with unwanted mental stress and inappropriate techniques that harm their well-being[8].

Better known as KC, the practice receives wide recommendations from pediatricians and other physicians for its beneficial effects on infant development. The practice of KC receives diverse levels of understanding from caretakers because some regions lack sufficient medical resources, and new parents rarely meet with physicians for education about medical practices[9]. YouTube provides simple access to its platform, which can help caregivers gain essential knowledge to deliver optimal care to their newborns. YouTube is a difficult platform for information-quality evaluation because consumers face overwhelming content[10]. Poorly presented educational videos presented as KC information materials could confuse viewers because wrong practices result in lost benefits from the intervention. The results from this investigation serve a vital function by validating the accuracy, effective, and helpfulness of YouTube educational material for its target viewers[11].

The research examines YouTube's educational quality for new parents and caregivers by analyzing one critical health topic. Evaluating KC education videos on YouTube will guide enhancements in online health communication protocols that specifically benefit new parents and preterm infant caregivers. The investigation can discover specific zones that teach content creators how to upgrade their educational YouTube videos, thus elevating the standard of knowledge on this platform. This research investigation possesses value to healthcare providers, healthcare establishments, and public health institutions utilizing YouTube for educational delivery.

Organizations that understand the current video content about KC will make better decisions for developing high-quality educational materials that effectively transmit health practices to the public[9]. The main purpose of this research project

was to examine the educational video quality about KC for newborns, which exists on the YouTube platform. The research examines video information accuracy to verify its compliance with modern research findings and healthcare standards. The research must evaluate the reliability of sources used in YouTube videos by assessing which ones use information from trusted healthcare experts, scholarly research, or respected medical organizations.

MATERIALS AND METHODS

Researchers explored the reliability and quality metrics of educational KC videos for newborns on the YouTube platform. The research methodology used in this study is concordant with the methods described in previous research that evaluated YouTube as a medical information source. This research reviewed 100 YouTube videos focusing on KC through systematic selection methods. The diagram of the analyzed videos was depicted in Figure 1. The research team evaluated these educational YouTube videos by examining their content accuracy, clear communication, and their creators' qualifications, video quality, and user engagement metrics. The quality of the searched videos was classified as follow: 144-240p: low quality, 360-480p: medium quality, 720p and higher: high quality (Figure 2).

YouTube searches from the keywords "KC for newborns" and its related words" "skin-to-skin care" and "newborn care" resulted in the selection of the analyzed videos. The selected videos appeared under the "relevance" default sorting setting to choose content that matched the search requirements. The research considerations included videos that were (1) in the English language, (2) dedicated to educational themes about KC and lasted between (3) 1 and 60 minutes. Videos between one minute and sixty minutes were analyzed while those shorter or longer than this time range were excluded to maintain only informative substantial video content. The analysis excluded duplicate files, promotional materials, and videos that lacked proper identification of the source or creator information. After applying the criteria, the team obtained 100 videos they selected for the study.

Features and Usability of the YouTube Videos

Several video characteristics were recorded systematically. A complete set of video data included information about its duration, content structure, information type demonstration style and expert involvement, and creator experience level and distribution platform statistics. Researchers sorted the videos into two creator categories: (1) professional creators who possessed healthcare backgrounds or belonged to medical institutions and (2) non-professional creators who were either non-medical individuals or belonged to non-medical organizations.

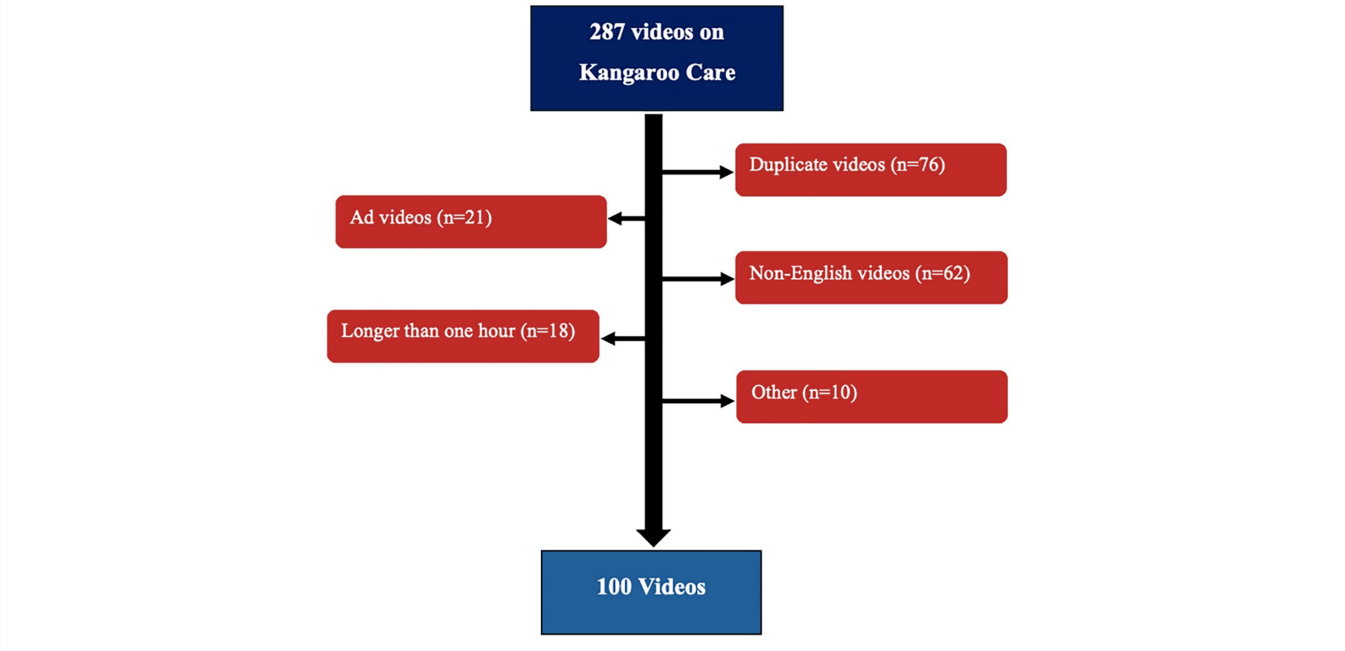


Figure 1. Diagram of the analyzed videos. The "Other" defines the videos that lacked proper identification of the source or creator information

Assessment of Quality-Reliability of Videos

The Global quality scale (GQS) (Table 1) and modified DISCERN (mDISCERN, instrument name) tool (Table 2) determined the reliability and quality of the produced videos[12]. The GQS functions as a five-point Likert scale to determine the entire quality level of online health information. The assessment tool checks several parameters that evaluate user interface accessibility, information organization structure, and content quality standards. The GQS uses five rating categories that begin at poor quality and end at excellent quality. Research participants evaluated the video's KC presentation effectiveness using the GQS to determine its usefulness for parents and caregivers.

Health-related content is assessed through the mDISCERN tool, which verifies the accuracy and reliability of information centered on treatment decisions. The research group transformed the evaluation instrument for use with YouTube health information but maintained its original framework to match this special online content. Each item in the mDISCERN tool is evaluated through a five-point Likert scale to determine if the video provides clear aims while using reliable sources to maintain balance and acknowledge uncertainties. Higher scores indicate greater reliability. The mDISCERN tool assessed the quality of information delivery through YouTube videos regarding KC. The analysis divided videos into categories using scores from mDISCERN and GQS instruments. Useful videos met high-quality standards, while insufficient content received the other classification. The two independent reviewers [GQS (1) and (2) - mDISCERN (1) and (2)] performed the classification process and settled conflicts by discussing the differences for mutual agreement on the categorization results.

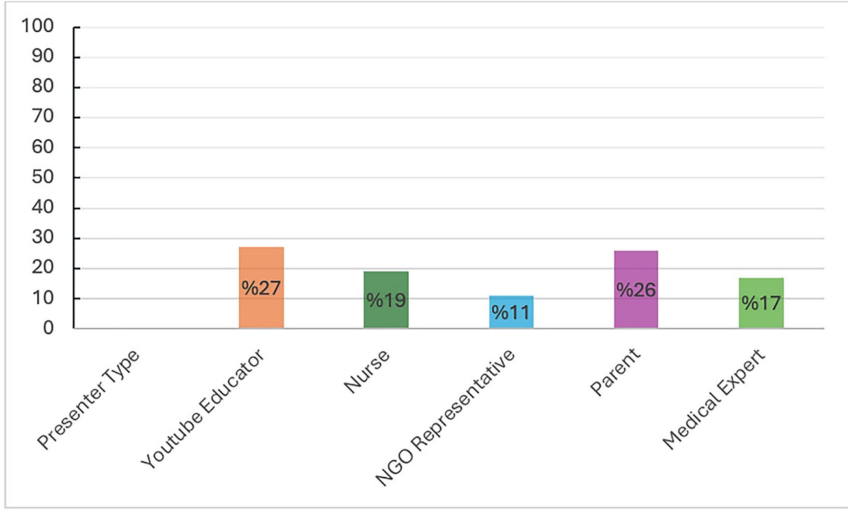


Figure 2. Video presenter classification

NGO: Non-governmental organization

Table 1. Global quality scale (GQS) items

#	Item
1	Poor quality, poor flow of the site, most information missing, not at all useful for patients
2	Generally poor quality and poor flow, some information listed but many important topics missing, of very limited use to patients
3	Moderate quality, suboptimal flow, some important information is adequately discussed but others poorly discussed, somewhat useful for patients
4	Good quality and generally good flow, most of the relevant information is listed, but some topics not covered, useful for patients
5	Excellent quality and excellent flow, very useful for patients

Table 2. Modified DISCERN (mDISCERN, instrument name) scale items

#	Item
1	Are the aims clear and achieved?
2	Are reliable sources of information used?
3	Is the information presented balanced and unbiased?
4	Are additional sources of information listed for patient reference?
5	Are areas of uncertainty mentioned?

Statistical Analysis

A statistical evaluation was conducted on the gathered data to find changes between videos produced by specialists and those from non-specialists. The researchers used descriptive statistics to compile data about video characteristics by presenting means and standard deviations for length, views, likes, and comments. The Gaussian distribution of the parametric data was analyzed with the Kolmogorov-Smirnov test. The evaluation of GQS and mDISCERN scores connecting professional and non-professional groups relied on independent t-tests to establish these differences for ongoing variables, including video duration and viewing activity metrics. Cronbach's alpha method measured the scorer agreement in video assessment, showing an excellent match when the value reached 0.8 or above. The researchers applied SPSS version 25.0 to execute all statistical methods at a p-value below 0.05.

Ethics approval for this study was not deemed necessary as it involved the analysis of publicly available data from YouTube videos, which do not involve human or animal participants. Since YouTube videos are publicly accessible and do not require permission from the content creators for viewing or analysis, no formal permission was sought from the platform. The study adhered to ethical guidelines by ensuring that all data collected were publicly available and did not involve personal or sensitive information.

RESULTS

Figure 3 shows the distribution of the various presenters in the analyzed videos. The largest number of presenters is "YouTube educator" (27%) while the second largest is "non-governmental organizations" (NGO) Representative" (26%). Among them, "Nurses" are 19%, "Medical Experts (physicians)" are 17% and "Parents" are 11%.

The distribution of videos using animations or graphics was also analyzed, and it was found that 52% of the videos included animations or other forms of visual addition while 48% did not.

In Table 3, general features of the analyzed videos were summarized including video length, view counts, time passed since the first upload and the daily view counts.

Table 4 shows the mean, standard deviation, and reliability of the score of mDISCERN and GQS. In mDISCERN (1), the

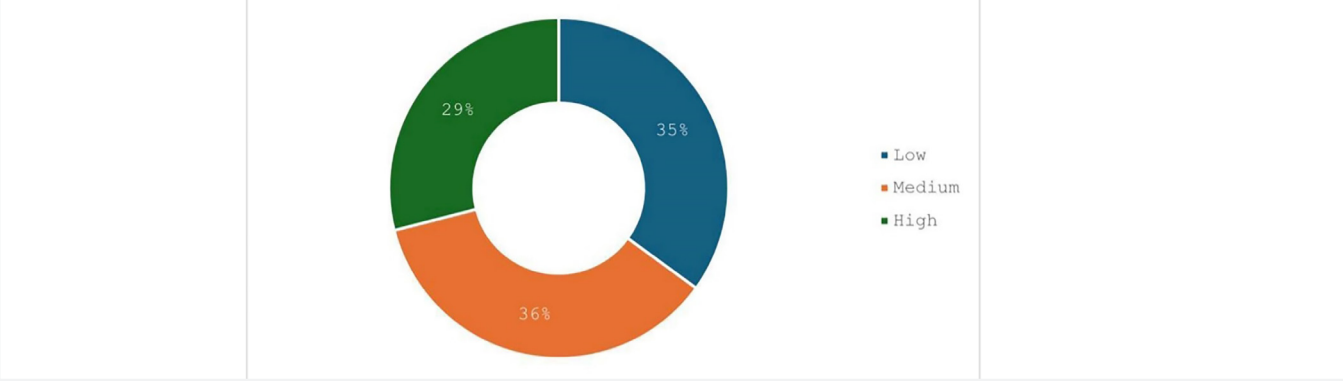


Figure 3. Evaluation of video quality

Figure 3 categorizes the quality of the videos produced as low, medium, or high. Regarding the intensity of the videos, 36% of them can be classified as having "medium" level, 35% of them are "low" and 29% of them are "high" Quality scale: 144-240p, low, 360-480p, medium,720p and higher: high

Table 3. General attributes of the analyzed videos

	Mean ± SD
Video length (minutes)	34.08±14.219
Views (count)	220588,89±144850,989
Time since upload (days)	1760,93±1150,943
Daily views (count)	2398,56±1340,593
Comments (count)	943,77±552,235
Likes (count)	21726,68±14480,347

SD: Standard deviation

Table 4. Interrater reliability assessment

	Mean ± SD	p-value	r	Cronbach's alpha
mDISCERN (1)	3.17±1.42	0.013	0.473	0.702
mDISCERN (2)	3.10±1.42			
GQS (1)	3.31±1.35	0.300	0.688	0.696
GQS (2)	3.13±1.43			

(1), (2) represents two independent evaluators. mDISCERN: Modified DISCERN, instrument name, GQS: Global quality scale, SD: Standard deviation

mean of the results is 3.17 and the p-value is 0.013, thus, the conclusion is that the results have statistical significance. Reliability of the mDISCERN (1) is moderate (Cronbach's alpha: 0.702) and reliability of mDISCERN (2) is not reported. GQS (1) has a mean of 3.31 and reliability coefficient of 0.696, which suggests reliability in the subjective quality assessment. GQS (2) has a mean of 3.13, but this study's p-value and reliability are not mentioned. Coefficients of correlation of mDISCERN (1) and GQS (1) are quite high and hence, it can be said that there is a significant positive correlation between these two measures of content quality.

Table 5 presents the difference of mDISCERN and GQS of the physician and non-physician presenters. Regarding the mDISCERN score, the mean for physicians is 3.10 and the mean for non-physicians is 3.16, and since the p-value is less than 0.05, there is statistical significance. The GQS score is also quite similar to the above scores (3.22 for physicians and 3.24 for non-physicians) and had a significant p-value. The slight variation in the scores indicates that both the presenter groups equally create videos of similar perceived quality, but the statistical analysis suggests that there is a small, yet significant difference between the two.

Table 6 shows the comparison of video characteristics between physician and non-physician presenters. Physicians make videos of 31.62 minutes on average, while non-physicians make videos of 37.09 minutes on average, making a near significant difference, where non-physicians are likely to produce content with longer duration.

For the non-physician group, the mean number of views stands at 228,076; for the physicians, the mean number is 214,463, but the p-value of 0.642 implies that it is not statistically significant. In the same way, comments and likes are quite similar in both groups with p-values of 0.896 and 0.512 respectively.

DISCUSSION

This research discovered important information about YouTube educational videos teaching KC for newborns. A study evaluation of 100 YouTube videos established important findings by using the GQS and the mDISCERN tool. The analysis of the presented YouTube educational videos on KC for newborns raises the awareness of the site's importance for sharing health-related information and demonstrates the variation in the quality of the materials. The present study also showed that even though many videos can be informative, their credibility and efficiency depend on the presenter, video quality, and if they include educational aids like animations and graphics. Videos created by physicians presented superior quality standards than those managed by non-medical personnel. The finding validates existing research which establishes medical professional-created videos present students with the most dependable and properly organized educational materials. The fact that educators and NGO representatives are among the most active users indicates that both online education and NGO are prominent in health-related video production. However, the appearance of physicians and nurses can be considered as a sign of a professional input; however, their numbers are not very high, which may mean that the representatives of the healthcare sector are not very active in the use of videos for public health promotion. Based on our findings, the majority of the videos have average production quality, meaning that they may have average picture and sound quality, but poor editing. This clearly indicates that techniques that are time-consuming and costly when producing health education videos are not very common because they may not capture the viewers' attention or are not very credible. Moreover, animation-assisted videos comprise almost half of the total videos, which means that although most of the creators use graphical elements to supplement their content, whereas the other half does not include any graphic elements in their videos apart from the usual video player interface. Incorporation of animations and graphics enhances viewership, understanding and recall of contents, especially where they are sensitive or technical in nature. Therefore, 48% of the videos that do not contain such elements may use other approaches like direct

Table 5. Comparison of scores between physicians and non-physicians

	Physicians	Non-physicians	p-value
	Mean ± SD	Mean ± SD	
mDISCERN	3.10±1.03	3.16±1.04	0.01
GQS	3.22±0.99	3.24±1.05	0.03

mDISCERN: Modified DISCERN, instrument name, GQS: Global quality scale, SD: Standard deviation

Table 6. Content analysis of physicians and non-physicians' videos

	Physicians	Non-physicians	p-value
	Mean ± SD	Mean ± SD	
Video length	31.62±14.59	37.09±13.29	0.055
Views	214463,07±152841,47	228076,00±135775,95	0.642
Comments	950,36±553,16	935,71±557,23	0.896
Likes	22590,76±15476,44	20670,58±13257,07	0.512

SD: Standard deviation

explanation, modelling, or text descriptions. The distribution is nearly equal to 50/50, which means that although animations are helpful, they do not seem to be widely implemented, which can be due to a lack of resources, knowledge of how to create them, or preference for the content format. Further studies could be done to compare the level of engagement, the amount of knowledge retained, or perceived credibility when using animated videos against non-animated videos.

The mDISCERN scores measured a reliability level of 4.56±0.50 for professional videos but non-professional videos received significantly lower scoring at 3.32±0.85. The research by Yapar Gümüş and Kaykı[1] demonstrated similar findings showing the better reliability characterizing YouTube videos when physicians create content instead of non-physicians for newborn care information. Osman et al.[7] demonstrated that health-related videos made by qualified professionals deliver better quality content which proves that healthcare sources determine the accuracy of medical videos.

Professional-made medical videos achieved higher GQS scores (4.66±0.47) than non-professional-made videos (3.30±0.91) because professionals arranged their contents more effectively. The findings support Haslam et al.[9] who stated that structured and easy-to-understand videos create effective health education systems. Videos created by professionals demonstrate better presentation of KC through systematic explanations including specific information about biological advantages combined with step-by-step guidance and proven medical recommendations that caregivers need to know.

Many video content pieces focused on general KC education while giving primary consideration to its advantages regarding preterm infant and neonatal intensive care unit newborn benefits. Engagement likely increases through the combination of scientific information with healthcare provider and parental experiences in these videos. Hernández-García and Giménez-Júlvez[13] along with other researchers discovered videos containing a mixture of factual information and personal stories improved both the relatability and popularity of health videos such as coronavirus disease 2019 prevention materials in Spanish.

Professional-made videos received most viewer engagement through likes and comments according to the research findings. The videos accumulated more viewer interactions thus showing audiences considered them to be trusted reliable information sources. The results match those presented by Lee et al.[6] showing that educational material effectiveness and trustworthy quality depend on video content quality together with user engagement metrics like comments and ratings. The level of viewer interactions functions as an important indicator to evaluate how useful and credible videos appear in transmitting important health-related practices. Live videos produced by non-professional creators exhibited inadequate standards even though the general video quality improved when professionals provided content. Most of these videos are based on the experiences of people and not evidence-based, which may lead to misinformation or lack of information in most cases. These educational videos usually do not show their sources correctly or they produce incorrect and obsolete content. Kohler and Dietrich[5] along with other studies showcased that user-generated content on YouTube faces accuracy and completion limitations according to their research. The need exists to introduce both better regulations for content creators and viewer-friendly reliability assessment features to ensure accurate health information availability. On the other hand, our findings showed that the presenter's professionalism does not affect the video views or the interactions, and other factors could be at play, such as the quality of production or the content of what is presented. However, while accepting the fact that the content created by professionals is generally credible, one of the limitations found was the inconsistency in the duration of the videos. Whereas, the professional videos offered more detailed information, they were also longer; the retention of such videos may not be as high as the short ones. On the other hand, non-professional videos were shorter, but they were not sufficient in providing adequate information for learning. This implies that videos should be lengthy to offer the relevant content but short to capture the attention of the viewers.

The research revealed one main drawback in the videos because professional and non-professional videos had varying durations. Some videos made by professionals provided detailed information whereas shorter videos from non-professional sources showed limited explanations regarding KC. The videos made by non-professionals displayed shorter durations because they usually presented personal stories instead of relying on evidence-based medical practices. According to Campbell-Yeo et al.[2] video length and depth determine educational effectiveness as an educational resource.

The research results illustrate why creator qualifications play a critical role in establishing the quality level of instructional videos discovered on YouTube. Users encounter various levels of reliability and usefulness on the platform since non-professional creators produce content alongside physician-contributed videos. Medical employees together with healthcare organizations should develop along with distribute higher-quality evidence-based content to give parents and caregivers dependable information about KC procedures.

Study Limitations

Several restrictions would influence the interpretation of the results in this research. The analysis's exclusion of non-English-

language videos may prevent the study findings from being relevant to non-English-speaking audiences. Since YouTube operates globally, important educational content in languages other than English was possibly omitted, thus reducing the study's ability to generalize conclusions across broader audiences.

The research analyzed only the 100 videos that appeared first in the YouTube search results. Video selection process maintains high relevance but might fail to demonstrate the complete availability of KC content on the YouTube platform. Useful educational videos with lower search rankings were probably omitted from this review.

Video quality and reliability assessment process depend on subjectivity since independent reviewers must provide their ratings. The researchers attempted to decrease scoring bias through agreement-based evaluations between reviewers and verification methods, but recognition differences between evaluators remain possible.

The study did not address the continuous evolution of YouTube content during its research period. Emerging video quality and relevance changes because of increased viewer response could influence study results.

CONCLUSION

This study has important implications for the stakeholders of the healthcare, content production, and public health sectors. The major conclusion of this study is that the videos prepared by the medical personnel have higher content credibility and effectiveness, therefore physicians and institutions are urged to produce credible and evidence-based videos in improving the quality of online medical education. YouTube videos created by people who are not medical professionals may be entertaining, but they do not have the accuracy needed for medical information sharing. Besides, the quality of production has a very significant influence over viewers' engagement. The videos that are professionally produced with good quality will attract more views, likes, and comments. Furthermore, animations and graphics were positively linked to comprehension and recall of content, suggesting that authors and other content developers, especially in the health sector, should attempt to use other visuals to support the educational goals. Judging from the case of KC and other health-related issues, due to the lack of truthfulness and credibility of YouTube videos, the future studies and policies should therefore shift to finding ways of promoting credible, reviewed articles. Last but not least, digital literacy should be promoted so that social media users, like YouTube viewers, can discern the content they are subjected to.

Ethics

Ethics Committee Approval: Ethics approval for this study was not deemed necessary as it involved the analysis of publicly available data from YouTube videos, which do not involve human or animal participants. Since YouTube videos are publicly accessible and do not require permission from the content creators for viewing or analysis, no formal permission was sought from the platform.

Informed Consent: The study adhered to ethical guidelines by ensuring that all data collected were publicly available and did not involve personal or sensitive information.

Footnotes

Authorship Contributions

Concept: M.T.A., Data Collection or Processing: M.T.A., Analysis or Interpretation: M.T.A., N.S., Literature Search: M.T.A., N.S., Writing: M.T.A., N.S., Literature Search: M.T.A., N.S., Writing: M.T.A.

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

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

REFERENCES

1. Yapar Gümüş C, Kaykı G. Youtube videos as a source of information for families on 'basic newborn care'. Kırıkkale Uni Med J. 2024;26:284-8.
2. Campbell-Yeo M, Dol J, Disher T, Benoit B, Chambers CT, Sheffield K, Boates T, et al. The power of a parent's touch: evaluation of reach and impact of a targeted evidence-based YouTube video. J Perinat Neonatal Nurs. 2017;31:341-9.
3. Bueno M, Nishi ÉT, Costa T, Freire LM, Harrison D. Blood sampling in newborns: a systematic review of YouTube videos. J Perinat Neonatal Nurs. 2017;31:160-5.
4. Harrison D, Reszel J, Dagg B, Aubertin C, Bueno M, Dunn S, et al. Pain management during newborn screening: using YouTube to disseminate effective pain management strategies. J Perinat Neonatal Nurs. 2017;31:172-7.
5. Kohler S, Dietrich TC. Potentials and limitations of educational videos on YouTube for science communication. Frontiers in Communication. 2021;6:581302.
6. Lee CS, Osop H, Goh DH-L, Kelni G. Making sense of comments on YouTube educational videos: a self-directed learning perspective. Online Information Review. 2017;41:611-25.
7. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. BMC Med Educ. 2022;22:382.
8. Drozd B, Couvillon E, Suarez A. Medical YouTube videos and methods of evaluation: literature review. JMIR Med Educ. 2018;4:e3.
9. Haslam K, Doucette H, Hachey S, MacCallum T, Zwicker D, Smith-Brilliant M, et al. YouTube videos as health decision aids for the public: an integrative review. Can J Dent Hyg. 2019;53:53-66.

10. Hernández-García I, Giménez-Júlvez T. Characteristics of YouTube videos in spanish on how to prevent COVID-19. Int J Environ Res Public Health. 2020;17:4671.

11. Harling J, Kaneza N. Informational videos to improve maternal knowledge of and comfort with using expressed breast milk in premature infants in rural Burundi. J Afric Neonatal. 2024;2:71-6.

12. Delli K, Livas C, Vissink A, Spijkervet FK. Is YouTube useful as a source of information for Sjögren's syndrome? Oral Dis. 2016;22:196-201.

13. Hernández-García I, Giménez-Júlvez T. Assessment of health information about COVID-19 prevention on the internet: infodemiological study. JMIR Public Health Surveill. 2020;6:e18717.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.35492
Nam Kem Med J 2025;13(3):231-245

Comprehensive Mapping of Psychosocial Burden in Breast Cancer: A Multicenter Cross-Sectional Study in Türkiye

Meme Kanserinde Psikososyal Yükün Kapsamlı Değerlendirilmesi: Türkiye'de Çok Merkezli Kesitsel Bir Çalışma

İvo GÖKMEN[1], Didem DİVRİKLİOĞLU[2], Erkan ÖZCAN[3], Veli ÇAKICI[4], Alper COŞKUN[5], Nazan DEMİR[6], Sernaz TOPALOĞLU[2]

[1]*Mehmet Akif Ersoy State Hospital, Clinic of Medical Oncology, Çanakkale, Türkiye*
[2]*Trakya University Faculty of Medicine, Department of Internal Medicine, Division of Medical Oncology, Edirne, Türkiye*
[3]*Kastamonu Training and Research Hospital, Clinic of Medical Oncology, Kastamonu, Türkiye*
[4]*Çanakkale Onsekiz Mart University Faculty of Medicine, Department of Internal Medicine, Division of Medical Oncology, Çanakkale, Türkiye*
[5]*Bursa Uludağ University Faculty of Medicine, Department of Internal Medicine, Division of Medical Oncology, Bursa, Türkiye*
[6]*Sultan 1st Murat State Hospital, Clinic of Medical Oncology, Edirne, Türkiye*

ABSTRACT

Aim: This study aimed to investigate the sociodemographic and clinical factors associated with depression and anxiety in women diagnosed with breast cancer and to identify predictors that may contribute to psychological distress.

Materials and Methods: This multicenter cross-sectional study included 460 breast cancer patients assessed via structured interviews and medical records. Depression and anxiety symptoms were measured using validated tools (Patient Health Questionnaire-9 and Generalized Anxiety Disorder-7), with scores ≥10 indicating clinical significance. Logistic regression analyses identified independent predictors.

Results: Clinically significant depression and anxiety were observed in 24.6% and 27.2% of the participants, respectively. Depression was independently associated with younger age [adjusted odds ratio (AOR): 4.68], being childless (AOR: 2.47), low income (AOR: 3.35), limited healthcare access (AOR: 3.34), and low social support (AOR: 6.38). Clinical predictors included premenopausal status (AOR: 2.86), poor sleep (AOR: 2.18), lymphedema (AOR: 2.55), advanced cancer stage (AOR: 1.65), and active chemotherapy (AOR: 2.61). Anxiety was similarly linked to younger age (AOR: 2.93), poor access to care (AOR: 3.84), low social support (AOR: 4.34), and ongoing treatments including chemotherapy and hormone therapy.

Conclusion: Depression and anxiety are prevalent among breast cancer patients and are strongly associated with both sociodemographic disadvantages and clinical disease burden. Routine psychological screening should be integrated into oncology care to support patient well-being and optimize outcomes.

Keywords: Breast cancer, depression, anxiety, psychosocial burden, PHQ-9, GAD-7

ÖZ

Amaç: Meme kanseri tanısı almış kadınlarda depresyon ve anksiyete ile ilişkili sosyodemografik ve klinik etkenleri araştırmak ve psikolojik sıkıntıya yol açabilecek belirleyicileri tanımlamaktır.

Gereç ve Yöntem: Bu çok merkezli kesitsel çalışmaya, yapılandırılmış görüşmeler ve tıbbi kayıtlar aracılığıyla değerlendirilen 460 meme kanseri hastası dahil edildi. Depresyon ve anksiyete semptomları, geçerliliği kanıtlanmış araçlar (Hasta Sağlığı Anketi-9 ve Yaygın Anksiyete Bozukluğu-7) ile ölçüldü; ≥10 puan klinik olarak anlamlı kabul edildi. Bağımsız belirleyicileri saptamak amacıyla lojistik regresyon analizleri yapıldı.

Address for Correspondence: İvo GÖKMEN MD, Mehmet Akif Ersoy State Hospital, Clinic of Medical Oncology, Çanakkale, Türkiye
E-mail: ivo_georgiev1@hotmail.com **ORCID ID:** orcid.org/0000-0002-0382-5396
Received: 15.04.2025 **Accepted:** 11.05.2025 **Publication Date:** 07.10.2025

Cite this article as: Gökmen İ, Divriklioğlu D, Özcan E, Çakıcı V, Coşkun A, Demir N, et al. Comprehensive mapping of psychosocial burden in breast cancer: a multicenter cross-sectional study in Türkiye. Nam Kem Med J. 2025;13(3):231-245

Bulgular: Katılımcıların sırasıyla %24,6'sında klinik düzeyde depresyon ve %27,2'sinde anksiyete tespit edildi. Depresyon; genç yaş [düzeltilmiş olabilirlik oranı (AOR): 4,68], çocuksuzluk (AOR: 2,47), düşük gelir (AOR: 3,35), sınırlı sağlık hizmeti erişimi (AOR: 3,34) ve düşük sosyal destek (AOR: 6,38) ile bağımsız olarak ilişkiliydi. Klinik belirleyiciler arasında premenopozal durum (AOR: 2,86), kötü uyku kalitesi (AOR: 2,18), lenfödem (AOR: 2,55), ileri evre kanser (AOR: 1,65) ve aktif kemoterapi (AOR: 2,61) yer aldı. Anksiyete de benzer şekilde genç yaş (AOR: 2,93), yetersiz sağlık hizmeti erişimi (AOR: 3,84), düşük sosyal destek (AOR: 4,34) ve kemoterapi ile hormon tedavisi gibi devam eden tedavilerle ilişkiliydi.

Sonuç: Depresyon ve anksiyete, meme kanseri hastalarında yaygındır ve hem sosyodemografik dezavantajlar hem de klinik hastalık yükü ile güçlü şekilde ilişkilidir. Rutin psikolojik tarama, onkolojik bakıma entegre edilmeli; böylece hasta iyilik hali desteklenerek sonuçlar optimize edilebilir.

Anahtar Kelimeler: Meme kanseri, depresyon, anksiyete, psikososyal yük, PHQ-9, GAD-7

INTRODUCTION

Breast cancer is the most common cancer in women globally, with approximately 2.3 million new cases annually, representing 24.5% of female cancers and 15% of cancer-related deaths. By 2050, incidence is projected to surpass 3.2 million, a 38% increase from 2022[1,2]. This rise presents growing challenges for diagnosis and treatment, especially in resource-limited settings[3].

Beyond its physical toll, breast cancer imposes a significant psychological burden. Studies report that 29-47% of the patients experience psychiatric conditions such as depression, anxiety, and distress, with up to 8% reporting suicidal ideation and 2-6% attempting suicide, particularly in terminal stages[4,5]. One in five women continues to experience depressive symptoms two years after diagnosis[6].

The severity and frequency of these symptoms often reflect underlying sociodemographic factors. Younger age, low education, limited social support, and lack of a stable relationship are consistently linked to greater psychological distress[7-9]. In low- and middle-income settings, these effects are intensified by poor healthcare access, mental health stigma, and fatalistic cultural views of cancer[10,11].

These psychological factors do not only reduce quality of life but also impair treatment adherence, exacerbate fatigue and pain, and may increase mortality[12,13]. By worsening disease progression and treatment outcomes, these psychological factors increase the overall burden of breast cancer, which remains one of the leading causes of lost healthy years measured in disability-adjusted life years worldwide[14]. These findings underscore the need for psychological assessment in comprehensive breast cancer care. Yet in many oncology settings, focus remains on physical symptoms, while depression and anxiety are often overlooked or poorly managed[15].

This study aimed to quantify the prevalence of psychological distress in breast cancer patients, examine related sociodemographic and clinical factors, and emphasize the importance of integrating psychosocial support into routine oncology care. It is among the few studies in Türkiye to use validated tools the Patient Health Questionnaire-9 (PHQ-9) and the Generalized Anxiety Disorder-7 (GAD-7) widely used in psycho-oncology and known for strong psychometric reliability[16,17]. Their standardized use enhances data accuracy and generalizability, supporting evidence-based care and informing national health policy.

MATERIALS AND METHODS

Study Design and Setting

This multicenter, cross-sectional descriptive study was conducted in six healthcare institutions across Türkiye. Participant recruitment began after ethical approval on January 6, 2025, and continued until mid-March. Data were collected and analyzed for sociodemographic, clinical, and psychological variables.

The study was conducted at three university-affiliated medical centers Trakya University Faculty of Medicine Hospital, Bursa Uludağ University Faculty of Medicine Hospital, and Çanakkale Onsekiz Mart University Faculty of Medicine Hospital and three state hospitals: Çanakkale Mehmet Akif Ersoy State Hospital, Kastamonu Training and Research Hospital, and Edirne Sultan 1st Murat State Hospital.

Participants and Inclusion Criteria

Eligible participants were women diagnosed with non-metastatic breast cancer who voluntarily agreed to participate. Inclusion criteria were: (1) age 18 years or older, (2) no current or past diagnosis of psychiatric disorder or dementia, (3) no history of alcohol or substance dependence, and (4) sufficient Turkish language proficiency for effective communication. Pregnancy and the presence of another malignancy were exclusion criteria. The sample size was calculated using Cochran's formula, assuming a 30% prevalence of depression/anxiety, with a 95% confidence level and 5% margin of error, resulting in a minimum requirement of 323 participants. This estimate was based on the lower end of previously reported prevalence rates (4-5) and reflects a conservative approach. However, to enhance the study's statistical power and enable subgroup analyses, the final sample size was increased to 460.

Data Collection Procedure

Data were collected through face-to-face interviews in outpatient clinics using structured questionnaires, which included three sections: sociodemographic data, clinical information, and standardized psychological tools (PHQ-9 and GAD-7). Participants were either receiving active oncological treatment (chemotherapy, radiotherapy, or endocrine therapy) or were in routine follow-up after primary treatment. All interviews were conducted by the researchers themselves, all of whom were directly involved in the clinical care of participants. Written informed consent was obtained from all participants after they were informed about the study's purpose and procedures. Interviews took place in private rooms to ensure confidentiality and lasted approximately 10-20 minutes, depending on responsiveness.

Variable Definition and Grouping Strategy

To ensure analytical clarity and minimize multicollinearity risk, variables were organized into two main domains: (1) individual and sociodemographic characteristics, and (2) clinical variables related to disease progression and treatment. The inclusion of a broad range of variables was based on sample size capacity and subgroup balance.

Multicategorical variables were dichotomized when preliminary analyses showed no significant differences or when distributions were highly unbalanced. Grouping decisions followed conceptually meaningful thresholds and clinically relevant cut-offs reported in the literature [e.g., age: <50 vs. ≥50 years; body mass index (BMI): <30 vs. ≥30 kg/m^2; marital status: married vs. single/divorced].

"Low income" was defined based on the net national minimum wage in Türkiye as of January 2025 (TRY 22,104.67; ~USD 631 at 1 USD: 35 TRY). Monthly household income below twice this amount (<TRY 44,209) was categorized as low income. Participants with income equal to or exceeding this threshold (≥TRY 44,209) were classified as having "moderate–high" income. Menopausal status was based on menstrual history; women with ongoing menstruation or <12 months of amenorrhea were classified as premenopausal. Subjective variables—physical activity, healthcare access, and perceived social support—were categorized using predefined criteria. Detailed definitions are provided in table footnotes.

Psychological Assessment Scales

Depression and anxiety symptoms were evaluated using two validated psychometric instruments: the PHQ-9 and the GAD-7. The PHQ-9 screens for major depressive disorder, while the GAD-7 assesses symptoms of generalized anxiety. Both scales measure symptom frequency over the past two weeks using a four-point Likert scale ranging from 0 (not at all) to 3 (nearly every day).

Total scores range from 0-27 for the PHQ-9 and 0–21 for the GAD-7, with higher scores indicating greater symptom severity. A cutoff score of ≥10 is widely accepted for both instruments as the threshold for clinically significant depression or anxiety. The Turkish versions of both scales have demonstrated strong psychometric validity in previous studies[18,19]. In this sample, both scales demonstrated high internal consistency, with Cronbach's alpha values of 0.81 for the PHQ-9 and 0.89 for the GAD-7, indicating strong reliability.

This study received ethical approval from the Non-Interventional Research Ethics Committee of Trakya University Faculty of Medicine (decision no: 01/14, date: 06.01.2025). Institutional permissions were obtained from all participating centers prior to study initiation. All procedures adhered to the principles of the Declaration of Helsinki, and informed written consent was obtained from all participants.

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA). Categorical variables were summarized as frequencies (n) and percentages (%), while continuous variables were expressed as means and standard deviations. Depression and anxiety were dichotomized based on established cut-off scores (≥10) for the PHQ-9 and GAD-7.

Initial group comparisons (e.g., depression vs. non-depression; anxiety vs. non-anxiety) were conducted using Pearson's chi-square test. Variables with p-values <0.10 were included in univariate logistic regression to calculate crude odds ratios with 95% confidence intervals (CIs).

Variables significant at $p<0.05$ in univariate analysis were entered into a multivariate logistic regression model to estimate adjusted odds ratios (AORs). Model fit was assessed using the Hosmer-Lemeshow test, explanatory power by Nagelkerke R^2, and classification accuracy was calculated. Statistical significance was set at $p<0.05$.

RESULTS

Sample Characteristics

A total of 460 women with non-metastatic breast cancer were included. Mean age was 54.8±12.1 years (range: 23-87), and 10.9% were under 40. Most participants were married (73.5%), postmenopausal (65.9%), and unemployed (77.2%). Regarding education, 51.5% had only primary education, while 19.1% held a university degree or higher. The average education duration was 8.0±4.8 years.

In terms of income, 28.3% reported a household income below the minimum wage. Urban residency was reported

by 70.7%, and 14.6% described healthcare access as "very difficult". Among the participants, 27.2% had no children, and 24.1% reported low perceived social support. Clinically, 34.1% were premenopausal. Obesity (BMI ≥30 kg/m²) was seen in 37.6%, and 27.4% had low physical activity. Poor sleep quality was reported by 17.8%. Cancer staging showed 44.1% in Stage II and 34.6% in Stage III.

Surgery was performed in 90.7% of cases, with breast-conserving surgery being most common (45.9%). Chemotherapy had been administered to 88.7%, and 12.6% were receiving active chemotherapy during data collection. Radiotherapy was given to 68.3%. Lymphedema was present in 23.7%, with 2.8% reporting severe symptoms. Time since diagnosis was less than two years in 59.3% of cases (complete descriptive data are provided in Table 1 A,B).

Table 1. Sociodemographic and clinical characteristics of the participants (n= 460)

Table 1. (A) Sociodemographic variables of the participants

Variable	n (%) or Mean ± SD
Age (years)	54.8±12.1
Age group	
≥70	57 (12.4%)
60-69	108 (23.5%)
50-59	130 (28.3%)
40-49	115 (25.0%)
18-39	50 (10.9%)
Marital status	
Married	338 (73.5%)
Single	28 (6.1%)
Divorced	94 (20.4%)
Number of children	
≥3	32 (7.0%)
2	193 (42.0%)
1	110 (23.9%)
0	125 (27.2%)
Years of education	8.0±4.8
Education level	
Primary school	237 (51.5%)
Middle school	49 (10.7%)
High school	86 (18.7%)
University and above	88 (19.1%)
Employment status	
Unemployed	335 (77.2%)
Employed	105 (22.8%)
Income level	
≥5x minimum wage	30 (6.5%)
3-4x minimum wage	127 (27.6%)
1-2x minimum wage	173 (37.6%)
<minimum wage	130 (28.3%)
Residential area	
Urban	325 (70.7%)
Rural	135 (29.3%)
Access to healthcare service	
Very easy	82 (17.8%)
Easy	221 (48.0%)
Difficult	90 (19.6%)
Very difficult	67 (14.6%)
Family history of breast cancer	
No	355 (77.2%)
Yes	105 (22.8%)
Perceived social support level	
Adequate	171 (37.2%)
Moderate	178 (38.7%)
Low	111 (24.1%)
Height (cm)	160±6.8
Weight (kg)	71.8±13.5

"Very easy" access refers to walking-distance healthcare facilities or the ability to reach physicians directly by phone. "Easy" access includes reasonable public transportation within the city (e.g., bus, minibus). "Difficult" access refers to long travel distances, irregular transportation, or financial barriers. "Very difficult" access reflects situations such as living in rural or remote areas, requiring referrals, or facing infrastructural limitations that hinder continuity of care. Perceived social support was not assessed using a validated scale. Instead, it was determined through brief face-to-face conversations in which patients described how emotionally, practically, and socially supported they felt. Based on this self-report, support levels were categorized as low, moderate, or high. SD: Standard deviation, n (%): Number and percentage of participants in each category, Urban: City or town, Rural: Village or small settlement. Access to healthcare was classified into four levels based on participants' responses during structured interviews

Table 1. (B) Clinical and treatment-related characteristics of the participants

Variable	n (%)
ECOG performance status	
0	397 (86.3%)
≥1	63 (13.7%)
Menopausal status	
Postmenopausal	303 (65.9%)
Premenopausal	157 (34.1%)
Body mass index	
<18.5	17 (3.7%)
18.5-24.9	156 (33.9%)
25-29.9	114(24.8%)
≥30	173(37.6%)
Physical	
High	27 (5.9%)
Moderate	307 (66.7%)
Low	126 (27.4%)

Table 1. (B) Continued	
Variable	**n (%)**
Sleep quality	
Good	211 (45.9%)
Moderate	167 (36.3%)
Poor	82 (17.8%)
Presence of lymphedema	
None	351 (76.3%)
Mild	96 (20.9%)
Severe	13 (2.8%)
Breast cancer stage	
Stage 1	98 (21.3%)
Stage 2	203 (44.1%)
Stage 3	159 (34.6%)
Type of surgery	
BCS	209 (45.9%)
Single MRM	181 (39.3%)
Bilateral MRM	27 (5.9%)
No Surgery	43 (9.3%)
Lymph node surgery	
SLNB	234 (50.9%)
ALND	183 (39.8%)
No Surgery	43 (9.3%)
Chemotherapy status	
Did not receive	52 (11.3%)
Received	350 (76.1%)
Active	58 (12.6%)
Radiotherapy status	
Did not receive	146 (31.7%)
Received	314 (68.3%)
Hormone therapy status	
Active	208 (45.2%)
Received	94 (20.4%)
Did not receive	158 (34.3%)
Duration since diagnosis	
≥120 months	27 (5.9%)
61-120 months	66 (14.3%)
25-60 months	94 (20.4%)
0-24 months	273 (59.3%)

Physical activity level was classified into three categories based on participants weekly total duration and frequency of activity. "Low" level reflected a predominantly sedentary lifestyle with less than 150 minutes of physical activity per week or activity on fewer than one day per week. "Moderate" level referred to 2-3 days of moderate-intensity activities (e.g., walking, household tasks), totaling approximately 150-300 minutes per week. "High" level indicated at least 4-5 days of regular activity per week, exceeding 300 minutes in total and including structured or vigorous physical exercise. Sleep quality was self-reported and categorized as "good," "moderate," or "poor" based on restfulness, sleep interruptions, and daytime fatigue. "Active chemotherapy" referred to patients undergoing chemotherapy at the time of data collection. Time since diagnosis was calculated from the date of pathology-confirmed diagnosis to the date of participation. n (%): number and percentage of participants in each category, ECOG: Eastern Cooperative Oncology Group, BCS: Breast-conserving surgery, MRM: Modified radical mastectomy, SLNB: Sentinel lymph node biopsy, ALND: Axillary lymph node dissection

Depression and Anxiety Scores

The mean PHQ-9 score was 6.61±4.42, and the mean GAD-7 score was 5.32±4.21. Using a cut-off of ≥10, 113 participants (24.6%) screened positive for clinically significant depression, and 125 (27.2%) for anxiety. Regarding depression severity, 19.8% reported mild symptoms, 3.3% moderate, and 1.5% severe. For anxiety, 24.8% had mild and 3.0% had moderate to severe symptoms. Among those below the clinical threshold, minimal symptoms were most common 37.0% for depression and 48.5% for anxiety (see Table 2 for full distribution details).

Factors Associated with Depression

Separate multivariate logistic regression models were conducted to assess associations between depression and sociodemographic (Table 3A) and clinical variables (Table 3B).

In the sociodemographic model, younger age (<50 years) (AOR: 4.68, 95% CI: 2.53-8.67), childlessness (AOR: 2.47, 95% CI: 1.40-4.37), low income (AOR: 3.35, 95% CI: 1.72-6.52), limited healthcare access (AOR: 3.34, 95% CI: 1.95-5.70), and low perceived social support (AOR: 6.38, 95% CI: 3.61-11.26) were significantly associated with higher odds of depression. Although marital status was significant in univariate analysis (crude odds ratio: 1.97, 95% CI: 1.25-3.12), it did not remain in the multivariate model. Model performance showed a

Table 2. Distribution of depression and anxiety levels based on PHQ-9 and GAD-7 scores

Scales	Mean ± SD	n	(%)
PHQ-9 (total score)	6.61±4.42	460	(100%)
≥10 (presence of depression)		**113**	**(24.6%)**
10-14	11.51±1.18	91	(19.8%)
15-19	16±0.92	15	(3.3%)
20-27	20.57±0.78	7	(1.5%)
<10 (absence of depression)			
0-4	2.22±1.41	170	(37.0%)
5-9	6.91±1.31	177	(38.5%)
GAD-7 (Total score)	5.32±4.21	460	(100%)
≥10 (presence of anxiety)		**125**	**(27.2%)**
10-14	10.45±0.77	114	(24.8%)
15-21	16.07±1.49	14	(3.0%)
<10 (absence of anxiety)			
0-4	1.69±1.26	233	(48.5%)
5-9	6.32±1.25	113	(24.6%)

"Presence of depression" and "presence of anxiety" were defined by a cut-off score of ≥10. The subgroups represent severity classifications: PHQ-9 → 10–14 (moderate), 15–19 (moderately severe), 20–27 (severe); GAD-7 → 10–14 (moderate), 15–21 (severe). Scores <10 indicate minimal or mild symptoms

PHQ-9: Patient Health Questionnaire-9, GAD-7: Generalized Anxiety Disorder-7, SD: Standard deviation

Nagelkerke R^2 of 0.307, classification accuracy of 83.7%, and a Hosmer-Lemeshow p-value of 0.005, indicating limited model fit.

In the clinical model, significant predictors included premenopausal status (AOR: 2.86, 95% CI: 1.78-4.61), poor sleep quality (AOR: 2.18, 95% CI: 1.09-4.38), lymphedema (AOR: 2.55, 95% CI: 1.53-4.25), advanced cancer stage (AOR: 1.65, 95% CI: 1.03-2.64), and active chemotherapy (AOR: 2.61, 95% CI: 1.39-4.89). BMI and physical activity were significant only in univariate analysis. The clinical model showed good fit (Hosmer–Lemeshow p= 0.696), with a Nagelkerke R^2 of 0.174 and classification accuracy of 75.9%.

Table 3. Factors associated with depression based on sociodemographic and clinical variables

Table 3. (A) Associations between depression and sociodemographic variables

	All patients	Depression present	Depression absent		Bivariate logistic regression analysis	Multivariate logistic regression models
Variable	**n (%)**	**n (%)**	**n (%)**	**p-value**	**COR (95% CI)**	**AOR (95% CI)**
Age (years)						
≥50	294 (63.9%)	56 (19.0%)	238 (81.0%)	**0.001**	Ref.	
<50	166 (36.1%)	57 (34.3%)	109 (65.7%)		2.22 (1.44-3.42)**	4.68 (2.53-8.67)**
Marital status						
Married	338 (73.5%)	71 (21.0%)	267 (79.9%)	**0.005**	Ref.	
Not married	122 (26.5%)	42 (34.4%)	80 (65.6%)		1.97 (1.25-3.12)**	-
Parental status						
With children	335 (72.8%)	69 (20.6%)	266 (79.4%)	**0.002**	Ref.	
Childless	125 (27.2%)	44 (35.2%)	81 (64.8%)		2.09 (1.33-3.29)**	2.47(1.40-4.37)**
Educational attainment						
Low	286 (62.2%)	62 (21.7%)	224 (78.3%)	0.074	Ref.	
High	174 (37.8%)	51 (29.3%)	123 (70.7%)		1.50 (0.97-2.31)	-
Employment status						
Employed	105 (22.8%)	30 (28.6%)	75 (71.4%)	0.302		
Unemployed	355 (77.2%)	83 (23.4%)	272 (76.6%)		-	-
Income level						
Low income	157 (34.1%)	29 (18.5%)	128 (81.5%)	**0.03**	Ref.	
Moderate-high	303 (65.9%)	84 (27.7%)	219 (72.3%)		1.69(1.05-2.72)*	3.35(1.72-6.52)**
Residential area						
Urban	325 (70.7%)	78 (24.0%)	247 (76.0%)	0.721		
Rural	135 (29.3%)	35 (25.9%)	100 (74.1%)		-	-
Healthcare accessibility						
Easy	303 (65.9%)	60 (19.8%)	243 (80.2%)	0.001	Ref.	
Difficult	157 (34.1%)	53 (33.8%)	104 (66.2%)		2.06 (1.34-3.19)**	3.34(1.95-5.70)**
Family history of breast cancer						
No	355 (77.2%)	89 (25.1%)	266 (74.9%)	0.7		
Yes	105 (22.8%)	24 (22.9%)	81 (77.1%)		-	-
Perceived social support level						
Adequate/moderate	349 (75.9%)	61 (17.5%)	288 (82.5%)	**<0.001**	Ref.	
Low	111 (24.1%)	52 (46.8%)	59 (53.2%)		4.16 (2.62-6.62)**	6.38(3.61-11.26)**

*p<0.05, **p<0.001, Ref.: Reference category for odds ratio comparisons, "Educational attainment" was grouped as follows: "Low" includes primary and middle school education; "High" includes high school and university education. "Healthcare accessibility" was dichotomized: "Easy" includes both "Very Easy" and "Easy"; "Difficult" includes both "Difficult" and "Very Difficult." COR: Crude odds ratio, AOR: Adjusted odds ratio, CI: Confidence interval. P-values are based on chi-square tests comparing depression rates between groups. Hosmer-Lemeshow test: p: 0.005; Nagelkerke R^2: 0.307, Overall Classification Accuracy: 83.7%

Table 3. (B) Associations between depression and clinical/treatment-related variables

	All patients	Depression present	Depression absent		Bivariate logistic regression analysis	Multivariate logistic regression models
Variable	**n (%)**	**n (%)**	**n (%)**	**p-value**	**COR (95% CI)**	**AOR (95% CI)**
ECOG performance status						
0	397 (86.3%)	92 (23.3%)	305 (76.8%)	0.085	Ref.	
≥1	63 (13.7%)	21 (33.3%)	42 (66.7%)		1.65 (0.93-2.94)	-
Menopausal status						
Postmenopausal	303 (65.9%)	58 (19.1%)	245 (80.9%)	**0.001**	Ref.	
Premenopausal	157 (34.1%)	55 (35.0%)	102 (65.0%)		2.27 (1.47-3.52)**	2.86(1.78-4.61)**
Body mass index						
<30	287 (62.4%)	57(19.9%)	230 (80.1%)	**0.009**	Ref.	
≥30	173 (37.6%)	56 (32.4%)	117 (67.6%)		1.93 (1.25-2.97)*	-
Physical activity level						
High/moderate	334 (72.6%)	72 (21.6%)	262 (78.4%)	**0.03**	Ref	
Low	126 (27.4%)	41 (32.5%)	85 (67.5%)		1.65 (1.14-2.76)*	-
Sleep quality						
Good/Moderate	378 (82.2%)	80 (21.2%)	298 (78.8%)	**0.001**	Ref.	
Poor	82 (17.8%)	33 (40.2%)	49 (59.8%)		2.51 (1.51-4.16)**	2.18 (1.09-4.38)*
Presence of lymphedema						
Absent	351 (76.3%)	74 (21.1%)	277 (78.9%)	**0.003**	Ref.	
Present	109 (23.7%)	39 (35.8%)	70 (64.2%)		2.08 (1.31-3.33)*	2.55 (1.53-4.25)**
Breast cancer stage						
Early (I-II)	301 (65.4%)	63 (20.9%)	238 (%79.1%)	0.017	Ref	
Advanced (III)	159 (34.6%)	50 (31.4%)	109 (68.6%)		1.73 (1.12-2.67)*	1.65 (1.03-2.64)*
Type of surgery (excluding non-surgical cases)						
BCS	209 (50.1%)	47 (22.5%)	162 (77.5%)	0.646		
Mastectomy#	208 (49.9%)	51 (24.5%)	157 (75.5%)		-	-
Lymph node surgery (excluding non-surgical cases)						
SLNB	234 (56.1%)	57 (24.4%)	177 (75.6%)	0.727		
ALNB	183 (43.9%)	41 (22.4%)	142 (77.6%)		-	-
Chemotherapy status						
Not receiving active	402 (87.4%)	89 (22.1%)	313 (77.9%)	**0.003**	Ref	
Active	58 (12.6%)	24 (41.4%)	34 (58.6%)		2.48 (1.4-4.4)*	2.61 (1.39-4.89)**
Radiotherapy status						
Did not receive	146 (31.7%)	38 (26.0%)	108 (74.0%)	0.642		
Received	314 (68.3%)	75 (23.9%)	239 (76.1%)		-	-
Hormone therapy status						
Not receiving active	252 (54.8%)	61 (24.2%)	191 (75.8%)	0.913		
Active	208 (45.%)	52 (25.0%)	156 (75.0%)		-	-
Time since diagnosis						
>24 months	187 (40.7%)	41 (21.9%)	146 (78.1%)	0.321		
≤24 months	273 (59.3%)	72 (26.4%)	201 (73.6%)		-	-

*p<0.05, **p<0.001, Ref.: Reference category for odds ratio comparisons, BMI: Body mass index, SLNB: Sentinel lymph node biopsy, ALND: Axillary lymph node dissection, BCS: Breast-conserving surgery, Mastectomy# includes both modified radical and bilateral mastectomy, HRT: Hormone replacement therapy, COR: Crude odds ratio, AOR: Adjusted odds ratio, CI: Confidence interval. p-values are based on chi-square tests comparing the two groups. Hosmer-Lemeshow test: p= 0.696, Nagelkerke R^2: 0.174, Overall classification accuracy: 75.9

Factors Associated with Anxiety

Separate multivariate logistic regression models were conducted to examine associations between anxiety and both sociodemographic (Table 4A) and clinical variables (Table 4B).

In the sociodemographic model, participants under 50 years (AOR: 2.93, 95% CI: 1.81-4.77), those with higher education (AOR: 2.43, 95% CI: 1.49-3.96), limited access to healthcare (AOR: 3.84, 95% CI: 2.33-6.34), and low perceived social support (AOR: 4.34, 95% CI: 2.57-7.34) showed significantly higher odds of anxiety. Other factors such as marital status, employment, income, and parental status were not significant in the adjusted model. Model fit was acceptable, with a Nagelkerke R^2 of 0.220, classification accuracy of 77.4%, and Hosmer–Lemeshow p= 0.079.

Table 4. Factors associated with anxiety based on sociodemographic and clinical variables

Table 4. (A) Associations between anxiety and sociodemographic variables

	All patients	Anxiety present	Anxiety absent		Bivariate logistic regression analysis	Multivariate logistic regression models
Variable	**n (%)**	**n (%)**	**n (%)**	**p-value**	**COR (95% CI)**	**AOR (95% CI)**
Age (years)						
≥50	294 (63.9%)	58 (19.7%)	236 (80.3%)	**0.001**	Ref.	
<50	166 (36.1%)	67 (40.4%)	99 (59.6%)		2.75 (1.81-4.21)**	2.93 (1.81-4.77)**
Marital status						
Married	338 (73.5%)	93 (27.5%)	245 (72.5%)	0.613		
Not married	122 (26.5%)	32 (26.2%)	90 (73.8%)		-	-
Parental status						
With children	335 (72.8%)	88 (26.3%)	247 (73.7%)	0.381		
Childless	125 (27.2%)	37 (29.6%)	88 (70.4%)		-	-
Educational attainment						
Low	286 (62.2%)	63 (22.0%)	223 (78.0%)	0.002	Ref.	
High	174 (37.8%)	62 (35.6%)	112 (64.4%)		1.95 (1.29-2.97)**	2.43 (1.49-3..96)**
Employment status						
Employed	105 (22.8%)	32 (30.5%)	73 (69.5%)	0.385		
Unemployed	355 (77.2%)	93 (26.2%)	262 (73.8%)		-	-
Income level						
Low income	157 (34.1%)	40 (25.5%)	117 (74.5%)	0.582		
Moderate-high	303 (65.9%)	85 (28.1%)	218 (71.9%)		-	-
Residential area						
Urban	325 (70.7%)	88 (27.1%)	237 (72.9%)	0.9		
Rural	135 (29.3%)	37 (27.4%)	98 (72.6%)		-	-
Healthcare accessibility						
Easy	303 (65.9%)	60 (19.8%)	243 (80.2%)	**0.001**	Ref.	
Difficult	157 (34.1%)	65 (41.4%)	92 (58.6%)		2.86 (1.87-4.37)**	3.84 (2.33-6.34)**
Family history of breast cancer						
No	355 (77.2%)	96 (27.0%)	259 (73.0%)	0.901		
Yes	105 (22.8%)	29 (27.6%)	76 (72.4%)		-	-
perceived social support level						
Adequate/Moderate	349 (75.9%)	76 (21.8%)	273 (78.2%)	**0.001**	Ref.	
Low	111 (24.1%)	49 (44.1%)	62 (55.9%)		2.83 (1.81-4.46)**	4.34 (2.57-7.34)**

*p<0.05, **p<0.001, Ref.: Reference category for odds ratio comparisons, "Educational attainment" was grouped as follows: "Low" includes primary and middle school education, "High" includes high school and university education. "Healthcare accessibility" was dichotomized: "Easy" includes both "Very Easy" and "Easy", "Difficult" includes both "Difficult" and "Very Difficult." COR: Crude odds ratio, AOR: adjusted odds ratio, CI: Confidence interval. p-values are based on chi-square tests comparing depression rates between groups. Hosmer-Lemeshow test: p: 0.079, Nagelkerke R^2: 0.220, Overall classification accuracy: 77.4%

Table 4. (B) Associations between anxiety and clinical/treatment-related variables						
	All patients	Anxiety present	Anxiety absent		Bivariate logistic regression analysis	Multivariate logistic regression models
Variable	n (%)	n (%)	n (%)	p-value	COR (95% CI)	AOR (95% CI)
ECOG performance status						
0	397 (86.3%)	87 (21.9%)	310 (78.1%)	0.321		
≥1	63 (13.7%)	15 (23.8%)	48 (72.2%)		-	-
Menopausal status						
Postmenopausal	303 (65.9%)	60 (19.8%)	243 (80.2%)	**0.001**	Ref.	
Premenopausal	157 (34.1%)	65 (41.4%)	92 (58.6%)		2.86 (1.87-4.27)**	3.05 (1.92-4.87)**
Body mass index (BMI)						
<30	287 (62.4%)	71 (24.7%)	216 (75.3%)	0.132		
≥30	173 (37.6%)	54 (31.2%)	119 (68.8%)		-	-
Physical activity level						
High/moderate	334 (72.6%)	79 (23.7%)	255 (76.3%)	0.02	Ref	
Low	126 (27.4%)	46 (36.5%)	80 (63.5%)		1.35 (1.19-2.24)*	-
Sleep quality						
Good/Moderate	378 (82.2%)	93 (24.6%)	285 (75.4%)	0.009	Ref.	
Poor	82 (17.8%)	32 (39.0%)	50 (61.0%)		1.96 (1.18-3.23)*	2.42 (1.38-4.25)*
Presence of lymphedema						
Absent	351 (76.3%)	91 (25.9%)	260 (74.1%)	0.324		
Present	109 (23.7%)	34 (31.2%)	75 (68.8%)		-	-
Breast cancer stage						
Early (I-II)	301 (65.4%)	69 (22.9%)	232 (77.1%)	**0.006**	Ref	
Advanced (III)	159 (34.6%)	56 (35.2%)	103 (64.8%)		1.82 (1.19-2.78)*	1.85 (1.16-2.96)*
Type of surgery (excluding non-surgical cases)						
BCS	209 (50.1%)	56 (26.8%)	153 (73.2%)	0.305		
Mastectomy#	208 (49.9%)	46 (22.1%)	162 (77.9%)		-	-
Lymph node surgery (excluding non-surgical cases)						
SLNB	234 (56.1%)	60 (25.6%)	174 (74.4%)	0.567		
ALNB	183 (43.9%)	42 (23.0%)	141 (77.0%)		-	-
Chemotherapy status						
Not receiving active	402 (87.4%)	94 (23.4%)	308 (76.6%)	**0.001**	Ref	
Active	58 (12.6%)	31 (53.4%)	27 (46.6%)		3.76 (2.12-6.62)**	2.09 (1.05-4.15)*
Radiotherapy status						
Did not receive	146 (31.7%)	46 (31.5%)	100 (68.5%)	0.117		
Received	314 (68.3%)	79 (25.2%)	235 (74.8%)		-	-
Hormone therapy status						
Not receiving active	252 (54.8%)	54 (21.4%)	198 (78.6%)	**0.009**	Ref	
Active	208 (45.%)	67 (32.1%)	141 (67.9%)		1.75 (1.15-2.71)*	-
Time since diagnosis						
>24 months	187 (40.7%)	36 (19.3%)	151 (80.7%)	**0.002**	Ref.	
≤24 months	273 (59.3%)	89 (32.6%)	184 (67.4%)		2.12 (1.31-3.15)*	2.84 (2.1-3.76)**

*p<0.05, **p<0.001, Ref.: Reference category for odds ratio comparisons, BMI: Body mass index, SLNB: Sentinel lymph node biopsy, ALND: Axillary lymph node dissection, BCS: Breast-conserving surgery, Mastectomy# includes both modified radical and bilateral mastectomy. COR: Crude odds ratio, AOR: Adjusted odds ratio, CI: Confidence interval. p-values are based on chi-square tests comparing the two groups. Hosmer- Lemeshow test: p: 0.076; Nagelkerke R2: 0.200, Overall classification accuracy: 77.2%

In the clinical model, significant predictors included premenopausal status (AOR: 3.05, 95% CI: 1.92-4.87), poor sleep quality (AOR: 2.42, 95% CI: 1.38-4.25), advanced-stage cancer (AOR: 1.85, 95% CI: 1.16-2.96), active chemotherapy (AOR: 2.09, 95% CI: 1.05-4.15), and shorter time since diagnosis (≤24 months) (AOR: 2.84, 95% CI: 2.10-3.76). Physical activity, BMI, and hormone therapy were significant only in univariate analysis. The model showed acceptable fit (Hosmer-Lemeshow p= 0.076), with a Nagelkerke R^2 of 0.200 and classification accuracy of 77.2%.

A visual summary of the multivariate logistic regression models is presented in Figure 1 for depression and Figure 2 for anxiety, based on AORs and 95% CIs for both sociodemographic and clinical predictors.

Reciprocal Association Between Depression and Anxiety

A significant bidirectional association was found between depression and anxiety. Among those with depression, 57.5% (65/113) also reported anxiety, compared to 17.9% (60/347) without depression. Conversely, 52.0% (65/125) of participants with anxiety showed depressive symptoms, versus 14.3% (48/335) without anxiety (p<0.001 for both).

Bidirectional logistic regression confirmed that the presence of anxiety significantly increased the odds of depression, and vice versa (OR: 6.48; 95% CI: 4.07–10.32; p<0.001). Both models showed a classification accuracy of 76.5% and a Nagelkerke R^2 of 0.195, indicating moderate explanatory power.

DISCUSSION

Anxiety is a prolonged state of alertness to perceived threats, while depression manifests as low mood, apathy, and reduced motivation[20]. In our study, 24.6% of the women with non-metastatic breast cancer showed depressive symptoms, and 27.2% experienced anxiety. These figures are consistent with European data (depression 20-35%, anxiety 25-40%), but rates are markedly higher in low- and middle-income countries up to 62.6% and 77.4% in Morocco, and 83% for depression in Pakistan[21,22]. This gap may result from limited healthcare access, cultural norms, lack of validated tools, and methodological variation. Stigma, fatalism, and weak social support further hinder help-seeking and exacerbate distress[21-25]. Altogether, these factors underscore the complex, multifactorial roots of psychological burden in underserved regions.

There was a notable overlap between depressive and anxiety symptoms, suggesting a bidirectional relationship. Patients experiencing one often reported the other. This co-occurrence may partly reflect shared symptom domains in assessment tools, potentially inflating comorbidity estimates. Still, some studies propose a linear progression typically with anxiety preceding depression[26,27]. Our lower depression rates may align with the Learned Helplessness Model, which posits that prolonged anxiety can evolve into depression over time[28]. Depression prevalence can reach 66.1% during the remission phase[29]. These findings highlight the need for continuous monitoring of symptoms and psychological support throughout cancer care. Furthermore, identifying high-risk groups remains essential within this symptom interaction framework.

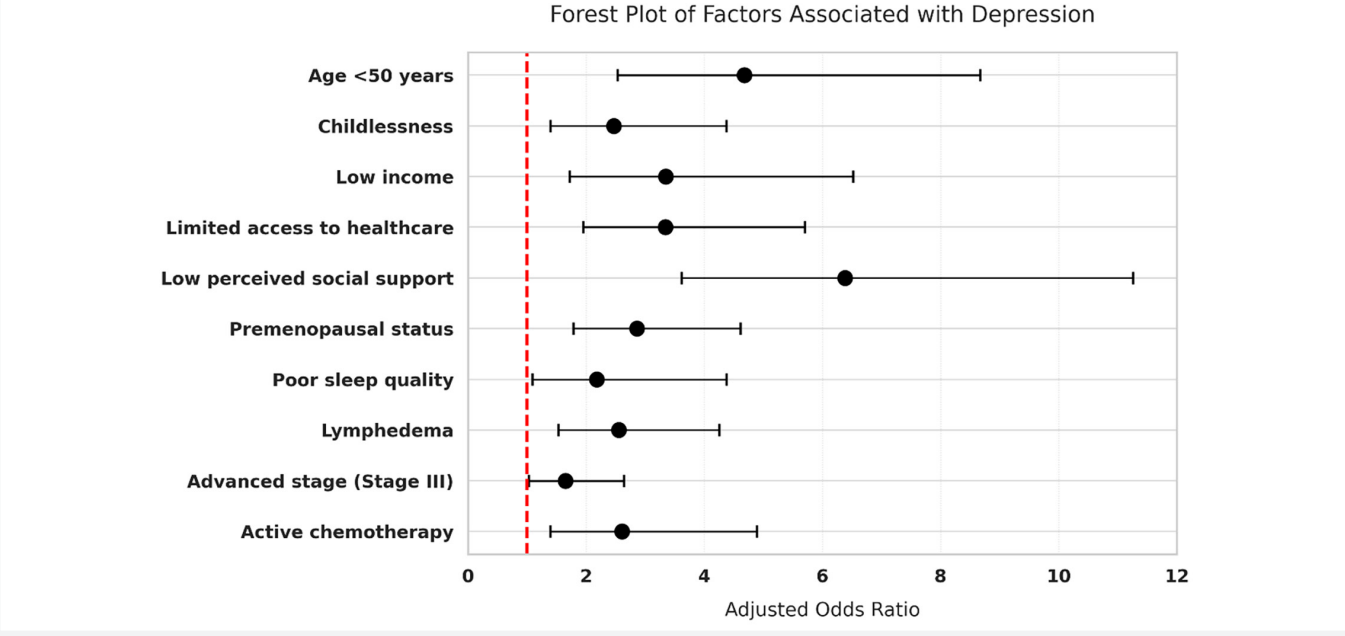


Figure 1. Forest plot of sociodemographic and clinical variables independently associated with depression

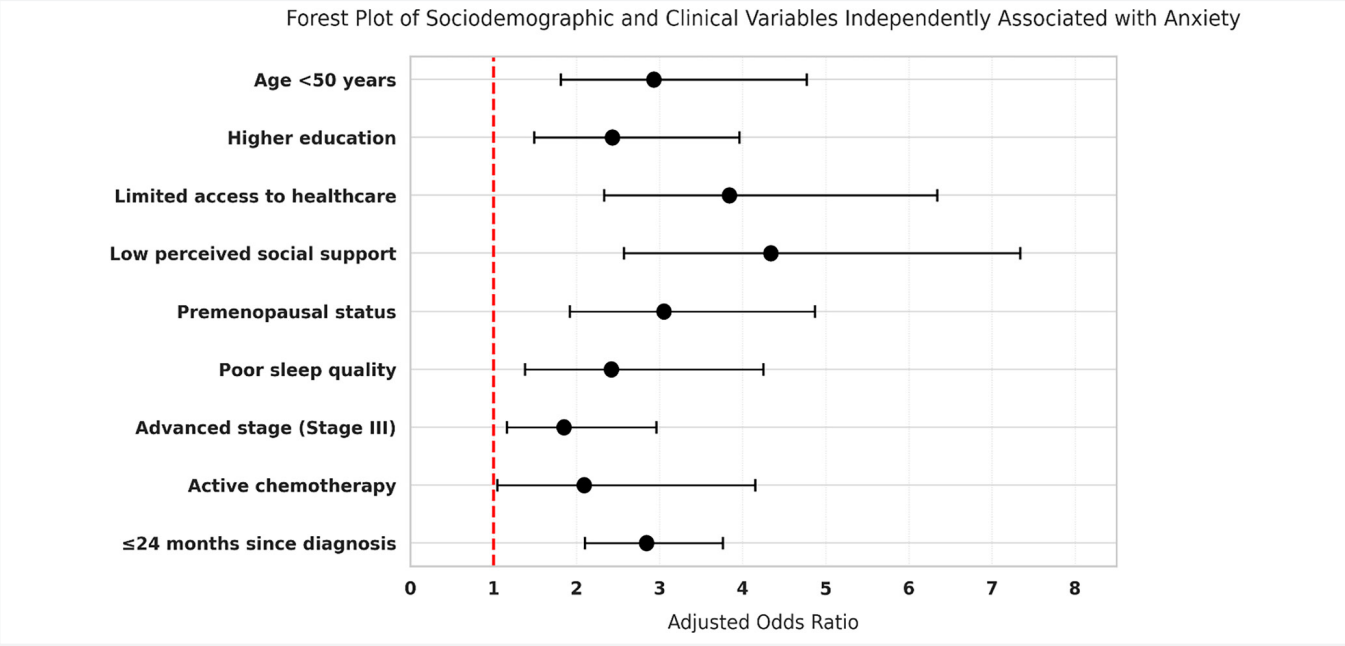


Figure 2. Forest plot of sociodemographic and clinical variables independently associated with anxiety

Depression and anxiety were significantly associated with sociodemographic factors, particularly younger age. Among the participants under 50 years of age, depression and anxiety rates were 34.3% and 40.4%, compared to 19.0% and 19.7% in older adults (p= 0.001; AOR: 4.68 and 2.93). This disparity may reflect greater psychological burden linked to fertility concerns, career disruption, and body image issues among younger women. Similar patterns have been reported in previous studies[30-32]. However, some studies suggest psychological symptoms increase with age, with depression risk rising 0.57% per year of age and 2.25-fold after 55[21,33]. These differences may be linked to factors like loneliness, chronic illness, or social isolation in older adults. Thus, age seems to be a context-sensitive rather than universal predictor of psychological vulnerability.

Perceived social support was the strongest psychosocial predictor of depression (46.8% vs. 17.5%, AOR: 6.38) and anxiety (44.1% vs. 21.8%, AOR: 4.34). These findings align with previous research emphasizing its protective role in psychological well-being[23,34,35]. Family structures often serve as key sources of support. In this context, childlessness emerged as a significant risk factor for depression (AOR: 2.47). Marital status was significant in univariate analysis (p= 0.005; COR: 1.97) but not multivariate, suggesting that relationship quality may matter more than marital status alone. Prior studies confirm this, showing that dysfunctional relationships can exacerbate distress and that parenthood does not always offer protection[22,36,37]. This is especially true when caregiving burdens are high, or when children are young, dependent, or emotionally impacted by the illness. Thus, social support should be evaluated not by the presence of family members alone, but by the emotional quality of those relationships, informing more personalized psychosocial care.

Socioeconomic status is a fundamental structural determinant of psychological symptoms. In our study, low income was significantly associated with depression (27.7% vs. 18.5%, AOR: 3.35), aligning with previous findings[8,22,23,38]. However, income was not significantly associated with anxiety in multivariate analysis, possibly due to confounding by factors like social support. Limited healthcare access strongly predicted both depression (33.8% vs. 19.8%, AOR: 3.34) and anxiety (41.4% vs. 19.8%, AOR: 3.84), consistent with prior research[39]. This suggests that reduced access heightens uncertainty and perceived loss of control, thereby intensifying psychological distress. Education level was unrelated to depression (p= 0.074) but unexpectedly associated with higher anxiety risk (35.6% vs. 22.0%, AOR: 2.43), contrary to studies suggesting a protective role[40]. This may reflect differences in awareness, expectations, and coping styles among highly educated individuals.

Several sociodemographic variables commonly associated with psychological symptoms in previous studies were not significant in our multivariate analysis. Employment status, though often discussed, may have variable effects depending

on job security, autonomy, and emotional demands all of which can influence psychological vulnerability[35,41]. Similarly, no significant link between residence and depression or anxiety was observed. While rural areas are commonly associated with higher psychological burden due to isolation and limited care access[39], this may not apply in settings where rural urban gaps are smaller. Family history of breast cancer also showed no significant association, although some studies have reported greater distress in such cases[42]. These findings emphasize that the psychological impact of sociodemographic factors is not universal but shaped by context, individual perception, and cultural norms. Emerging evidence also suggests that stressful life events may increase breast cancer risk, highlighting the broader role of psychosocial stressors in both emotional and biological processes[43].

However, sociodemographic factors alone do not fully explain psychological vulnerability. Clinical factors also significantly influence emotional outcomes. Premenopausal status was significantly associated with higher rates of depression (35.0% vs. 19.1%, p= 0.001; AOR: 2.86) and anxiety (41.4% vs. 19.8%, p= 0.001; AOR: 3.05). Hormonal fluctuations may contribute to mood vulnerability by affecting regulatory pathways, although psychosocial factors likely play a more prominent role[38,44]. This underscores the importance of early identification of younger patients experiencing greater emotional burden.

Sleep quality is closely associated with psychological symptoms. Poor sleep was significantly linked to both depression (35.8% vs. 18.2%, p= 0.002; AOR: 2.18) and anxiety (38.3% vs. 17.6%, p= 0.001; AOR: 2.42). Sleep disturbances may disrupt mood regulation via neurohormonal pathways. Literature describes a bidirectional relationship: disrupted sleep alters serotonin and cortisol, while depressive and anxious symptoms impair sleep initiation and maintenance[38,45]. In cancer care, fatigue, pain, and anticipatory anxiety may intensify this cycle[46]. A prospective study reported that depression, fatigue, and sleep disturbances often co-occur as a symptom cluster in breast cancer patients[47]. Routine screening of sleep quality may help identify at-risk patients before more severe symptoms develop.

Stage III patients showed higher rates of depression (31.4% vs. 20.9%, p: 0.017; AOR: 1.65) and anxiety (35.2% vs. 22.9%, p: 0.006; AOR: 1.85), even in non-metastatic cases. Another study found that Stage IV patients had nearly twice the depression risk of earlier stages (OR: 1.9, p: 0.003)[48]. These findings suggest that disease stage acts not only biologically but also as a psychological stressor, driven by uncertainty, intensive treatment, and prognosis concerns.

Receiving chemotherapy was significantly associated with depression (41.4% vs. 22.1%, p: 0.003; AOR: 2.61) and anxiety (53.4% vs. 23.4%, p: 0.001; AOR: 2.09). Side effects like hair loss, nausea, fatigue, early menopause, and neuropathy may lower quality of life and trigger depressive symptoms. Additionally, the treatment's cyclical nature and frequent hospital visits may reinforce the "patient role," heightening feelings of lost control. Our findings align with previous research showing chemotherapy's emotional burden extends beyond physical effects to include symbolic and psychological dimensions[38,49-51].

Surgical type showed no significant association with psychological symptoms (p>0.05), implying that postoperative complications especially lymphedema may be more influential[22]. Lymphedema was significantly associated with depression (35.8% vs. 21.1%, p: 0.003; AOR: 2.55), likely due to chronic pain, mobility limitations, and body image concerns, as previous studies suggest[52,53]. Many women avoid form-fitting clothing, which may undermine self-image and social confidence[54].

Anxiety was more common within the first two years post-diagnosis (32.6% vs. 19.3%, p: 0.002; AOR: 2.84, 95% CI: 2.10-3.76). This period may reflect a psychologically vulnerable window due to diagnostic shock, treatment adjustment, and abrupt lifestyle changes. Previous studies support this; for instance, in a five-year follow-up, anxiety peaked before treatment (38.0%) and fell to 25.3% by the first year's end[55]. The absence of a significant link between time since diagnosis and depression suggests symptoms may develop gradually, underscoring the need for sustained psychosocial monitoring.

Clinical factors like inactivity, high BMI, hormone therapy, and poor performance status were initially significant but lost relevance in multivariate analysis. Other variables type of surgery, family history, and residential setting showed no significant association. Yet literature indicates they may still affect depression and anxiety. At least 2.5 hours of weekly physical activity is associated with lower depression risk[22]. High BMI correlates with fatigue, poorer quality of life, and depression[56]. Aromatase inhibitors may increase depression risk by 27-41%[57]. Mastectomy without reconstruction is also tied to higher rates of depressive symptoms[58]. The complex, interrelated nature of psychological distress highlights the need to assess mental health factors within a broad multivariate framework.

Study Limitations

This study benefits from a large sample, multicenter design, and validated psychological measures. However, several limitations warrant consideration. Its cross-sectional design limits causal inference and captures only time-specific associations. Depression and anxiety were self-reported, introducing

potential biases like social desirability or recall error. Key psychosocial variables e.g., perceived support and healthcare access were measured with non-validated, study-specific tools, limiting comparability. Important domains like body image, sexual function, death anxiety, and post-traumatic stress were not assessed. Lastly, because participation was voluntary, individuals with higher psychological burden may have been underrepresented, which may affect the generalizability of the results.

CONCLUSION

This study shows that depression and anxiety are common even among non-metastatic breast cancer patients. Psychosocial and clinical factors like younger age, weak support, financial strain, and ongoing treatment worsen emotional distress. Psychological vulnerability often begins at diagnosis and deepens with uncertainty, isolation, and treatment stress. These findings call for a rethink of oncology models that prioritize tumor control but overlook mental well-being. Better outcomes demand a holistic approach with routine psychosocial screening and support. Addressing mental and physical health together may improve adherence, coping, and reduce disparities. This multicenter study not only informs future research but also urges urgent integration of psychosocial care into oncology where mental health is essential, not optional.

Ethics

Ethics Committee Approval: This study received ethics approval from the Non-Interventional Research Ethics Committee of Trakya University Faculty of Medicine (decision no: 01/14, date: 06.01.2025).

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

Acknowledgements: The authors would like to thank the participants.

Footnotes

Authorship Contributions

Concept: İ.G., S.T., Design: İ.G., Data Collection or Processing: İ.G., D.D., E.Ö., V.Ç., A.C., N.D., Analysis or Interpretation: İ.G., S.T., Literature Search: İ.G., Writing: İ.G.

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

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

REFERENCES

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74:229-63.
2. Kim J, Harper A, McCormack V, Sung H, Houssami N, Morgan E, et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries. Nat Med. 2025;31:1154-62.
3. Barrios CH. Global challenges in breast cancer detection and treatment. Breast. 2022;62(Suppl 1):S3-S6.
4. Baider L, Andritsch E, Uziely B, Ever-Hadani P, Goldzweig G, Hofmann G, et al. Do different cultural settings affect the psychological distress of women with breast cancer? A randomized study. Eur J Cancer Care (Engl). 2003;12:263-73.
5. İzci F, İlgün AS, Fındıklı E, Özmen V. Psychiatric symptoms and psychosocial problems in patients with breast cancer. J Breast Health. 2016;12:94-101.
6. Avis NE, Levine BJ, Case LD, Naftalis EZ, Van Zee KJ. Trajectories of depressive symptoms following breast cancer diagnosis. Cancer Epidemiol Biomarkers Prev. 2015;24:1789-95.
7. Zhu L, Ranchor AV, Helgeson VS, van der Lee M, Garssen B, Stewart RE, et al. Benefit finding trajectories in cancer patients receiving psychological care: predictors and relations to depressive and anxiety symptoms. Br J Health Psychol. 2018;23:238-52.
8. Alfonsson S, Olsson E, Hursti T, Lundh MH, Johansson B. Socio-demographic and clinical variables associated with psychological distress 1 and 3 years after breast cancer diagnosis. Support Care Cancer. 2016;24:4017-23.
9. Wei M, Guo L, Zhu Y, Guo Y, Yv S, Namassevayam G, et al. Type C personality and depression among newly diagnosed breast cancer patients: the mediating role of sense of coherence. Neuropsychiatr Dis Treat. 2019;15:3519-29.
10. Hammoudeh W, Hogan D, Giacaman R. From a death sentence to a disrupted life: Palestinian women's experiences and coping with breast cancer. Qual Health Res. 2017;27:487-96.
11. Jassim GA, Whitford DL. Understanding the experiences and quality of life issues of Bahraini women with breast cancer. Soc Sci Med. 2014;107:189-95.
12. Mausbach BT, Schwab RB, Irwin SA. Depression as a predictor of adherence to adjuvant endocrine therapy (AET) in women with breast cancer: a systematic review and meta-analysis. Breast Cancer Res Treat. 2015;152:239-46.
13. Vin-Raviv N, Akinyemiju TF, Galea S, Bovbjerg DH. Depression and anxiety disorders among hospitalized women with breast cancer. PLoS One. 2015;10:e0129169.
14. Zhang S, Jin Z, Bao L, Shu P. The global burden of breast cancer in women from 1990 to 2030: assessment and projection based on the global burden of disease study 2019. Front Oncol. 2024;14:1364397.
15. Günther MP, Kirchebner J, Schulze JB, Götz A, von Känel R, Euler S. Uncovering barriers to screening for distress in patients with cancer via machine learning. J Acad Consult Liaison Psychiatry. 2022;63:163-9.
16. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16:606-13.
17. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006;166:1092-7.
18. Sarı YE, Tamam L, Ekinci S, Özdel K, Başoğlu C. Hasta sağlık anketi-9'un (PHQ-9) Türkçe geçerlik ve güvenilirlik çalışması. Cukurova Med J. 2018;43:666-74.
19. Konkan R, Şenormancı Ö, Güçlü O, Aydın E, Sungur MZ. Validity and reliability study for the Turkish version of the generalized anxiety disorder-7 (GAD-7) scale. Arch Neuropsychiatry. 2013;50:53-8.

20. WHO. Depression. 2022; Available from: 1.https:// www. who. int/ healthtopics/depression.

21. Javan Biparva A, Raoofi S, Rafiei S, Masoumi M, Doustmehraban M, Bagheribayati F, et al. Global depression in breast cancer patients: systematic review and meta-analysis. PLoS One. 2023;18:e0287372.

22. Mahlaq S, Lahlou L, Rammouz I, Abouqal R, Belayachi J. Correction: factors associated with psychological burden of breast cancer in women in Morocco: crosssectional study. BMC Womens Health. 2024;24:61.

23. Soqia J, Al-Shafie M, Agha LY, Alameer MB, Alhomsi D, Saadoun R, et al. Depression, anxiety and related factors among Syrian breast cancer patients: a cross-sectional study. BMC Psychiatry. 2022;22:796.

24. Sadaqa D, Farraj A, Naseef H, Alsaid H, Al-Shami N, AbuKhalil AD. Risk of developing depression among breast cancer patients in Palestine. BMC Cancer. 2022;22:295.

25. Salem H, Daher-Nashif S. Psychosocial aspects of female breast cancer in the Middle East and North Africa. Int J Environ Res Public Health. 2020;17:6802.

26. Naser AY, Hameed AN, Mustafa N, Alwafi H, Dahmash EZ, Alyami HS, et al. Depression and anxiety in patients with cancer: a cross-sectional study. Front Psychol. 2021;12:585534.

27. Safaie N, Zeinali H, Ghahramanfard F, Mirmohammadkhani M, Moonesan M. Anxiety and depression among new cancer patients. J Family Med Prim Care. 2022;11:4146-50.

28. Else-Quest NM, LoConte NK, Schiller JH, Hyde JS. Perceived stigma, self-blame, and adjustment among lung, breast and prostate cancer patients. Psychol Health. 2009;24:949-64.

29. Fortin J, Leblanc M, Elgbeili G, Cordova MJ, Marin MF, Brunet A. The mental health impacts of receiving a breast cancer diagnosis: a meta-analysis. Br J Cancer. 2021;125:1582-92.

30. Alacacioglu A, Ulger E, Varol U, Yildiz I, Salman T, Bayoglu V, et al. Depression, anxiety and sexual satisfaction in breast cancer patients and their partners-Izmir oncology group study. Asian Pac J Cancer Prev. 2014;15:10631-6.

31. Omari B, Riash SA, Arfat A, Amer R, Mosleh S, Shawahna R. Assessing the factors associated with body image perception and quality of life of Palestinian women undergoing breast cancer treatment: a cross-sectional study. BMC Womens Health. 2024;24:565.

32. Vahdaninia M, Omidvari S, Montazeri A. What do predict anxiety and depression in breast cancer patients? A follow-up study. Soc Psychiatry Psychiatr Epidemiol. 2010;45:355-61.

33. Bennett B, Goldstein D, Lloyd A, Davenport T, Hickie I. Fatigue and psychological distress--exploring the relationship in women treated for breast cancer. Eur J Cancer. 2004;40:1689-95.

34. Civilotti C, Botto R, Maran DA, Leonardis B, Bianciotto B, Stanizzo MR. Anxiety and depression in women newly diagnosed with breast cancer and waiting for surgery: prevalence and associations with socio-demographic variables. Medicina (Kaunas). 2021;57:454.

35. Taskila T, Lindbohm ML, Martikainen R, Lehto US, Hakanen J, Hietanen P. Cancer survivors' received and needed social support from their work place and the occupational health services. Support Care Cancer. 2006;14:427-35.

36. Ding W, Ruan G, Lin Y, Zhu J, Tu C, Li Z. Dynamic changes in marital status and survival in women with breast cancer: a population-based study. Sci Rep. 2021;11:5421.

37. Krajc K, Miroševič Š, Sajovic J, Klemenc Ketiš Z, Spiegel D, Drevenšek G, et al. Marital status and survival in cancer patients: a systematic review and meta-analysis. Cancer Med. 2023;12:1685-708.

38. Karabulut Gul S, Tepetam H, Gursel OK, Alanyali S, Oruc AF, Tugrul F, et al. Investigating the levels of depression, anxiety, sexual disorders, and other influencing factors in breast cancer patients: Turkish radiation oncology integrative group study (12-05). Medicine (Baltimore). 2023;102:e35280.

39. Xu RH, Wang LL, Zhou LM, Wong EL, Wang D. Urban-rural differences in financial toxicity and its effect on cancer survivors' health-related quality of life and emotional status: a latent class analysis. Support Care Cancer. 2022;30:4219-29.

40. Habimana O, Mukeshimana V, Ahishakiye A, Makuza P, Hategekimana V, Muhayimana C, et al. standardization of education of patients with cancer in a low- and middle-income country: a quality improvement project using the cancer and you booklet. J Glob Oncol. 2019;5:1-6.

41. Han X, Li Q, Wang C, Li Y. The association of occupational stress and depressive symptoms among employed persons with benign breast disease: the mediating role of psychological capital. Psychopathology. 2019;52:205-11.

42. Liu Y, Cao C. The relationship between family history of cancer, coping style and psychological distress. Pak J Med Sci. 2014;30:507-10.

43. İriağaç Y, Çavdar E, Karaboyun K, Avcı O, Beyazyüz E, Şeber ES. The role of stressful life events in breast cancer etiology. Balıkesir Atatürk City Hospital Medical Journal. 2024;1:1-6.

44. Biddle M, Knox D. The role of estrogen receptor manipulation during traumatic stress on changes in emotional memory induced by traumatic stress. Psychopharmacology (Berl). 2023;240:1049-61.

45. Scott AJ, Webb TL, Martyn-St James M, Rowse G, Weich S. Improving sleep quality leads to better mental health: a meta-analysis of randomised controlled trials. Sleep Med Rev. 2021;60:101556.

46. Ritterband LM, Bailey ET, Thorndike FP, Lord HR, Farrell-Carnahan L, Baum LD. Initial evaluation of an internet intervention to improve the sleep of cancer survivors with insomnia. Psychooncology. 2012;21:695-705.

47. Bower JE, Ganz PA, Irwin MR, Arevalo JM, Cole SW. Fatigue and gene expression in human leukocytes: increased NF-κB and decreased glucocorticoid signaling in breast cancer survivors with persistent fatigue. Brain Behav Immun. 2011;25:147-50.

48. Tsaras K, Papathanasiou IV, Mitsi D, Veneti A, Kelesi M, Zyga S, et al. Assessment of depression and anxiety in breast cancer patients: prevalence and associated factors. Asian Pac J Cancer Prev. 2018;19:1661-9.

49. Gokce Ceylan G, Gok Metin Z. Symptom status, body perception, and risk of anxiety and depression in breast cancer patients receiving paclitaxel: a prospective longitudinal study. Support Care Cancer. 2022;30:2069-79.

50. Habimana S, Biracyaza E, Mpunga T, Nsabimana E, Kayitesi F, Nzamwita P, et al. Prevalence and associated factors of depression and anxiety among patients with cancer seeking treatment at the Butaro Cancer Center of Excellence in Rwanda. Front Public Health. 2023;11:972360.

51. Radhakrishnan R, Selvaraj H, Chidambaram K, Kv A, James A, Thangavel S. Prevalence of depression and anxiety disorders among cancer patients: an insight from a single institute. Cureus. 2023;15:e42831.

52. Yusof KM, Mohd Sidik S, Mahmud R, Abdullah M, Avery-Kiejda KA, Rosli R. Association of psychological distress with arm morbidity symptoms in breast cancer survivors: outcomes from the use of PHQ-9 and GAD-7 questionnaires. Breast Cancer. 2023;30:810-9.

53. Taghian NR, Miller CL, Jammallo LS, O'Toole J, Skolny MN. Lymphedema following breast cancer treatment and impact on quality of life: a review. Crit Rev Oncol Hematol. 2014;92:227-34.

54. Boing L, Pereira GS, Araújo CDCR, Sperandio FF, Loch MDSG, Bergmann A, et al. Factors associated with depression symptoms in women after breast cancer. Rev Saude Publica. 2019;53:30.

55. Lopes C, Lopes-Conceição L, Fontes F, Ferreira A, Pereira S, Lunet N, et al. Prevalence and persistence of anxiety and depression over five years since breast cancer diagnosis-The NEON-BC prospective study. Curr Oncol. 2022;29:2141-53.

56. Dorfman CS, Fisher HM, Thomas S, Kelleher SA, Winger JG, Mitchell NS, et al. Breast cancer survivors with pain: an examination of the relationships between body mass index, physical activity, and symptom burden. Support Care Cancer. 2023;31:604.

57. Choi HL, Jeong SM, Jeon KH, Kim B, Jung W, Jeong A, et al. Depression risk among breast cancer survivors: a nationwide cohort study in South Korea. Breast Cancer Res. 2024;26:188.

58. Garg SP, Chintalapati NV, Sandepudi K, Marzouk S, Ho KC, Ko JH, et al. Racial differences in incidence of anxiety and depression among mastectomy and breast reconstruction patients using the all of us database. Surgeries. 2024;5:986-96.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.09815
Nam Kem Med J 2025;13(3):246-252

Optimizing Blood Management in Total Knee Arthroplasty: An Analytical Study on the Impact of Preoperative and Intraoperative Factors

Total Diz Artroplastisinde Kan Yönetiminin Optimize Edilmesi: Ameliyat Öncesi ve Ameliyat Sırasındaki Faktörlerin Etkisine İlişkin Analitik Bir Çalışma

Murat TAŞCI[1], Mehmet Talha AYDIN[1], Mehmet Mete ORUÇ[2], Serdar Kamil ÇEPNİ[1]

[1]University of Health Sciences Türkiye, Ümraniye Training and Research Hospital, Clinic of Orthopedics and Traumatology, İstanbul, Türkiye
[2]Medeniyet University, Gözepe Prof. Dr. Süleyman Yalçın City Hospital, Department of Orthopedics and Traumatology, İstanbul, Türkiye

ABSTRACT

Aim: In surgeries like arthroplasty that involve joint resection, preoperative evaluations of the patient's comorbidities and medications are conducted to reduce blood loss. This study aimed to address factors that can influence the need for blood transfusion due to significant blood loss, in conjunction with preventative measures applied before and after surgery.

Materials and Methods: Patients who underwent knee arthroplasty from 2017 to 2022, and for whom necessary follow-up parameters were available, were included in the study. Patients with incomplete data, bleeding diathesis, or preoperative albumin, platelet, prothrombin time, and international normalized ratio levels outside the normal range were excluded.

Results: The study included a total of 479 patients. A positive and very weak significant correlation was found between body mass index and intraoperative blood loss. Patients with diabetes showed significantly higher intraoperative blood loss, drainage volume, total blood loss, and the difference in preoperative and postoperative hemoglobin compared to those without diabetes. There were no statistically significant differences in intraoperative blood loss, drainage volume, or total blood loss between patients with and without cardiovascular disease. Patients not receiving tranexamic acid (TXA) had higher intraoperative blood loss compared to those receiving intravenous (IV) TXA and lower compared to those receiving local TXA. The IV group had lower intraoperative blood loss compared to the local group.

Conclusion: In conclusion, patient-related factors and pre- and post-operative measures are crucial in managing blood loss in knee arthroplasty. The goal should be to eliminate the need for transfusions and thus prevent possible complications associated with blood transfusions.

Keywords: Total knee arthroplasty, blood transfusion, blood loss, comorbidities

ÖZ

Amaç: Eklem rezeksiyonu içeren artroplasti gibi ameliyatlarda, kan kaybını azaltmak için hastanın eşlik eden hastalıkları ve ilaçları hakkında ameliyat öncesi değerlendirmeler yapılır. Bu çalışma, ameliyattan önce ve sonra uygulanan önleyici tedbirlerle birlikte, önemli kan kaybı nedeniyle kan transfüzyonu ihtiyacını etkileyebilecek faktörleri ele almayı amaçlamaktadır.

Gereç ve Yöntem: 2017-2022 yılları arasında diz artroplastisi geçiren ve gerekli takip parametreleri mevcut olan hastalar çalışmaya dahil edildi. Eksik verileri olan, kanama diatezi olan veya ameliyat öncesi albümin, trombosit, protrombin zamanı ve uluslararası normalize oranı seviyeleri normal aralığın dışında olan hastalar hariç tutuldu.

Address for Correspondence: Murat TAŞCI MD, University of Health Sciences Türkiye, Ümraniye Training and Research Hospital, Clinic of Orthopedics and Traumatology, İstanbul, Türkiye
E-mail: drtascimurat@gmail.com **ORCID ID:** orcid.org/0000-0001-5391-7662
Received: 10.03.2025 **Accepted:** 11.05.2025 **Publication Date:** 07.10.2025

Cite this article as: Taşcı M, Aydın MT, Oruç MM, Çepni SK. Optimizing blood management in total knee arthroplasty: an analytical study on the impact of preoperative and intraoperative factors. Nam Kem Med J. 2025;13(3):246-252

Bulgular: Çalışmaya toplam 479 hasta dahil edildi. Vücut kitle indeksi ile ameliyat sırasındaki kan kaybı arasında pozitif ve çok zayıf anlamlı bir korelasyon bulundu. Diyabetli hastalarda, diyabeti olmayanlara kıyasla anlamlı derecede daha yüksek ameliyat sırasındaki kan kaybı, drenaj hacmi, toplam kan kaybı ve ameliyat öncesi ve sonrası hemoglobin farkı görüldü. Kardiyovasküler hastalığı olan ve olmayan hastalar arasında intraoperatif kan kaybı, drenaj hacmi veya toplam kan kaybında istatistiksel olarak anlamlı bir fark görülmedi. Traneksamik asit (TXA) almayan hastalarda intravenöz (IV) TXA alanlara kıyasla daha yüksek, lokal TXA alanlara kıyasla daha düşük intraoperatif kan kaybı görüldü. IV grupta lokal gruba kıyasla daha düşük intraoperatif kan kaybı görüldü.

Sonuç: Sonuç olarak, diz artroplastisinde kan kaybını yönetmede hasta ile ilgili faktörler ve ameliyat öncesi ve sonrası önlemler çok önemlidir. Amaç, kan transfüzyonu ihtiyacını ortadan kaldırmak ve böylece kan transfüzyonuyla ilişkili olası komplikasyonları önlemek olmalıdır.

Anahtar Kelimeler: Total diz protezi, kan transfüzyonu, kan kaybı, eşlik eden hastalıklar

INTRODUCTION

Total knee arthroplasty (TKA) is commonly performed for advanced-stage knee osteoarthritis. Following TKA, patients may encounter postoperative complications including superficial or deep infections at the surgical site, hematoma formation, deep vein thrombosis, or significant bleeding that necessitates transfusion[1]. In surgeries like arthroplasty that involve joint resection, preoperative evaluations of the patient's comorbidities and medications are conducted to reduce blood loss. Various methods can also be applied intraoperatively to minimize blood loss. These methods often include the use of tourniquets, and local or intravenous (IV) administration of tranexamic acid (TXA)[2,3]. In knee replacement surgeries, bleeding generally occurs postoperatively. The amount of blood loss can be assessed intraoperatively as well as postoperatively by measuring the amount drained[4]. Hemoglobin (HGB) levels are monitored, and the need for transfusion is determined based on the drop in HGB levels. There are certain risks associated with transfusions after significant blood loss. According to a study by Politis et al.[5] in 2022 using data from the International Haemovigilance Network Database, 25% of reported adverse reactions (ARs) were severe, resulting in 368 deaths. Among the 284 cases of transfusion-transmitted infections, 187 were bacterial in nature, 84 were viral, and 13 stemmed from parasites or fungi, resulting in nine fatalities. ARs related to the respiratory system, including transfusion-associated circulatory overload, transfusion-related acute lung injury, and transfusion-associated dyspnea, accounted for 8.3% of all ARs, 20.1% of severe incidents, and 52.2% of mortality cases[5]. Other studies have also reported transmission of infectious diseases such as Hepatitis C virus or human immunodeficiency virus and acute reactions related to transfusions[6,7]. This study aimed to address factors that can influence the need for blood transfusion due to significant blood loss, in conjunction with preventative measures applied before and after surgery.

MATERIALS AND METHODS

The study received ethics approval from the Ethics Committee of the University of Health Sciences Türkiye, Ümraniye Training and Research Hospital (decision no: B.10.TKH.4.34.H.GP.0.01/332, date: 17.10.2024). Patients who underwent knee arthroplasty from 2017 to 2022, had a preoperative HGB level above 10, and for whom necessary follow-up parameters were available, were included in the study. In our clinic, in accordance with the recommendation of the hematology department, a preoperative HGB level of ≥10 g/dL is required prior to major surgeries such as TKA, in order to minimize complications associated with blood transfusion.

Patients with incomplete data, bleeding diathesis, or preoperative albumin, platelet, prothrombin time, and international normalized ratio levels outside the normal range were excluded. Demographic variables including age and sex, along with the presence of comorbidities such as hypertension, diabetes, and cardiovascular disease were assessed. Additionally, factors such as the use and duration of a tourniquet, administration of TXA (both locally and IV), intraoperative blood loss, volume of postoperative drainage, necessity for transfusion, and values of preoperative and postoperative HGB and hematocrit (HCT) were examined. Intraoperative blood loss was estimated by combining suction volume and weighed swabs, adjusted for irrigation fluid.

All data were obtained from the hospital's Health Information System. Acetylsalicylic acid (ASA) medication was discontinued 5 days prior to surgery for patients using it. HGB levels were routinely checked on the first and second postoperative days. All patients were monitored for at least two days postoperatively until wound and HGB levels normalized. In our clinic, if the postoperative HGB level is ≤8 g/dL or if the patient exhibits clinical signs of anemia, blood transfusion is planned in accordance with the transfusion protocols recommended by the hematology department. This HGB threshold is also supported by the literature[8,9]. Patients requiring more than one unit of red blood cells (RBCs) also received one unit of fresh frozen plasma. Routine deep vein thrombosis prophylaxis was applied to all patients postoperatively and continued until discharge. Clexane 0.4 mL subcutaneous was routinely administered as prophylaxis, along with ankle exercise for mechanical prophylaxis starting from the first day. A mobilization and

quadriceps exercise program was implemented on the second postoperative day. Patients who developed wound discharge in the form of serous fat necrosis from clexane were switched to 100 mg ASA. Drains were removed routinely when the output fell below 50 cc. Drain output was recorded before each emptying and was always maintained under negative pressure.

Statistical Analysis

Statistical analysis were conducted using the Number Cruncher Statistical System 2007 software (Kaysville, Utah, USA). Descriptive statistics including mean, standard deviation, median, frequency, ratio, minimum, and maximum were employed to analyze the study data. The distribution of data was evaluated using the Shapiro-Wilk test. The Kruskal-Wallis test was employed to compare variables across three or more independent groups, such as IV, local, or no use of TXA. For comparisons between two independent groups, such as the presence or absence of comorbidities, the Mann-Whitney U test was utilized. The Spearman's correlation analysis was conducted to assess relationships between continuous variables, including HGB levels and blood loss. Multiple linear regression analysis was performed to identify the factors influencing the difference between preoperative and postoperative HGB values. Additionally, ROC curve analysis was used to determine cut-off values for quantitative variables, such as HGB levels associated with an increased risk of transfusion.

Significance levels were set at $p<0.01$ and $p<0.05$.

RESULTS

The study included a total of 479 patients. Mean age was 66.5±7.6, with preoperative HGB levels averaging 11.46±1.14 and postoperative levels at 9.59±1.35. Hypertension was present in 45.9% of the patients, diabetes in 39.7%, and cardiovascular diseases in 42%. Mean body mass index (BMI) was found to be 38.1±8.5. All mean values are shown in Table 1. A positive and very weak significant correlation was found between BMI and intraoperative blood loss ($r=0.118$, $p<0.05$). However, no statistically significant correlation was found between BMI and postoperative HGB, postoperative HCT, operation duration, drainage amount, total blood loss, and the difference in HGB levels before and after surgery ($p>0.05$).

Tourniquet duration was negatively and weakly correlated with intraoperative blood loss ($r=-0.268$, $p<0.01$), amount from drainage ($r=-0.122$, $p<0.01$), and total blood loss ($r=-0.205$, $p<0.01$). There was a positive and weak significant relationship between tourniquet duration and the difference in preoperative and postoperative HCT ($r=0.247$, $p<0.01$) and HGB ($r=0.170$, $p<0.01$).

A very high and positive significant correlation was observed between intraoperative blood loss and total blood loss ($r=0.927$, $p<0.01$), while there was a positive and very weak significant correlation with the difference in preoperative and postoperative HGB ($r=0.164$, $p<0.01$).

Similarly, while there was a very high and positive significant correlation between the amount from drainage and total blood loss ($r=0.931$, $p<0.01$), there was a positive and very weak significant correlation with the difference in preoperative and postoperative HGB ($r=0.147$, $p<0.01$).

No statistically significant association was observed between the duration of the operation and variables such as intraoperative blood loss, drainage volume, total blood loss, and the differences in preoperative and postoperative HCT and HGB levels ($p>0.05$).

The presence of hypertension was not associated with statistically significant differences in intraoperative blood loss, the difference between preoperative and postoperative HGB and HCT levels, or total blood loss ($p>0.05$). However, patients with hypertension exhibited a significantly lower drainage volume compared to those without hypertension ($p=0.001$, $p<0.05$).

Patients with diabetes showed significantly higher intraoperative blood loss, drainage volume, total blood loss, and the difference in preoperative and postoperative HGB compared to those without diabetes ($p=0.001$, $p<0.05$).

No statistically significant differences were observed in intraoperative blood loss, drainage volume, or total blood loss between patients with and without cardiovascular disease ($p>0.05$). However, the difference between preoperative and postoperative HGB and HCT levels was significantly greater

Table 1. Average values

Parameter	Mean ± SD	Minimum-maximum (median)
Age	66.53±7.69	18-87 (67)
Body mass index	38.1±2.85	30-46 (38)
Preop HGB	11.46±1.14	9-15 (11.1)
Postop HGB	9.59±1.35	6.4-13.4 (9.7)
Preop HCT	35.11±5.16	19.2-44 (35.5)
Postop HCT	28.98±4	0-38.5 (29)
Tourniquet time	43.43±43.7	0-120 (60)
Operation time	107.34±30.41	30-180 (105)
Intraoperative blood loss	367.12±177.53	150-800 (400)
Blood from drain	241.65±148.54	50-600 (200)
Total blood loss	608.77±297.51	200-1300 (600)
Preop-postop HCT difference	6.44±3.94	-3.2-24 (6.4)
Preop-postop HGB difference	1.87±1.08	-2-9 (1.8)

HGB: Hemoglobin, HCT: Hematocrit, SD: Standard deviation

in patients with cardiovascular disease compared to those without (p=0.001, p<0.05).

In our study, TXA was administered IV to 276 patients, topically to 73 patients, and 130 patients did not receive TXA (Table 2).

Regarding the use of TXA, there were significant differences in intraoperative blood loss (p=0.001, p<0.05). Patients not receiving TXA had higher intraoperative blood loss compared to those receiving IV TXA (p=0.001, p<0.05) and lower compared to those receiving local TXA (p=0.001, p<0.05). The IV group had lower intraoperative blood loss compared to the local group (p=0.001, p<0.05). The IV group also had lower drainage volumes compared to the local group (p=0.001, p<0.05). Total blood loss was higher in the non-TXA group compared to the IV group (p=0.001, p<0.05) and lower compared to the local group (p=0.001, p<0.05). Total blood loss was also lower in the IV group compared to the local group (p=0.001, p<0.05).

The multiple linear regression analysis conducted to assess the impact of independent variables on the difference between preoperative and postoperative HGB levels yielded statistically significant results (F: 40.790, p<0.001). There was a positive and moderate significant relationship between independent variables and the preoperative-postoperative HGB difference (R: 0.640, p<0.001). The independent variables in the model explained 41% of the total variance of the preoperative-postoperative HGB difference (p<0.01).

Upon examining the regression coefficients, it was observed that diabetes (β: 0.088; p<0.001) and local TXA application (β: 0.132; p<0.001) had a positive effect on the preoperative-postoperative HGB difference, while transfusion (β: -0.138; p<0.001) had a negative and significant impact. Consequently, it was seen that patients with diabetes and those who received local TXA had a higher preoperative-postoperative HGB difference compared to those who did not.

The ROC analysis identified a reliable cutoff point for the preoperative-postoperative HGB difference at 1.6, with a sensitivity of 93.5% and specificity of 55.9% (Table 3; Figure 1). "A HGB difference >1.6 g/dL predicted transfusion need with 93.5% sensitivity".

DISCUSSION

The characteristics of patients, such as preoperative comorbidities, BMI, or preoperative HGB and HCT levels, influence total blood loss and the need for transfusion. It has been shown in prospective, retrospective, and clinical trial studies that the preoperative HGB value is an indicator of postoperative transfusion needs[10-12]. There is no consensus on the threshold values for HGB and HCT. The importance of a patient's age, weight, and comorbidities varies across studies. In our study, we analyzed in detail the relationship between preoperative factors and perioperative practices aimed at reducing blood loss with blood loss indicators.

Table 2. Distribution of patients' comorbidities and TXA usage

		n	%
Hypertension status	Yes	220	45.9
	No	259	54.1
Diabetes status	Yes	190	39.7
	No	289	60.3
Cardiovascular comorbidity	Yes	201	42.0
	No	278	58.0
Tranexamic acid usage	No	130	27.1
	IV	276	57.6
	Local	73	15.2
Transfusion status	Yes	123	25.7
	No	356	74.3

TXA: Tranexamic acid

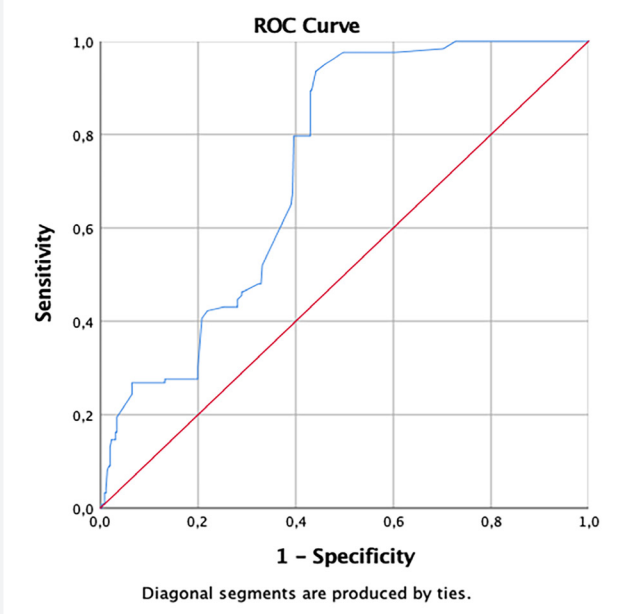


Figure 1. Preop-postop HGB difference: ROC analysis

HGB: Hemoglobin

Table 3. ROC analysis result: cut-off and AUC value

Parameter	Sensitivity (%)	Specificity (%)	Cut-off point	Area under curve
Preop-postop HGB difference	93.5%	55.9%	1.6	0.732

HGB: Hemoglobin, AUC: Area under the curve

In our study, no statistically significant associations were observed between BMI and various postoperative outcomes, including HGB levels, HCT levels, operation duration, blood loss from drainage, total blood loss, and the difference between preoperative and postoperative HGB levels ($p>0.05$). We demonstrated that patients with a higher BMI have higher intraoperative blood loss ($r= 0.118$, $p<0.05$) but this did not affect the total blood loss ($p>0.05$). Frisch et al.[13] in their study mentioned that patients with high BMI lose a smaller percentage of their total blood volume due to their increased blood losses and thus have lower transfusion rates. Bashaireh et al.[14] also showed that high BMI is not a risk factor for postoperative blood transfusion. Additionally, low body weight is associated with a smaller RBC volume, making it more difficult for lighter patients to compensate for blood loss[15]. In our study, consistent with the findings in the literature, the requirement for transfusion was not influenced by high BMI in patients. This observation is attributed to the higher blood volume in patients with elevated BMI, which allows them to tolerate greater intraoperative blood loss without necessitating transfusions.

Working under a tourniquet has, as expected, shown positive effects in reducing perioperative blood loss, as well as decreasing the amount from drainage and total blood loss. Studies, like those by Alcelik et al.[16], have shown that using a tourniquet reduces both intraoperative and total blood loss . Another study with a small patient group by Tan et al.[17] compared groups with and without a tourniquet and found no significant differences were observed in drainage volume, total blood loss, postoperative HGB levels, or the requirements for blood transfusion between the two groups.

Our study found that increases in intraoperative blood loss and the amount from drainage led to increases in both total blood loss and the difference between preoperative and postoperative HGB levels. However, the duration of the operation did not create a statistically significant difference in perioperative and postoperative parameters related to blood loss.

While no significant differences were noted in blood loss measures among hypertensive patients, they did exhibit significantly lower drainage volumes ($p=0.001$). This could suggest a physiological adaptation in hypertensive individuals that minimizes visible bleeding, potentially due to vascular changes such as increased arterial stiffness. Alternatively, it may reflect more conservative surgical drainage practices in patients with hypertension, due to concerns over their altered hemodynamic responses. These findings highlight the need for tailored surgical management strategies for hypertensive patients and warrant further investigation to better understand the interplay between hypertension and surgical outcomes.

Berenholtz et al.[18] identified chronic complications of diabetes as an independent risk factor for blood transfusions in their study. Similarly, Slover et al.[19] study with a large patient series demonstrated that accompanying comorbidities increase the need for blood transfusions. A systematic review by Barr et al.[20] also indicated that diabetes and cardiovascular diseases were among the comorbidities that increased the need for transfusions. While specific comorbidities may vary, they are generally associated with a decreased capacity to tolerate anemia While no significant differences were noted in blood loss measures among hypertensive patients, they did exhibit significantly lower drainage volumes ($p=0.001$). This could suggest a physiological adaptation in hypertensive individuals that minimizes visible bleeding, potentially due to vascular changes such as increased arterial stiffness. Alternatively, it may reflect more conservative surgical drainage practices in patients with hypertension, due to concerns over their altered hemodynamic responses. These findings highlight the need for tailored surgical management strategies for hypertensive patients and warrant further investigation to better understand the interplay between hypertension and surgical outcomes.

While total blood loss did not differ significantly between patients with and without cardiovascular disease, those with cardiovascular conditions experienced a more pronounced drop in HGB and HCT levels. This discrepancy may be attributed to several factors. First, patients with cardiovascular disease often have compromised cardiac output and tissue perfusion, which can exacerbate the physiological impact of any blood loss, making them more susceptible to hemodynamic fluctuations. Additionally, these patients might have pre-existing anemia or be on medications like anticoagulants that could amplify the impact of blood loss on HGB and HCT levels despite similar volumes of blood loss. This observation underscores the need for cautious perioperative management and possibly more aggressive postoperative monitoring and intervention in patients with underlying cardiovascular disease.

Notably, in patients with diabetes, statistically significant results were noted, including increased intraoperative blood loss, drainage volume, total blood loss, and a heightened difference in HGB levels from preoperative to postoperative measurements.

Some surgeons prefer to use TXA IV or topically during TKA, while others do not use it at all. The use of TXA in controlling perioperative and postoperative bleeding has been shown to be safe and effective[21]. TXA can be administered using various methods including IV, topical, and oral applications[22]. Furthermore, studies have shown that it reduces postoperative HGB drop and the need for transfusions[23]. Fillingham et al.[24] have demonstrated that TXA is superior to placebo in reducing

blood loss and the risk of transfusion. Recent studies advocate for the routine use of IV TXA in arthroplasty surgeries due to its effectiveness in reducing bleeding. Our study observed that patients who did not receive TXA had higher intraoperative blood loss compared to the IV group, but lower than those who received topical applications. The volume of blood drained was lower in the IV group compared to the topical group. Additionally, total blood loss was greater in the topical group than in the IV group, yet lower in patients who did not receive TXA compared to those who underwent topical applications. The observed discrepancy, where higher blood loss is associated with local TXA, may reflect inconsistencies in application or insufficient dosages compared to IV administration.

Drainage use is also quite common to reduce the need for transfusions and the incidence of bleeding[25]. The American Academy of Orthopedic Surgeons (AAOS) recommends the use of drains in arthroplasty, stating that there is no difference in complications or outcomes between patients with or without drains[26]. However, some studies have indicated that there may be certain drawbacks associated with a decrease in HGB levels. Our study found a highly significant positive relationship between the amount of drainage and total blood loss ($r= 0.931$, $p<0.01$). Yet, the implication of this relationship in clinical practice is unclear, aligning with the AAOS's ambivalence towards the utility of drains in reducing transfusion needs. There is also a very weak but statistically significant positive relationship between the preoperative and postoperative HGB difference ($r=0.147$, $p<0.01$). Madan et al[23]. noted that the amount drained varied from 0-700 mL; there was no statistical significance between the amount drained and either HGB drop or the need for blood transfusion (p-value=0.401). However, they observed a significant increase in HGB reduction and the requirement for blood transfusions associated with drainage. Watanabe et al.[27] findings further revealed that the average decrease in HGB was more pronounced in the drainage group one day post-surgery. Conversely, other research has presented contrary evidence, indicating that there were no statistically significant differences in HGB levels or the frequency of blood transfusions during the initial six weeks post-operation[28]. This highlights the complexity of the effects of drainage on postoperative outcomes and emphasizes the need for further research to clarify its efficacy and application in arthroplasty.

Study Limitations

We acknowledge certain limitations in our study. The relatively small number of patients and the retrospective nature of data collection are among these limitations. We believe that prospective randomized controlled trials with larger patient populations would be useful in identifying factors that affect blood loss and the need for transfusions. Additional limitations of this study include its single-center design, potential variability in surgical techniques among different surgeons, and the absence of detailed data regarding the dosage and timing of TXA administration.

CONCLUSION

In conclusion, patient-related factors and pre- and post-operative measures are crucial in managing blood loss in knee arthroplasty. The goal should be to eliminate the need for transfusions and thus prevent possible complications associated with blood transfusions. To achieve this, strategies involving multiple measures should be developed and implemented. Additionally, patients with chronic preoperative anemia should have their anemia managed before planning surgical treatment.

Ethics

Ethics Committee Approval: The study received ethics approval from the Ethics Committee of the University of Health Sciences Türkiye, Ümraniye Training and Research Hospital (decision no: B.10.TKH.4.34.H.GP.0.01/332, date: 17.10.2024).

Informed Consent: It has been shown in prospective, retrospective, and clinical trial studies that the preoperative HGB value is an indicator of postoperative transfusion needs.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Yohe N, Funk A, Ciminero M, Erez O, Saleh A. Complications and readmissions after total knee replacement in octogenarians and nonagenarians. Geriatr Orthop Surg Rehabil. 2018;9:2151459318804113.
2. Cai DF, Fan QH, Zhong HH, Peng S, Song H. The effects of tourniquet use on blood loss in primary total knee arthroplasty for patients with osteoarthritis: a meta-analysis. J Orthop Surg Res. 2019;14:348.
3. Sadigursky D, Araujo LM, Fernandes RJC. Efficacy of tranexamic acid in reducing blood loss ın total knee arthroplasty. Acta Ortop Bras. 2018;26:63-6.
4. Tucci G, Amorese V, Romanini E. Closed suction drainage after orthopaedic surgery: evidence versus practice. Journal of Orthopaedics and Traumatology. 2006;7:29-32.
5. Politis C, Wiersum-Osselton J, Richardson C, Grouzi E, Sandid I, Marano G, et al. Adverse reactions following transfusion of blood components, with a focus on some rare reactions: reports to the International Haemovigilance Network Database (ISTARE) in 2012-2016. Transfus Clin Biol. 2022;29:243-9.

6. Dean CL, Wade J, Roback JD. Transfusion-transmitted infections: an update on product screening, diagnostic techniques, and the path ahead. J Clin Microbiol. 2018;56:e00352-18.

7. Hatayama Y, Matsumoto S, Hamada E, Kojima N, Hara A, Hino N, et al. Analysis of acute transfusion reactions and their occurrence times. Yonago Acta Med. 2018;61:87-90.

8. Chaudhry YP, MacMahon A, Rao SS, Mekkawy KL, Toci GR, Oni JK, et al. Predictors and outcomes of postoperative hemoglobin of <8 g/dL in total joint arthroplasty. J Bone Joint Surg Am. 2022;104:166-71.

9. Lee QJ, Mak WP, Yeung ST, Wong YC, Wai YL. Blood management protocol for total knee arthroplasty to reduce blood wastage and unnecessary transfusion. J Orthop Surg (Hong Kong). 2015;23:66-70.

10. Evans S, O'Loughlin E, Bruce J. Retrospective audit of blood transfusion and comparison with haemoglobin concentration in patients undergoing elective primary and revision lower limb arthroplasty. Anaesth Intensive Care. 2011;39:480-5.

11. Guerin S, Collins C, Kapoor H, McClean I, Collins D. Blood transfusion requirement prediction in patients undergoing primary total hip and knee arthroplasty. Transfus Med. 2007;17:37-43.

12. Mesa-Ramos F, Mesa-Ramos M, Maquieira-Canosa C, Carpintero P. Predictors for blood transfusion following total knee arthroplasty: a prospective randomised study. Acta Orthop Belg. 2008;74:83-9.

13. Frisch N, Wessell NM, Charters M, Peterson E, Cann B, Greenstein A, et al. Effect of body mass index on blood transfusion in total hip and knee arthroplasty. Orthopedics. 2016;39:e844-9.

14. Bashaireh K, Aljararhih O, Alawneh K. Impact of body mass index on hemoglobin level and blood transfusion in total knee arthroplasty: a retrospective case control study. Ann Med Surg (Lond). 2020;55:180-4.

15. Khanna MP, Hébert PC, Fergusson DA. Review of the clinical practice literature on patient characteristics associated with perioperative allogeneic red blood cell transfusion. Transfus Med Rev. 2003;17:110-9.

16. Alcelik I, Pollock RD, Sukeik M, Bettany-Saltikov J, Armstrong PM, Fismer P. A comparison of outcomes with and without a tourniquet in total knee arthroplasty: a systematic review and meta-analysis of randomized controlled trials. J Arthroplasty. 2012;27:331-40.

17. Tan Y, Guo S, Wang H, Tie K, Qin J, Zhao X, et al. The effects of tourniquet use on blood loss and perioperative complications in total knee arthroplasty. BMC Musculoskelet Disord. 2023;24:847.

18. Berenholtz SM, Pronovost PJ, Mullany D, Garrett E, Ness PM, Dorman T, et al. Predictors of transfusion for spinal surgery in Maryland, 1997 to 2000. Transfusion. 2002;42:183-9.

19. Slover J, Lavery JA, Schwarzkopf R, Iorio R, Bosco J, Gold HT. Incidence and risk factors for blood transfusion in total joint arthroplasty: analysis of a statewide database. J Arthroplasty. 2017;32:2684-7.

20. Barr PJ, Donnelly M, Cardwell C, Alam SS, Morris K, Parker M, et al. Drivers of transfusion decision making and quality of the evidence in orthopedic surgery: a systematic review of the literature. Transfus Med Rev. 2011;25:304-16.e1-6.

21. Lin ZX, Woolf SK. Safety, Efficacy, and cost-effectiveness of tranexamic acid in orthopedic surgery. Orthopedics. 2016;39:119-30.

22. Cankaya D, Dasar U, Satilmis AB, Basaran SH, Akkaya M, Bozkurt M. The combined use of oral and topical tranexamic acid is a safe, efficient and low-cost method in reducing blood loss and transfusion rates in total knee arthroplasty. J Orthop Surg (Hong Kong). 2017;25:2309499016684725.

23. Madan FH, Khamis E, Alhassan MA, Alrashid M, Saleh A, Rahma M. Hemoglobin drop and the need for transfusion in primary knee arthroplasty. Cureus. 2022;14:e27659.

24. Fillingham YA, Ramkumar DB, Jevsevar DS, Yates AJ, Shores P, Mullen K, et al. The efficacy of tranexamic acid in total knee arthroplasty: a network meta-analysis. J Arthroplasty. 2018;33:3090-8.e1.

25. Al-Zahid S, Davies AP. Closed suction drains, reinfusion drains or no drains in primary total knee replacement? Ann R Coll Surg Engl. 2012;94:347-50.

26. Strong evidence supports not using a drain with total knee arthroplasty (TKA) because there is no difference in complications or outcomes. (2014)." ifadesi, AAOS Clinical Practice Guideline: Surgical Management of Osteoarthritis of the Knee. Available from: https://www.orthoguidelines.org/

27. Watanabe T, Muneta T, Yagishita K, Hara K, Koga H, Sekiya I. Closed suction drainage is not necessary for total knee arthroplasty: a prospective study on simultaneous bilateral surgeries of a mean follow-up of 5.5 years. J Arthroplasty. 2016;31:641-5.

28. Yin D, Delisle J, Banica A, Senay A, Ranger P, Laflamme GY, et al. Tourniquet and closed-suction drains in total knee arthroplasty. No beneficial effects on bleeding management and knee function at a higher cost. Orthop Traumatol Surg Res. 2017;103:583-9.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.71463
Nam Kem Med J 2025;13(3):253-260

Quantitative Evaluation of Hepatic Microstructural Changes in Hepatocellular Carcinoma Using Diffusion Tensor Imaging

Difüzyon Tensör Görüntüleme ile Hepatoselüler Karsinomda Hepatik Mikroyapısal Değişikliklerin Kantitatif Değerlendirilmesi

Cihan ÖZGÜR[1], Baran Serdar SUNAL[1], Derya KARABULUT[2], Osman KULA[2], Necdet SÜT[3], İdris KURT[4], Nermin TUNÇBİLEK[2]

[1]*Tekirdağ Namık Kemal University Faculty of Medicine, Department of Radiology, Tekirdağ, Türkiye*
[2]*Trakya University Faculty of Medicine, Department of Radiology, Edirne, Türkiye*
[3]*Trakya University Faculty of Medicine, Department of Biostatistics and Medical Informatics, Edirne, Türkiye*
[4]*Trakya University Faculty of Medicine, Department of Gastroenterology, Edirne, Türkiye*

ABSTRACT

Aim: Magnetic resonance imaging (MRI) is a fundamental imaging modality in the diagnosis, management, and follow-up of hepatocellular carcinoma (HCC). Diffusion-based imaging techniques have been shown to potentially play a significant role in the characterization of focal liver lesions. The aim of this study was to evaluate the contribution of diffusion-weighted imaging (DWI) and diffusion tensor imaging (DTI) in the diagnosis of HCC by comparing DWI and DTI parameters with those of cirrhotic and non-pathological liver parenchyma.

Materials and Methods: This retrospective study included 62 patients with HCC, 56 with cirrhosis, and 52 with non-pathological liver parenchyma who underwent 1.5-T MRI. DWI and DTI sequences were acquired using b-values of 50-800 sec/mm^2 and 20 diffusion-encoding directions. Fractional anisotropy (FA) and apparent diffusion coefficient (ADC) values were calculated and compared across the groups.

Results: ADC values were significantly lower and FA values significantly higher in HCC compared to cirrhotic and non-pathological parenchyma ($p<0.05$). Mean FA was 0.46±0.14 for HCC, 0.39±0.08 for non-cancerous parenchyma, and 0.40±0.07 for both cirrhotic and non-pathological parenchyma. An FA cut-off of 0.45 yielded 38.7% sensitivity and 94.2% specificity [area under the curve (AUC): 0.653, 95% confidence interval (CI): 0.554-0.751] vs. non-pathological parenchyma, and 62.9% sensitivity and 75.0% specificity (AUC: 0.649, 95% CI: 0.542-0.745) vs. cirrhotic parenchyma.

Conclusion: FA values reflected moderate anisotropy in HCC, cirrhotic, and non-pathological liver parenchyma. Although FA differed significantly between HCC and non-malignant tissues, its modest sensitivity limits its utility as a stand-alone biomarker.

Keywords: Cirrhosis, diffusion tensor imaging, diffusion weighted imaging, hepatocellular carcinoma, liver, magnetic resonance imaging

ÖZ

Amaç: Manyetik rezonans görüntüleme (MRG), hepatosellüler karsinomun (HSK) tanı, yönetim ve takip süreçlerinde temel bir görüntüleme yöntemidir. Difüzyon temelli görüntüleme tekniklerinin fokal karaciğer lezyonlarının karakterizasyonunda tamamlayıcı bir rol oynayabileceği gösterilmiştir. Bu çalışmanın amacı, HSK tanısında difüzyon ağırlıklı görüntüleme (DAG) ve difüzyon tensör görüntüleme (DTG) parametrelerinin katkısını, sirotik ve normal karaciğer parankimi ile karşılaştırarak değerlendirmektir.

Gereç ve Yöntem: Bu retrospektif çalışmaya, 1,5 T MRG yapılan 62 HSK hastası, 56 siroz hastası ve 52 normal karaciğer parankimi olan birey dâhil edildi. DAG ve DTG dizileri, 50-800 s/mm^2 b-değerleri ve 20 difüzyon kodlama yönü ile elde edildi. Fraksiyonel anizotropi (FA) ve görünen difüzyon katsayısı (ADC) değerleri hesaplanarak gruplar arasında karşılaştırıldı.

Address for Correspondence: Cihan ÖZGÜR MD, Tekirdağ Namık Kemal University Faculty of Medicine, Department of Radiology,Tekirdağ, Türkiye
E-mail: cihanozguritf@gmail.com **ORCID ID:** orcid.org/0000-0001-8163-9815
Received: 27.03.2025 **Accepted:** 22.05.2025 **Publication Date:** 07.10.2025

Cite this article as: Özgür C, Sunal BS, Karabulut D, Kula O, Süt N, Kurt İ, et al. Quantitative evaluation of hepatic microstructural changes in hepatocellular carcinoma using diffusion tensor imaging. Nam Kem Med J. 2025;13(3):253-260

Bulgular: HSK'da ADC değerleri anlamlı şekilde daha düşük, FA değerleri ise sirotik ve normal parankime kıyasla anlamlı şekilde daha yüksekti (p<0,05). Ortalama FA değeri HSK için 0,46±0,14, non-kanseröz parankim için 0,39±0,08, sirotik ve normal parankim için 0,40±0,07 idi. FA için 0,45 eşik değeri kullanıldığında, HSK'yı normal parankimden ayırt etmede duyarlılık %38,7, özgüllük %94,2 [eğri altındaki alan (AUC): 0,653, %95 güven aralığı (GA): 0,554-0,751); sirotik parankimden ayırt etmede ise duyarlılık %62,9, özgüllük %75,0 (AUC: 0,649, %95 GA: 0,542-0,745) olarak bulundu.

Sonuç: FA değerleri HSK, sirotik ve normal karaciğer parankiminde orta düzeyde anizotropiyi yansıtmaktadır. FA, HSK ile non-malign parankim arasındaki farkı yansıtmakla birlikte, sınırlı duyarlılığı nedeniyle tek başına tanısal biyobelirteç olarak kullanımı kısıtlıdır.

Anahtar Kelimeler: Difüzyon ağırlıklı görüntüleme, difüzyon tensör görüntüleme, hepatoselüler karsinom, karaciğer, manyetik rezonans görüntüleme, siroz

INTRODUCTION

Hepatocellular carcinoma (HCC) is the sixth most commonly diagnosed cancer and the third leading cause of cancer-related mortality worldwide, according to the latest GLOBOCAN data from the World Health Organization (https://gco.iarc.fr/). The age of onset for HCC varies depending on sex, geographic region, and associated risk factors[1]. Concomitant cirrhosis is found in 80% of cases of HCC[2]. Hepatitis B virus (HBV) is the most common cause of virus-associated HCC and is also the predominant etiology in cases without concomitant cirrhosis[3].

Diffusion tensor imaging (DTI) is widely used in neuroimaging, particularly for fiber tractography, but it has also been applied to the liver in a number of publications[4-17]. These studies have explored various aspects of liver DTI, including sequence optimization, hepatic isotropy, fibrosis assessment, and ischemia-reperfusion injury. However, only a few have focused on its use in the evaluation of focal liver lesions, and this area remains underexplored. Prior studies have predominantly evaluated apparent diffusion coefficient (ADC) values, while comprehensive analysis of diffusion tensor parameters particularly fractional anisotropy (FA) in HCC is still limited[4,5].

Although the liver parenchyma is generally considered isotropic, the anisotropic properties of hepatic tumors remain controversial[6,7]. FA, the most commonly utilized DTI-derived parameter in clinical practice, quantifies the degree of anisotropy. FA is a dimensionless value ranging from 0 to 1, where values closer to 1 indicate increasingly directional (anisotropic) diffusion[8].

The aim of this study was to compare DTI parameters of HCC with those of non-malignant liver parenchyma and present preliminary findings.

MATERIALS AND METHODS

Patients

This retrospective study included magnetic resonance imaging (MRI) examinations of 62 patients with HCC (13 females, 49 males; mean age: 65.08±8.19 years, range: 45-84), 56 patients with cirrhosis (19 females, 37 males; mean age: 58.69±11.43 years, range: 20-79), and 52 individuals with non-pathological liver parenchyma (19 females, 33 males; mean age: 59.75±15.78 years, range: 18-84). All imaging was performed at our institution between January and December 2019. In 22 patients (35.5%), HCC diagnosis was confirmed histopathologically. In the remaining 40 patients (64.5%), HCC diagnosis was established based on characteristic MRI findings according to the 2018 Liver Imaging Reporting and Data System (LI-RADS) criteria and multidisciplinary consensus. Among all patients, 49 (79%) were categorized as LI-RADS 5, 11 (17.7%) as LI-RADS 4, and one patient each (1.6%) as LI-RADS 3 and LI-RADS M. Among the 22 histopathologically confirmed HCC cases, 19 (86.4%) were classified as LI-RADS 5, 2 (9%) as LI-RADS 4, and 1 (4.5%) as LI-RADS 3. All final diagnoses were based either on histopathological confirmation or multidisciplinary consensus in accordance with established imaging criteria.

Ethical approval was obtained from the Trakya University Faculty of Medicine Scientific Research Ethics Committee (desicion no: 03/17, date: 03.02.2020).

MR Examination - Image Acquisition

MRI was performed using a 1.5T system (Aera, Siemens Medical Systems, Erlangen, Germany) with an 18-channel body matrix coil. Patients were positioned supine, head first, with arms at their sides. DTI sequences were acquired after T1- and T2-weighted imaging and before dynamic contrast-enhanced sequences. Imaging parameters for DTI were as follows: b-values = 50-800 s/mm^2, repetition time/echo time: 6300/69 ms, slice thickness: 6 mm, interslice gap: 1.2 mm, number of slices: 40, matrix size: 192 × 115, number of signal averages: 1, and 20 diffusion encoding directions. Breath-hold technique was used during acquisition.

Image Analysis-Calculation of DTI Parameters

All imaging data were processed on a dedicated workstation (Syngo.Via VB10B, Siemens). T2-weighted and contrast-enhanced images were used for anatomical reference. ADC maps were automatically generated by the system for each diffusion tensor dataset. Region of interest (ROI) placement

avoided major vascular and biliary structures and was limited to the posterior segment of the right hepatic lobe. For lesion measurements, circular ROIs of 2 cm^2 were used when feasible (Figure 1). In lesions where a 2 cm^2 ROI could not be placed, the largest possible ROI was selected. For large lesions, the average of three ROI measurements on the same slice was recorded. Lesions smaller than 1 cm were excluded due to limited spatial resolution. In patients with HCC, ROIs were placed to avoid cystic and necrotic areas, selecting the most appropriate solid portion of the lesion.

Statistical Analysis

Comparisons of DTI parameters between the HCC, cirrhotic, and non-pathological liver groups were performed using either the Student's t-test or the Mann-Whitney U test, based on the normality of data distribution. ROC curve analysis was conducted to determine optimal diagnostic cut-off values for distinguishing HCC, and corresponding sensitivity, specificity, positive predictive value, and negative predictive value were calculated. A p-value of less than 0.05 was considered statistically significant. Interobserver agreement for the HCC measurements was evaluated by calculating intraclass correlation coefficients (ICCs) between two radiologists, using a two-way random effects model with absolute agreement. ICCs were calculated separately for diffusion-weighted imaging (DWI) ADC, DTI ADC, and DTI FA values. All statistical analyses were conducted using SPSS software (Version 25, IBM Corp., Armonk, NY, USA).

RESULTS

HCC was more prevalent among the elderly and male patients, with a male-to-female ratio of 3.7. HBV was the most common etiological factor in both the HCC and cirrhotic groups, observed in 59.7% (n=37) and 53.6% (n=30) of cases, respectively.

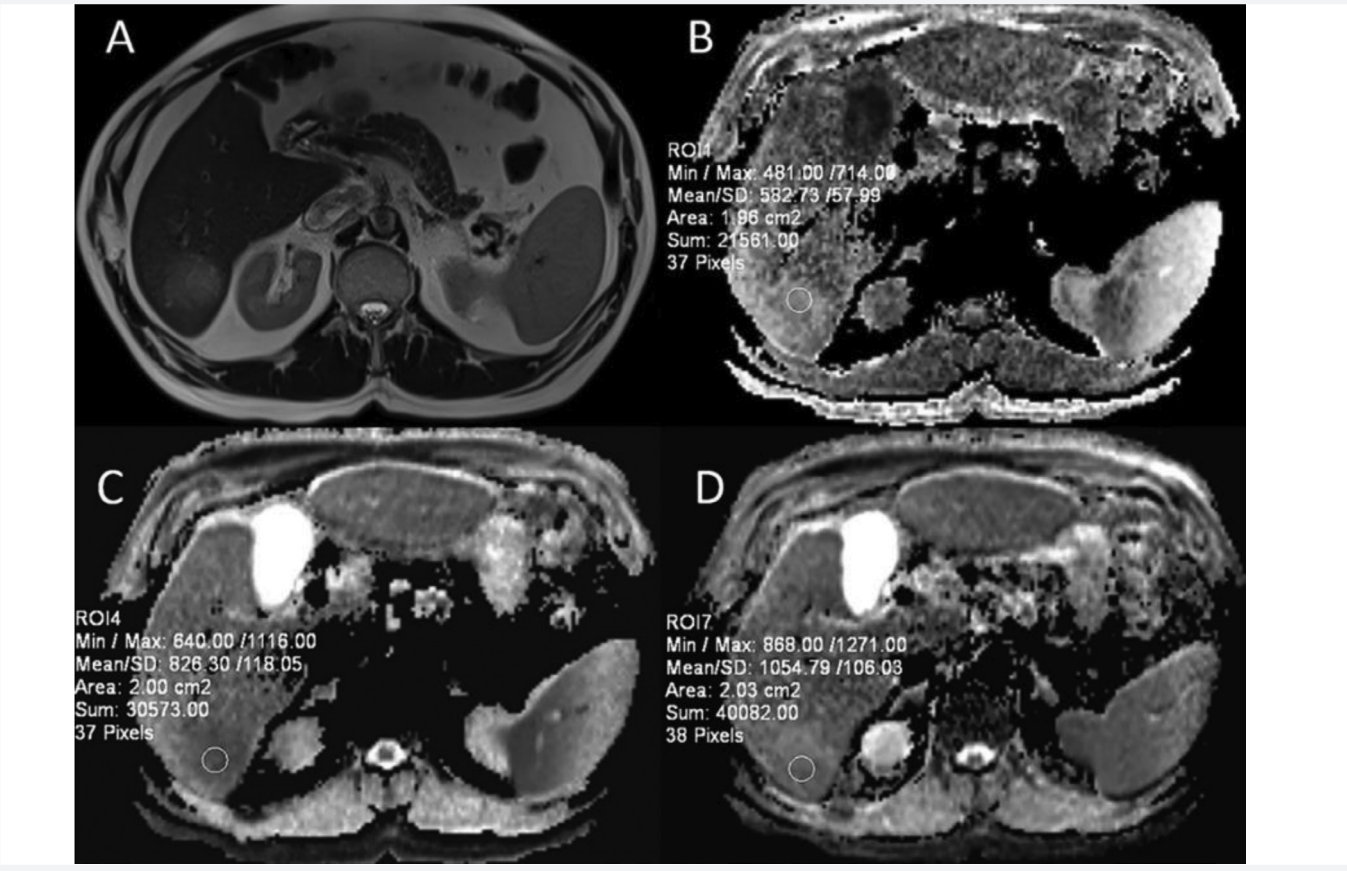


Figure 1. MRI images of a 60-year-old man with biopsy-proven hepatocellular carcinoma

T2-weighted images were used for anatomical orientation (A). Compared to surrounding liver parenchyma, hepatocellular carcinoma is seen as hyperintense (B) on the FA map (mean FA of 0.58) and hypointense on the ADC map (mean ADC of 0.83 x 10-3 mm^2/s) from diffusion tensor imaging data (C). On the ADC map from conventional diffusion imaging data (D), the mass is of intermediate signal intensity (mean ADC of 1.05 x 10-3 mm^2/s)

MRI: Magnetic resonance imaging, ADC: Apparent diffusion coefficient, FA: Fractional anisotropy, ROI: Region of interest, SD: Standard deviation

Patient Characteristics

Clinical and laboratory characteristics of all patient groups are summarized in Table 1.

MRI Features of HCC

In HCC patients, lesion characteristics were as follows: T1 hyperintensity in 37.1% (n=23), intermediate T2 hyperintensity in 82.3% (n=51), visually assessed diffusion restriction in 59.7% (n=37), arterial phase enhancement in 87.1% (n=54), capsular enhancement in the delayed phase in 85.5% (n=53), and washout in 88.7% (n=55). Satellite tumors were observed in 12.9% (n=8), lymphadenopathy in 25.8% (n=16), and vascular invasion in 27.4% (n=17). Intralesional fat was present in 11.3% (n=7), hemorrhage in 30.6% (n=19), and necrosis in 21% (n=13). The interobserver reliability analysis yielded ICC values of 0.858 [95% confidence interval (CI): 0.748-0.922] for DWI ADC, 0.853 (95% CI: 0.713-0.923) for DTI ADC, and 0.837 (95% CI: 0.712-0.910) for DTI FA, indicating good agreement between the two observers for the HCC measurements.

Conventional DWI Findings

Conventional DWI was performed in 45 out of 62 HCC patients. ADC values in HCC lesions were significantly lower than those in non-cancerous, cirrhotic, and non-pathological liver parenchyma. Using a cut-off ADC value of 0.99×10^{-3} mm^2/s to distinguish HCC from non-pathological parenchyma yielded a sensitivity of 71.1% and specificity of 69.2%. When the cut-off value for differentiating HCC from cirrhotic parenchyma was set at 0.95×10^{-3} mm^2/s, sensitivity and specificity were calculated as 62.2% and 80.4%, respectively.

Comparison with Child-Pugh Scores

A weak negative correlation was found between DWI-based ADC values and Child-Pugh scores (r=-0.396, p=0.008). No statistically significant correlation was found between ADC or FA values derived from DTI and Child-Pugh scores (p=0.695 and p=0.932, respectively).

Table 1. Clinical characteristics and laboratory data of the patients

Clinical factors and laboratory data	HCC group patients (n=62) (%)	Cirrhosis group patients (n=56) (%)
Sex		
Male	49 (79%)	37 (66.1%)
Female	13 (21%)	19 (33.9%)
Age (y, mean ± SD) (range)	65.1±8.2 (45-84)	58.7±11.2 (20-79)
Etiology		
HBV	37 (59.7%)	30 (53.6%)
HCV	3 (4.8%)	1 (1.8%)
Ethanol	8 (12.9%)	5 (8.9%)
HBV and ethanol	4 (6.5%)	2 (3.6%)
NASH	1 (1.6%)	1 (1.8%)
Cryptogenic or unknown etiology	8 (12.9%)	12 (21.4%)
Autoimmune hepatitis	1 (1.6%)	2 (3.6%)
Wilson disease	1 (1.8%)	
HBV-HDV coinfection	1 (1.8%)	
Primary biliary cirrhosis	1 (1.8%)	
Serum α-fetoprotein level (ng/mL) (mean ± SD)	1205.97±6117.35 (2-48000)	17.62±53.45 (1-350)
Cirrhosis	48 (77.4%)	56 (100%)
Child Pugh class		
A	31 (50%)	
B	20 (32%)	
C	9 (15%)	

HBV: Hepatitis B virus, HCC: Hepatocellular carcinoma, HCV: Hepatitis C virus, HDV: Hepatitis D virus, NASH: Non-alcoholic steatohepatitis, SD: Standard deviation, y: Years

Comparison Between Non-pathological Parenchyma and HCC

Patients with HCC had significantly lower ADC values and higher FA values compared to those with non-pathological parenchyma. Using a cut-off ADC value of 0.94×10^{-3} mm^2/s yielded a sensitivity of 33.9% (95% CI, 23.3-46.3) and specificity of 98.1% (95% CI, 89.9-99.7), with an area under the curve (AUC) of 0.656 (95% CI, 0.558-0.753). For FA, a threshold of 0.65 resulted in sensitivity and specificity of 38.7% (95% CI, 27.6-51.2) and 94.2% (95% CI, 81.8-97.0), respectively, with an AUC of 0.653 (95% CI, 0.554-0.751). In conventional DWI measurements, using a cut-off ADC value of 0.99×10^{-3} mm^2/s provided a sensitivity of 51.6% (95% CI, 39.4-63.6) and specificity of 69.2% (95% CI, 55.7-80.1), with an AUC of 0.688 (95% CI, 0.582–0.799). Group means are presented in Table 2 (Figure 2).

Comparison Between Cirrhotic Parenchyma and HCC

ADC values were significantly lower and FA values significantly higher in HCC lesions compared to cirrhotic parenchyma. Using a threshold of 1.07×10^{-3} mm^2/s for ADC, sensitivity and specificity were 56.5% (95% CI, 44.1-68.1) and 76.8% (95% CI, 64.2-85.9), respectively, with an AUC of 0.702 (95% CI, 0.603-0.794). For FA, a threshold of 0.43 yielded sensitivity and specificity of 62.9% (95% CI, 50.5-73.8) and 75.0% (95% CI, 62.3-84.5), respectively, with an AUC of 0.649 (95% CI, 0.542-0.745). In conventional DWI measurements, using a threshold of 0.95×10^{-3} mm^2/s provided a sensitivity of 62.2% (95% CI, 47.6-74.9) and specificity of 80.4% (95% CI, 68.2-88.7), with an AUC of 0.696 (95% CI, 0.579-0.803). No statistically significant difference was found in parenchymal DTI and FA measurements between HCC patients with or without underlying cirrhosis (p=0.832 and p=0.911, respectively) (Figure 3). Group means are presented in Table 3.

Comparison Between Cirrhotic and Non-pathological Liver Parenchyma

There was no statistically significant difference in DTI-ADC and FA values between cirrhotic and non-pathological liver parenchyma (p=0.469 and p=0.995, respectively).

Comparison Between HCC Lesion and Adjacent Non-tumoral Parenchyma

A paired sample t-test revealed significant differences between DTI-ADC and FA measurements obtained from HCC lesions and adjacent non-tumoral parenchyma in the same patients (p<0.009 and p=0.005, respectively). The corresponding mean values are summarized in Table 3.

DISCUSSION

In this preliminary study, ADC and FA values derived from DTI are presented in a relatively large cohort of patients with HCC. Our study was aimed to evaluate whether quantitative diffusion-based imaging parameters, particularly FA derived from DTI and ADC from conventional DWI, can reliably differentiate HCC from its surrounding parenchymal background. Rather than aiming to characterize the full radiologic spectrum of hepatic lesions, we focused on non-tumoral and cirrhotic liver tissue both clinically relevant comparators due to their established roles in the natural history and differential diagnosis of HCC. This design enabled us to investigate DTI's diagnostic performance within a well-defined and pathophysiologically coherent framework. Notably, unlike benign lesions, which are typically identified with high accuracy using conventional MRI protocols, cirrhotic liver parenchyma represents a diagnostically challenging substrate where subtle microstructural alterations may overlap with early tumorigenic changes. Therefore, establishing standardized DTI metrics in this context may serve as a crucial step toward integrating DTI into future multiparametric liver imaging strategies.

Table 2. Comparison of diffusion tensor imaging parameters between normal liver and HCC

	Non-pathological liver	HCC	p-value
ADC value (x 10^{-3} s/mm^2) (Mean ± SD)	1.14±0.13	1.04±0.22	<0.002
FA value (Mean ± SD)	0.40±0.07	0.46±0.14	0.004

ADC: Apparent diffusion coefficient, FA: Fractional anisotropy, HCC: Hepatocellular carcinoma, SD: Standard deviation

Table 3. Comparison of diffusion tensor imaging parameters between cirrhotic liver and HCC

	Cirrhotic liver	HCC	p-value
ADC value (x 10^{-3} s/mm^2) (Mean ± SD)	1.18±0.03	1.04±0.22	<0.001
FA value (Mean ± SD)	0.40±0.07	0.46±0.14	0.004

ADC: Apparent diffusion coefficient, FA: Fractional anisotropy, HCC: Hepatocellular carcinoma, SD: Standard deviation

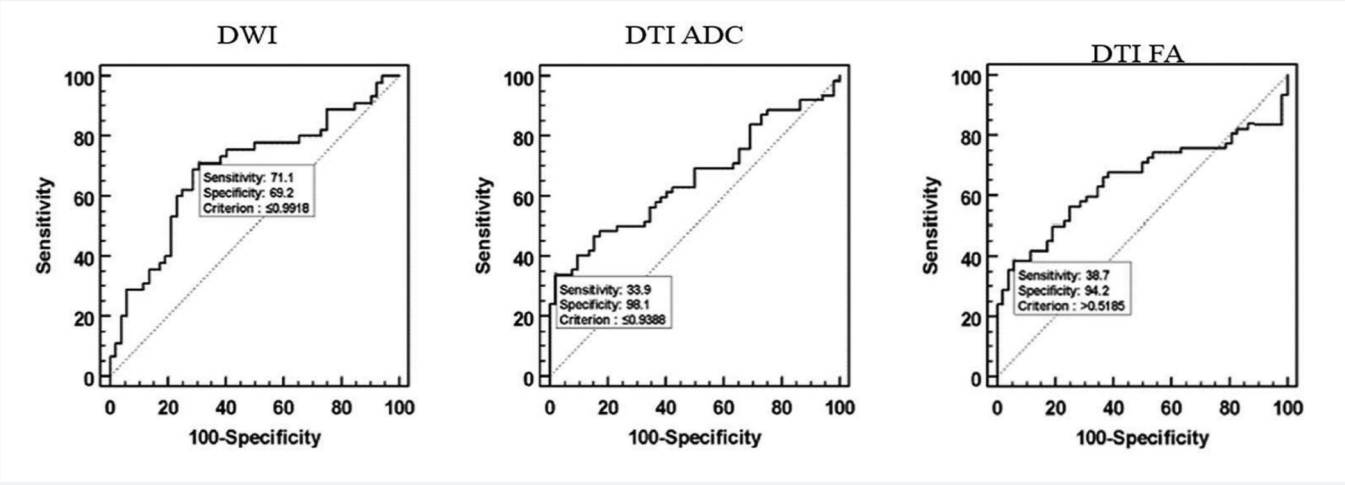


Figure 2. ROC curve on the group of patients with hepatocellular carcinoma and non-pathological liver parenchyma

DWI: Diffusion-weighted imaging, DTI: Diffusion tensor imaging, ADC: Apparent diffusion coefficient, FA: Fractional anisotropy

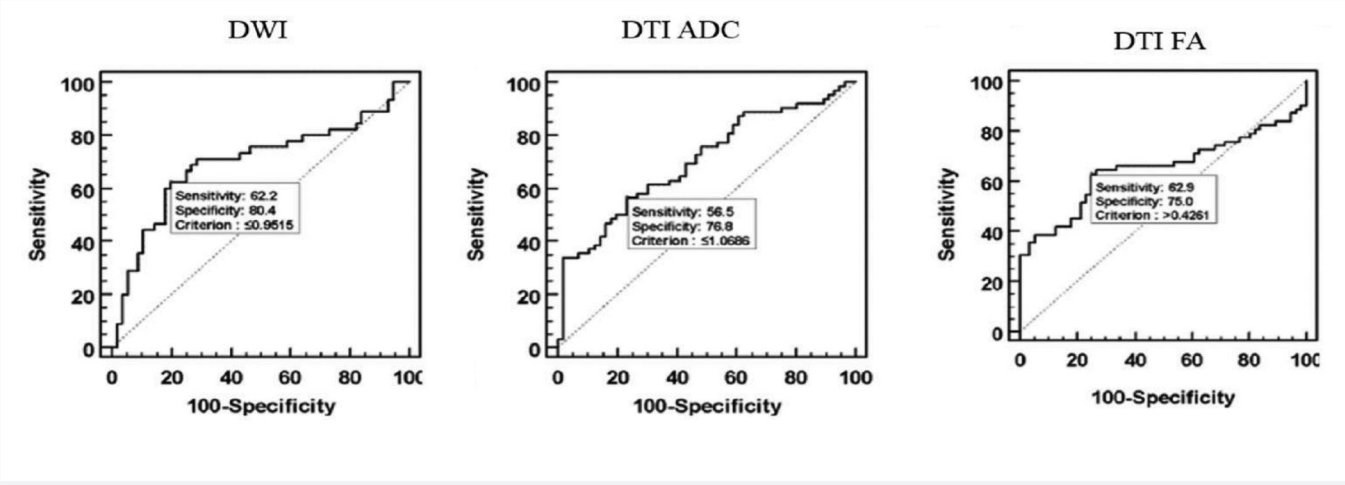


Figure 3. ROC curve on the group of patients with hepatocellular carcinoma and cirrhotic liver parenchyma.

DWI: Diffusion-weighted imaging, DTI: Diffusion tensor imaging, ADC: Apparent diffusion coefficient, FA: Fractional anisotropy

Although there are limited studies on the use of DTI in liver imaging, early findings suggest its potential utility[4-6,10-18]. Taouli et al.[6] reported that diffusion in non-pathological and cirrhotic liver parenchyma, as well as in focal hepatic lesions, was predominantly isotropic likely due to the liver's randomly organized microstructure, which contrasts with more structured organs like the brain or kidneys. In the same study, malignant lesions demonstrated lower ADC values than both benign lesions and liver parenchyma.

To date, only a limited number of studies have specifically investigated DTI findings in HCC[5,16-18]. Li et al.[5], using a 3T MRI system, focused on sequence optimization and the diffusion characteristics of HCC. Although our study was conducted at 1.5T, the observed diffusion patterns were largely comparable. While they recommended nine diffusion-encoding directions, we employed 20 to improve the accuracy and robustness of DTI[5].

Recent studies have further supported the clinical value of DTI in hepatic imaging, particularly for differentiating HCC from other hepatic lesions. In a study by Saleh et al.[16], FA and mean diffusivity (MD) were significantly lower in benign lesions compared to HCC, with FA values showing the highest diagnostic accuracy. Notably, FA >0.38 emerged as an independent predictor of HCC, outperforming both ADC and

MD in their regression model. These findings are in line with our observation that FA values in HCC are significantly elevated compared to non-malignant liver tissues, emphasizing the anisotropic nature of malignant lesions.

Mahmoud et al.[17] also demonstrated the diagnostic potential of liver DTI by comparing DTI-derived parameters with LI-RADS classification. They found a moderate positive correlation between FA and LI-RADS category, and a substantial negative correlation between DTI-ADC and LI-RADS, concluding that DTI-ADC and FA values perform better than conventional DWI-ADC in discriminating between benign and malignant lesions. This supports our findings, particularly in terms of the inverse relationship between FA and DTI-ADC values in malignant hepatic tissue.

Karim et al.[18] investigated the role of DTI in both lesion characterization and post-treatment response assessment. They reported that an FA cut-off value of >0.29 differentiated malignant from benign lesions with 95% sensitivity and 70% specificity.

In our study, which focused on distinguishing HCC from non-malignant parenchyma, we identified a higher optimal FA threshold of 0.43, yielding 62.9% sensitivity and 75% specificity. The lower sensitivity observed in our cohort may reflect the more subtle microstructural differences between malignant and cirrhotic tissues, as opposed to benign lesions, which are generally more isotropic. These variations highlight the importance of reference tissue selection and patient population characteristics in DTI analysis.

We found that the mean FA value in HCC (0.46±0.14) reflected moderate anisotropy, which was comparable to FA values in cirrhotic and non-pathological liver parenchyma (0.40±0.07). Although liver parenchyma is generally described as isotropic[6,7], the diffusion behavior of hepatic tumors remains controversial. Kinoshita et al.[19], in a study of malignant brain tumors, found a positive correlation between FA and both tumor cell density and Ki-67 index, suggesting that higher FA may reflect increased cellularity, although the biological basis for this relationship remains unclear.

Although our study demonstrated significantly higher FA values in HCC compared to non-pathological liver parenchyma, the relatively low sensitivity (38.7-62.9%) limits the stand-alone diagnostic utility of FA as a biomarker. This finding indicates that while elevated FA values can assist in suggesting malignancy, FA measurements alone may not reliably differentiate HCC in routine clinical practice. Therefore, FA should be interpreted alongside conventional MRI features and other quantitative diffusion parameters such as ADC. Future technical advances or multiparametric approaches may help overcome this limitation.

Study Limitations

This study has several limitations. First, its retrospective design may have introduced selection bias. Second, not all HCC and cirrhotic cases had histopathological confirmation; thus, the fibrosis stage in cirrhotic patients could not be evaluated. Third, ADC and FA measurements can vary depending on scanner hardware, observer expertise, and physiological motion. Previous reports have noted that cardiac motion can influence diffusion metrics, particularly in the left hepatic lobe, where systolic movement significantly elevates FA values[12]. To minimize motion artifacts, we focused our ROI placement on the posterior segment of the right lobe. While the majority of HCC lesions in our study were categorized as LI-RADS 5, a small number of LI-RADS 4 and indeterminate (LI-RADS 3) cases were also included. Although these cases were either histologically confirmed or supported by consensus diagnosis based on imaging features, their inclusion may introduce a degree of heterogeneity in diagnostic certainty, and should be considered when interpreting the results. Furthermore, the lack of significant difference in DTI measurements between cirrhotic and non-pathological parenchyma may reflect the heterogeneity of our cirrhosis group. Future studies should incorporate fibrosis grading and interobserver agreement analyses to better validate these findings.

CONCLUSION

In summary, this study demonstrated that DTI-derived parameters, particularly FA and ADC values, differ significantly between HCC lesions and non-lesional liver parenchyma. These findings suggest that DTI can provide additional microstructural information specific to HCC. However, the moderate sensitivity of FA measurements, and the absence of standardized acquisition protocols currently limit the clinical applicability of DTI as a stand-alone diagnostic tool. Rather than replacing conventional imaging criteria, DTI metrics should be considered complementary parameters that may enhance lesion characterization when used in conjunction with established MRI features. Future prospective studies involving larger, diverse patient populations and standardized imaging protocols are warranted to validate these preliminary findings and clarify the potential role of DTI in the non-invasive evaluation of HCC.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Trakya University Faculty of Medicine Scientific Research Ethics Committee (desicion no: 03/17, date: 03.02.2020).

Informed Consent: This is retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: İ.K., Concept: C.Ö., D.K., N.T., Design: C.Ö., O.K., N.S., İ.K., Data Collection or Processing: C.Ö., B.S.S., N.S., İ.K., Analysis or Interpretation: O.K., N.S., Literature Search: C.Ö., B.S.S., D.K., O.K., İ.K., N.T., İ.K., N.T., Writing: C.Ö., B.S.S., D.K.

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

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

REFERENCES

1. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American Association for the Study of Liver Diseases. Hepatology. 2018;68:723-50.
2. Bruix J, Sherman M, American Association for the Study of Liver D. Management of hepatocellular carcinoma: an update. Hepatology. 2011;53:1020-2.
3. Ozakyol A. Global epidemiology of hepatocellular carcinoma (HCC Epidemiology). J Gastrointest Cancer. 2017;48:238-40.
4. Erturk SM, Ichikawa T, Kaya E, Yapici O, Ozel A, Mahmutoglu AS, et al. Diffusion tensor imaging of cysts, hemangiomas, and metastases of the liver. Acta Radiol. 2014;55:654-60.
5. Li X, Liang Q, Zhuang L, Zhang X, Chen T, Li L, et al. Preliminary study of MR diffusion tensor imaging of the liver for the diagnosis of hepatocellular carcinoma. PLoS One. 2015;10:e0135568.
6. Taouli B, Vilgrain V, Dumont E, Daire JL, Fan B, Menu Y. Evaluation of liver diffusion isotropy and characterization of focal hepatic lesions with two single-shot echo-planar MR imaging sequences: prospective study in 66 patients. Radiology. 2003;226:71-8.
7. Girometti R, Esposito G, Bagatto D, Avellini C, Bazzocchi M, Zuiani C. Is water diffusion isotropic in the cirrhotic liver? A study with diffusion-weighted imaging at 3.0 Tesla. Acad Radiol. 2012;19:55-61.
8. Lee Y, Kim H. Assessment of diffusion tensor MR imaging (DTI) in liver fibrosis with minimal confounding effect of hepatic steatosis. Magn Reson Med. 2015;73:1602-8.
9. Chernyak V, Fowler KJ, Kamaya A, Kielar AZ, Elsayes KM, Bashir MR, et al. Liver imaging reporting and data system (LI-RADS) Version 2018: imaging of hepatocellular carcinoma in at-risk patients. Radiology. 2018;289:816-30.
10. Tosun M, Inan N, Sarisoy HT, Akansel G, Gumustas S, Gurbuz Y, et al. Diagnostic performance of conventional diffusion weighted imaging and diffusion tensor imaging for the liver fibrosis and inflammation. Eur J Radiol. 2013;82:203-7.
11. Taouli B, Chouli M, Martin AJ, Qayyum A, Coakley FV, Vilgrain V. Chronic hepatitis: role of diffusion-weighted imaging and diffusion tensor imaging for the diagnosis of liver fibrosis and inflammation. J Magn Reson Imaging. 2008;28:89-95.
12. Wong OL, Gloh Lo G, Lee R, Wa Li W, Lung Chan P, Ki Yu S, et al. The effect of respiratory and cardiac motion in liver diffusion tensor imaging (DTI). J Comput Assist Tomogr. 2014;38:352-9.
13. Taouli B, Martin AJ, Qayyum A, Merriman RB, Vigneron D, Yeh BM, et al. Parallel imaging and diffusion tensor imaging for diffusion-weighted MRI of the liver: preliminary experience in healthy volunteers. AJR Am J Roentgenol. 2004;183:677-80.
14. Cheung JS, Fan SJ, Gao DS, Chow AM, Man K, Wu EX. Diffusion tensor imaging of liver fibrosis in an experimental model. J Magn Reson Imaging. 2010;32:1141-8.
15. Ugurlar OY, Bugdaycı MS, Mahmutoglu AS, Alkım C, Ozguven BY, Ozel A, et al. The role of diffusion tensor imaging for the assessment of liver fibrosis and inflammation in chronic viral hepatitis: a preliminary study. Marmara Med J. 2020;33:97-101.
16. Saleh GA, Elmokadem AH, Razek AA, El-Morsy A, Hamdy O, Eleraky ES, et al. Utility of diffusion tensor imaging in differentiating benign from malignant hepatic focal lesions. Eur Radiol. 2023;33:1400-11.
17. Mahmoud BE, Metwally LI, Salama RM, Yehia M, Ibrahim IMH, Mikhael HSW. Does liver diffusion tensor imaging (L-DTI) has a role in differentiation of hepatic focal lesions? Analytic study for assessment of the value of L-DTI in differentiating hepatic focal lesions according to LI-RADS. Egypt J Radiol Nucl Med. 2022;53:154.
18. Karim MA, Nehad MF, Tharwat M. Characterization of hepatic focal lesions by diffusion tensor imaging and how far it can predict post-treatment response? Med J Cairo Univ. 2021;89:1263-73.
19. Kinoshita M, Hashimoto N, Goto T, Kagawa N, Kishima H, Izumoto S, et al. Fractional anisotropy and tumor cell density of the tumor core show positive correlation in diffusion tensor magnetic resonance imaging of malignant brain tumors. Neuroimage. 2008;43:29-35.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.44265
Nam Kem Med J 2025;13(3):261-268

Insights from 151 Consecutive Pancreaticoduodenectomies at a High-Volume Tertiary Center: A Cross-Sectional Observational Study

Yüksek Hacimli Bir Tersiyer Merkezde 151 Ardışık Pankreatikoduodenektomi: Kesitsel Gözlemsel Bir Çalışma

Ali Emre ATICI, Ayşegül Bahar ÖZOCAK, Mümin COŞKUN, Şevket Cumhur YEĞEN

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

ABSTRACT

Aim: Pancreaticoduodenectomy (PD) is a complex and essential procedure in the treatment of localized periampullary neoplasms. Due to its comlexity, postoperative morbidity and mortality remain significant concerns. The aim of this study was to evaluate our short-term outcomes, categorize them according to international standards, and compare our findings with those reported in the existing literature.

Materials and Methods: One hundred and fifty-one patients underwent classical PD for pancreatic tumors between February 2019 and May 2023 at the Department of General Surgery, Marmara University Faculty of Medicine, Department of General Surgery. Patients meeting the inclusion criteria were enrolled in the study. Clinical, operative, pathological, and short-term outcome data, prospectively recorded, were retrospectively analyzed.

Results: The mean age of the patients was 63.6 years, with 87 (57.6%) being male. The median operative time was 227.4±48.5 minutes. Clavien-Dindo grade 3 or higher complications were observed in 38 patients (25.2%). Intraabdominal abscesses were noted in 9 patients, chylous fistula in 9 patients, and postoperative bleeding in 6 patients. Postoperative pancreatic fistula was diagnosed in 54 patients (35.8%), of which 38 (70.3%) were classified as grade A and 16 (29.7%) as grade B. No grade C pancreatic fistulas were observed. The overall incidence of delayed gastric emptying was 26.5% (n=40). Two patients required reoperation: one for postoperative bleeding and the other for gastroenterostomy leakage. In the early postoperative period, five patients died.

Conclusion: Effective management of complications following complex surgeries plays a critical role in improving postoperative outcomes. The morbidity and mortality rates in our series were relatively lower compared to those reported in the literature.

Keywords: Pancreaticoduodenectomy, periampullary tumor, postoperative complications

ÖZ

Amaç: Pankreatikoduodenektomi (PD), lokalize periampüller neoplazmların tedavisinde karmaşık ve hayati bir cerrahi prosedür olmakla birlikte postoperatif morbidite ve mortalite oranları hala en önemli endişe konusu olmaya devam etmektedir. Bu çalışmanın amacı, postoperatif kısa dönem sonuçlarımızı değerlendirmek, uluslararası standartlara göre sınıflandırmak ve elde edilen bulguları mevcut literatürle karşılaştırmaktır.

Gereç ve Yöntem: 2019 Şubat ile 2023 Mayıs tarihleri arasında, Marmara Üniversitesi Tıp Fakültesi, Genel Cerrahi Anabilim Dalında periampuller tümör tanısı ile 151 hastaya klasik PD operasyonu yapıldı. Dahil edilme kriterlerini karşılayan hastaların prospektif olarak kaydedilen klinik verileri ve kısa dönem postoperatif sonuçları retrospektif olarak analiz edildi.

Bulgular: Hastaların ortalama yaşı 63,6 yıl olup, 87 hasta (%57,6) erkekti. Ortanca operasyon süresi 227,4±48,5 dakika olarak bulundu. Clavien-Dindo sınıflamasına göre derece 3 ve üzeri daha yüksek komplikasyonlar 38 hastada (%25,2) izlendi. Dokuz hastada intraabdominal apse, 9 hastada

Address for Correspondence: Ali Emre ATICI MD, Marmara University Faculty of Medicine, Department of General Surgery, İstanbul, Türkiye
E-mail: ali.emre@marmara.edu.tr **ORCID ID:** orcid.org/0000-0001-5094-0461
Received: 06.01.2025 **Accepted:** 22.05.2025 **Publication Date:** 07.10.2025

Cite this article as: Atıcı AE, Özocak AB, Coşkun M, Yeğen ŞC. Insights from 151 consecutive pancreaticoduodenectomies at a high-volume tertiary center: a crosssectional observational study. Nam Kem Med J. 2025;13(3):261-268

şilöz fistül ve 6 hastada postoperatif kanama saptandı. Elli dört hastada (%35,8) postoperatif pankreatik fistül gelişirken, 38'i (%70,3) derece A, 16'sı (%29,7) derece B fistül ile uyumlu idi. Derece C pankreatik fistül hiçbir hastamızda izlenmedi. 40 (%26,5) hastada geçikmiş mide boşalımı gözlendi. Bir hasta postoperatif kanama diğer bir hastada gastroenterostomi kaçağı nedeniyle postoperatif erken dönemde yeniden opere edildi. Toplam 5 (%3) hasta ise erken postoperatif dönemde hayatını kaybetti.

Sonuç: Karmaşık cerrahiler sonrasında komplikasyonların etkin yönetimi, postoperatif sonuçları iyileştirmede kritik bir rol oynamaktadır. Morbidite ve mortalite oranlarımız literatürde bildirilen verilere kıyasla nispeten daha düşük bulunmuştur.

Anahtar Kelimeler: Pankreatikoduodenektomi, periampüller tümör, postoperatif komplikasyonlar

INTRODUCTION

Pancreaticoduodenectomy (PD) is one of the most complex abdominal surgeries performed for malignant lesions of the pancreatic head and periampullary region[1]. German surgeon Kausch performed the first successful PD for periampullary tumors in 1909, but it was later popularized by Allen Whipple[2]. In the early years, due to high morbidity and mortality rates, only a limited number of procedures were performed. However, since the 1980s, with advancements in surgical techniques and perioperative patient care, mortality rates have significantly decreased and have now fallen below 5%. In contrast, morbidity rates have not decreased at the same rate and continue to range between 22% and 57%[1,2]. Postoperative morbidity, by causing prolonged hospital stay, increased treatment costs, and delays in adjuvant therapy, negatively impacts prognosis and survival[3]. Therefore, although it may not be possible to eliminate postoperative morbidity, identifying its causes, determining predictive factors, and taking preventive measures will significantly reduce morbidity rates, such as mortality rates.

Currently, the most frequently observed causes of morbidity after PD are post-operative pancreatic fistula (POPF), delayed gastric emptying (DGE), and post-pancreatectomy hemorrhage (PPH)[1]. To standardize the definition of these complications and enable inter-center comparisons, internationally recognized classification systems have been developed considering the severity of complications. After that, most efforts have focused on reducing them[3]. This study aimed to assess the short-term outcomes, classify our probable complications according to international definitions, and compare our results with the current literature.

MATERIALS AND METHODS

One hundred and eleven patients underwent classical PD (with 20-40 percent antrectomy) due to tumoral mass in the periampullary region at the Department of General Surgery, Marmara University Faculty of Medicine, from February 2019 to May 2023, and those who met the inclusion criteria were enrolled. Prospectively recorded patients' clinical, operative, pathologic, and short-term outcomes were analyzed retrospectively. The study protocol was approved by the Ethics Committee of Marmara University Faculty of Medicine, and was conducted by the principles of the Declaration of Helsinki (desicion no: 09.2023.892, date: 14.07.2023).

Study Population

Eligibility criteria included being over 18 years of age, having a tumor confined to the pancreatic head, and meeting the definition of resectable disease. Resectability was defined based on the criteria established by Isaji et al.[4] Patients who had received neoadjuvant chemotherapy, had distant metastasis, had synchronous or metachronous tumors, and in addition, to minimize variability related to postoperative outcomes, patients who had undergone complex procedures such as vascular or multiorgan resections (such as gastric or colonic) were excluded. Patient age, sex, body mass index (BMI), American Society of Anesthesiologists classification, Eastern Cooperative Oncology Group performance status, co-morbidities, surgical parameters, and early postoperative complications were recorded.

Surgery

All surgical procedures were carried out by two experienced hepatobiliary surgeons. The patients underwent resection with curative intent, including standard open-technique PD and lymphadenectomy[5]. Pyloric preservation was not performed on any patients. Anastomoses were constructed on a single jejunal loop repositioned upwards through transverse mesocolon into a subhepatic resection field in the reconstruction phase. Pancreaticojejunostomy was performed using either the modified Blumgart or Heidelberg technique[6,7]. Surgical team also performed a subjective assessment of the pancreatic tissue texture intraoperatively. Hepaticojejunostomy was performed with absorbable interrupted sutures. Then, a hand-sewn or linear stapler performed gastrojejunostomy to the same jejunal loop.

Postoperative Care

The nasogastric tube was removed within 24 hours after surgery and reinserted if nausea and vomiting occurred. Oral intake was initiated on the first postoperative day and increased based on tolerance. Prophylaxis for deep vein thrombosis was

started with low-molecular-weight heparin 12 hours before anesthesia induction. Antibiotic prophylaxis was administered to patients for 24 hours. No patient received a prophylactic somatostatin analog to prevent POPF. In the decision to remove the abdominal drains, the character of the drainage content and amylase levels were considered.

Definitions

Postoperative complications were classified from grade I to grade V according to the Clavien-Dindo (C-D) classification[8]. DGE and POPF were categorized according to the criteria of the International Study Group of Pancreatic Surgery[9,10]. Deaths related to postoperative complications were recorded.

Statistical Analysis

Descriptive statistical analysis was performed using the SPPS Software Program Version 26.0 (SPSS, Inc., Chicago, IL, USA). Continuous variables are expressed as mean ± standard deviation (SD). Categorical variables are described as frequency and percentage. Categorical variables were analyzed using the chi-square test or the Fisher's exact test, as appropriate, to assess associations between groups. Results were considered statistically significant at a p-value<0.05.

RESULTS

Mean ± SD age was 63.6 (11.9) years, and 87 patients (57.6%) were male. Mean ± SD BMI of the patients was 26.2 (4.4) kg/m^2. Diabetes mellitus was the most common co-morbidity present in 57 (37.7%) patients. Of the 151 patients, 107 required preoperative biliary drainage, of which 72 (47.7%) had endoscopic retrograde cholangiopancreatography (ERCP)-plastic biliary stent, and 35 (23.2%) had percutaneous biliary catheter placement. Eighty-eight (58.3%) had normal bilirubin levels pre-operatively. Baseline patient characteristics are shown in Table 1.

Intraoperative Findings

Median operative time was 227.4±48.5 minutes. All pancreaticojejenostomy anastomoses were performed duct-to-mucosa using either the modified Blumgart (n=125, 82%) or the Heidelberg technique (n=26, 17%). Out of these, stents were used in 71 (47.3%) cases, while stents were not applied in 79 (52.7%) cases. 69 (45.7%) patients were observed to have a soft pancreatic texture, while 76 (56.3%) patients had a firm pancreatic texture (as decided by the surgeon who performed the surgery). The pancreatic duct size was measured by a sterile plastic ruler intraoperatively. 43 (28.5%) had a duct size of less than 3 mm, while 96 (63.6%) had a duct size of more than 3 mm (Table 2).

Postoperative Results

The total rate of POPF was 35.8% (n=54). Out of 54 patients, 38 (70.3%) patients had grade A and 16 (29.7%) had grade B pancreatic fistula. Grade C pancreatic fistula was not observed. The incidence of DGE was 26.5% (n=40). Of 40 patients, 23 (57.5%) had grade A, 1 (37.5%) had grade B, and 2 (5%) had grade C (Table 3). Intra-abdominal abscess (n=9), chylous

Table 1. Demographic data

All cases n=151 (%)	
Age mean (SD)	63.6±11.9
Sex	
Female	64 (42.4)
Male	87 (57.6)
BMI mean (SD), kg/m^2	26.2±4.4
ASA score	
1	15 (9.9)
2	66 (43.7)
3	70 (46.3)
ECOG score	
0	51 (33.7)
1	80 (52.9)
2	20 (13.2)
Diabetes mellitus	
No	94 (62.3)
Yes	57 (37.7)
Total bilirubin mean (SD), mg/dL	
0-1.99	88 (58.3)
2-6	33 (21.9)
6<	27 (17.9)
Pre-operative biliary drainage	
No drainage	43 (28.5)
ERCP	72 (47.7)
PTBD	35 (23.2)
no data (missing)	1 (0.7)
Tumor location	
Pancreatic head	76 (50.3)
Papillary tumor	23 (15.2)
Uncinate process	19 (12.6)
Periampullary tumor	31 (20.5)

SD: Standard deviation, BMI: Body-mass index, ASA: American Society of Anesthesiologists, ECOG: Eastern Cooperative Oncology Group, ERCP: Endoscopic retrograde cholangiopancreatography, PTBD: Percutaneous transhepatic biliary drainage

fistula (n=9), and postoperative bleeding (n=6) were the most common complications. Two patients were re-operated due to bleeding and anastomotic leakage, respectively. In the early postoperative period, a total of 5 patients died. Three of these patients died due to sepsis, while the other two succumbed to postoperative bleeding. Mean postoperative hospital stay was 8.7±4.3 days (Table 4). When postoperative complications were evaluated in general rather than specifically, 38 (25.2%) had C-D grade 3 or above complications. No significant association was identified between pancreatic texture, Wirsung's duct diameter, anastomosis type, or intraoperative blood loss and the occurrence of complications (Table 5).

Table 6 summarizes the distribution of the histopathologic phenotype. The most common histopathological finding was adenocarcinoma, specifically the pancreaticobiliary subtype. Among the other histopathological findings, neuroendocrine carcinoma (n=8), chronic pancreatitis (n=4), autoimmune pancreatitis (n=3), serous cystadenoma (n=2), and mucinous cystadenoma (n=1) were observed, respectively.

Table 2. Intraoperative and perioperative parameters

All cases n=151 (%)	
Pancreatic texture	
Soft	69 (45.7)
Firm	76 (50.3)
No data (missing)	6 (4.0)
Pancreatic duct diameter, mm	
<3	43 (28.5)
3<	96 (63.6)
no data (missing)	12 (7.9)
Stent	
No	79 (52.3)
Yes	71 (47.0)
no data (missing)	1 (0.7)
Type of pancreaticojejunostomy	
Blumgart	125 (82.0)
Heidelberg	26 (17.0)
Operative time, mean ± SD	227.4±48.5
Blood loss, mL	
0-299	19 (12.6)
300-750	71 (47.0)
750<	21 (13.9)
SD: Standard deviation	

DISCUSSION

In the presented study, the rate of severe postoperative complications was 25.2%. Overall pancreatic fistula rate was 35.8%, and the DGE incidence was 26.5%. Two patients were re-operated, one due to postoperative bleeding and the other due to gastroenterostomy leakage. Mortality rate was 3.3%. As a result, morbidity and mortality rates were found to be comparable to the literature data and even relatively better.

Although the complexity of surgery and the challenges in managing postoperative complications led to a reluctance toward PD, the increase in surgical experience over time and advances in postoperative patient care have reduced mortality rates to below 5%[1,2]. However, morbidity rates remain relatively high, ranging from 30% to 50%[2].

POPF after PD is the most common cause of postoperative mortality. Additionally, it contributes directly or indirectly to other morbidities such as DGE, bleeding, and sepsis[1]. Factors such as pancreatic texture, pancreatic duct diameter,

Table 3. The rates of POPF and DGE

All cases n=151 (%)	
POPF	
No	97 (64.2)
Yes	54 (35.8)
POPF grade	
Grade A	38/54 (70.3)
Grade B	16/54 (29.7)
Grade C	0
DGE	
No	111 (73.5)
Yes	40 (26.5)
DGE grade	
Grade A	23/40 (57.5)
Grade B	15/40 (37.5)
Grade C	2/40 (5)
PPH	
No	145 (96.1)
Yes	6 (3.9)
PPH grade	
Grade A	4 (2.6)
Grade B	1 (0.6)
Grade C	1 (0.6)
POPF: Post-operative pancreatic fistula, DGE: Delayed gastric emptying, PPH: Post-pancreatectomy hemorrhage	

intraoperative blood loss, and pancreaticojejunal anastomosis technique are considered predictive for POPF. Its incidence remains between 3 and 45 % at high-volume centers[11]. In the current study, the POPF rate was 35.8%. Although it may appear to be a relatively high rate, most of these fistulas were biochemical leaks not classified as clinically relevant fistulas (70.3%). The incidence of biochemical fistula in our study was consistent with, and even slightly lower than, the rates reported in the literature[10]. The rate of clinically relevant POPF with intra-abdominal collection detected on postoperative control computed tomography scans was 29.7%. Of these patients, 56.3% (n=9) underwent placement of a drainage catheter by interventional radiology. No grade C fistulas were detected in any patient. The use of standard surgical techniques by the same surgical team and a relatively short average operative time are the major factors contributing to low clinically relevant fistula rates. Additionally, the pancreatic duct size greater than 3 millimeters in 63.6% of the cases may be another important reason for this.

DGE is another common cause of morbidity after PD. Its incidence rate ranges between 10% and 60%[12]. In the presented study, the postoperative DGE rate was 26.5%, with grade B and C DGE observed in 15 and 2 patients, respectively. Of the patients with clinically relevant DGE, 38.3% underwent gastroscopy, while 14.9% required percutaneous drainage and total parenteral nutrition due to intra-abdominal

Table 4. Overall postoperative complications

All cases n=151 (%)	
Intra-abdominal abscess/collection	9 (6.0)
Chylous fistula	9 (6.0)
Bleeding	6 (4.0)
Surgical site infection	4 (2.0)
Liver failure	4 (2.0)
Biliary leakage	3 (2.0)
Sepsis	3 (2.0)
Pneumonia	1 (0.6)
Anastomotic stenosis (GJ)	1 (0.6)
Renal failure	1 (0.6)
Pulmonary embolism	1 (0.6)
Hospital stay, day, mean ± SD	8.7±4.3
Re-operation	2 (1.3)
Mortality	5 (3.3)

GJ: Gastrojejunostomy, SD: Standard deviation

Table 5. Complication rates according to intraoperative parameters

	C-D <3	3≤ C-D**	p-value
	n (%)	n (%)	
	113 (74.8)	38 (25.2)	
Pancreatic texture			0.960
Soft	52 (75.4)	17 (24.6)	
Firm	57 (75)	19 (25)	
Wirsung diameter, mm			0.71
<3	31 (72.1)	12 (27.9)	
3<	72 (75.0)	24 (25.0)	
Stent			0.261
No	56 (70.9)	23 (29.1)	
Yes	56 (78.9)	15 (21.1)	
Type of pancreaticojejunostomy			0.21
Blumgart	92 (74.8)	31 (25.2)	
Heidelberg	21 (77.8)	6 (22.2)	
Blood loss, mL			0.15
0-299	17 (89.5)	2 (10.5)	
300-750	48 (67.6)	23 (32.4)	
750<	16 (76.2)	5 (23.8)	

**C-D 3a; n=27, C-D 3b; n=2, C-D 4a; n=2 , C-D 4b; n=2, C-D 5; n=5, C-D: Clavien-Dindo classification, p: Probability

collection. Many factors are blamed among the causes of etiopathogenesis, but the exact reason remains unclear[13,14]. Previous studies have shown a significant correlation between DGE and complications such as postoperative pancreatitis, pancreatic fistula, biliary fistula, and enteric leaks. It has been suggested that the developing local or abdominal inflammation is the underlying physiopathological mechanism[14]. Considering this possible relationship, the low rates of clinically relevant pancreatic fistulas suggest a positive impact on our low-grade B and C DGE rates.

PPH is a rare but one of the serious complications after pancreatic resection. In most case series, its incidence varies between 3% and 10%[15]. It has been categorized (grades A, B, and C) by the International Study Group based on the timing (early or late), severity (mild or severe), and location of the bleeding (intraluminal or extraluminal)[16]. Potentially life-threatening bleeding is defined as grade C. Bleeding occurring within the first 24 hours after the operation is classified as early hemorrhage, usually resulting from technical issues related to the hemostasis of the vascular-rich area or due to anticoagulant medications and an underlying coagulopathy[15]. Hemorrhages occurring in the late period are typically due to erosion of vascular structures caused by POPF, most commonly arising from the stump of the gastroduodenal artery[16]. In our case series, we encountered postoperative hemorrhage in six patients, one of which required reoperation. The remaining five patients were managed conservatively. Four of them were grade A, and one was grade B. In six patients, the bleeding occurred in the early postoperative period. Therefore, secondary causes such as pancreatic fistula could be excluded. Unfortunately, one patient who was re-operated due to PPH died in the postoperative period due to disseminated intravascular coagulation.

Table 6. Histopathologic findings

All cases n=151 (%)	
Adenocarcinoma	120 (79.4)
Pancreatobiliary	71
Ductal	21
Biliary	12
Undifferentiated	4
Intestinal	2
Mucinous	3
Adenosquamous	2
Mixt	4
Tubular	1
Neuroendocrine tumor/carcinoma	8 (5.3)
Chronic pancreatitis	4 (2.6)
Autoimmune pancreatitis	3 (2.0)
GIST	2 (1.3)
IPMN	2 (1.3)
Serous cystadenoma	2 (1.3)
Mucinous cyst	1 (0.7)
*Other	9 (6.0)

*Undif sarcoma 2, carcinosarcoma 1, squamous carcinoma 1, cohesive carcinoma 1, gastric cancer 1, gastric peptic ulcus 2, acinar cell carcinoma 1

GIST: Gastrointestinal stromal tumor, IPMN: Intraductal papillary neoplasm

Obstructive jaundice is the most common symptom in patients with periampullary mass[17]. However, it is increasingly acknowledged that routine preoperative biliary drainage cannot be recommended in patients with obstructive jaundice due to the increased rate of infectious complications associated with drainage procedures in these patients. Recent studies have indicated that biliary drainage before PD should be reserved only for patients with severe and long-standing jaundice, cholangitis, renal failure, or malnourishment or with indications for neoadjuvant therapy[18,19]. In our case series, ERCP and percutaneous transhepatic cholangiography were performed pre-operatively in 47.7% and 23.2% of the patients, respectively, and the remaining underwent direct surgery without preoperative biliary drainage. In subgroup analysis, the postoperative complication rate was higher in patients who underwent ERCP compared to those who underwent percutaneous transhepatic biliary drainage (PTBD) (44.4% vs. 28.6%), and this difference did not reach statistical significance (p=0.115). We believe that the development of symptomatic or asymptomatic ascending cholangitis in patients who underwent ERCP increases the risk of postoperative infectious complications[13]. Therefore, PTBD should be considered the preferred drainage method for patients requiring preoperative biliary drainage. The lack of a statistically significant difference may be due to the limited sample size in our study.

The need for reoperation occurred in only two patients, one for previously mentioned postoperative bleeding and the other for a gastrojejunostomy anastomotic leak. Five patients died in the postoperative early period. Three of these patients died due to sepsis. In all three cases, the clinical signs of sepsis developed on the first postoperative day, and the patients died due to septic shock and multiorgan failure. The common feature of these three patients was that they had undergone preoperative endoscopic biliary drainage and had a plastic biliary stent placed in the bile duct. Another shared characteristic was that they were all elderly patients. Due to these cases, we revised our antibiotic prophylaxis protocol in the clinic. Specifically, for patients who had undergone preoperative biliary drainage, we started using broad-spectrum antibiotics covering gram-negative bacteria one day before surgery. We continued this

treatment in the postoperative period until the results of intraoperative bile cultures were available. Following the protocol change regarding the use of prophylactic antibiotics in patients undergoing preoperative biliary drainage, no cases of clinical sepsis were observed on postoperative first day. One of the other two patients died in the early postoperative period due to respiratory failure caused by pneumonia, while the other died due to bleeding.

Finally, it was found that approximately one in every four patients required intervention during the postoperative period. Therefore, it would be appropriate to recommend that this surgery be performed in reference centers with level three intensive care units and with interventional radiology to perform percutaneous drainage for postoperative abdominal collections[20].

Study Limitations

The study's limitations include selection bias because of the study design, even if we have a prospectively recorded database. Second, excluding patients with borderline and locally advanced disease who received neoadjuvant treatment, even though their numbers were relatively low, might have had a relatively positive impact on our results. Third, it was conducted at a single center, which may limit the generalizability of the findings. Institutional practices, patient populations, and perioperative management protocols may differ across centers, potentially influencing outcomes. Therefore, multicenter studies are warranted to validate these results in broader clinical settings. Fourth, the lack of long-term follow-up data precludes assessment of delayed complications, recurrence rates, and long-term efficacy of the intervention.

CONCLUSION

As a result, the conclusions drawn primarily reflect short-term outcomes. Another limitation is that this study's cross-sectional nature may reflect the relatively good results of our team after gaining some experience in this complex surgery. This study allowed us to document and report our PD results in a standardized manner and compare them with the literature. Morbidity and mortality rates were relatively better compared to the literature data. Managing complications after complex surgeries is the most important factor affecting postoperative outcomes. In addition, using the same type of surgical technique by the same surgical team and the relatively short average operative time were the primary positive contributing factors to our results.

Ethics

Ethics Committee Approval: The study protocol was approved by the Ethics Committee of Marmara University Faculty of Medicine, and was conducted by the principles of the Declaration of Helsinki (desicion no: 09.2023.892, date: 14.07.2023).

Informed Consent: Prospectively recorded patients' clinical, operative, pathologic, and short-term outcomes were analyzed retrospectively.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Jakhmola CK, Kumar A. Whipple's pancreaticoduodenectomy: outcomes at a tertiary care hospital. Med J Armed Forces India. 2014;70:321-6.
2. Cameron JL, Riall TS, Coleman J, Belcher KA. One thousand consecutive pancreaticoduodenectomies. Ann Surg. 2006;244:10-5.
3. Joliat GR, Petermann D, Demartines N, Schäfer M. Prediction of complications after pancreaticoduodenectomy: validation of a postoperative complication score. Pancreas. 2015;44:1323-8.
4. Isaji S, Mizuno S, Windsor JA, Bassi C, Fernández-Del Castillo C, Hackert T, et al. International consensus on definition and criteria of borderline resectable pancreatic ductal adenocarcinoma 2017. Pancreatology. 2018;18:2-11.
5. Tol JA, Gouma DJ, Bassi C, Dervenis C, Montorsi M, Adham M, et al. Definition of a standard lymphadenectomy in surgery for pancreatic ductal adenocarcinoma: a consensus statement by the International Study Group on Pancreatic Surgery (ISGPS). Surgery. 2014;156:591-600.
6. Wang SE, Chen SC, Shyr BU, Shyr YM. Comparison of modified Blumgart pancreaticojejunostomy and pancreaticogastrostomy after pancreaticoduodenectomy. HPB (Oxford). 2016;18:229-35.
7. Shrikhande SV, Kleeff J, Büchler MW, Friess H. Pancreatic anastomosis after pancreaticoduodenectomy: how we do it. Indian J Surg. 2007;69:224-9.
8. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg. 2004;240:205-13.
9. Wente MN, Bassi C, Dervenis C, Fingerhut A, Gouma DJ, Izbicki JR, et al. Delayed gastric emptying (DGE) after pancreatic surgery: a suggested definition by the International Study Group of Pancreatic Surgery (ISGPS). Surgery. 2007;142:761-8.
10. Bassi C, Marchegiani G, Dervenis C, Sarr M, Abu Hilal M, Adham M, et al. The 2016 update of the International Study Group (ISGPS) definition and grading of postoperative pancreatic fistula: 11 Years After. Surgery. 2017;161:584-91.
11. Fu SJ, Shen SL, Li SQ, Hu WJ, Hua YP, Kuang M, et al. Risk factors and outcomes of postoperative pancreatic fistula after pancreatico-duodenectomy: an audit of 532 consecutive cases. BMC Surg. 2015;15:34.
12. Nojiri M, Yokoyama Y, Maeda T, Ebata T, Igami T, Sugawara G, et al. Impact of the gastrojejunal anatomic position as the mechanism of delayed gastric emptying after pancreatoduodenectomy. Surgery. 2018;163:1063-70.

13. Atici AE, Ozocak AB, Kayaci AE, Ozturk EG, Kararmaz A, Yegen SC. Impact of marmara-yegen cuttıng gastrojejunostomy on delayed gastrıc emptyıng after pancreatoduodenectomy: ınıtıal results. Langenbecks Arch Surg. 2024;409:291.

14. Arora L, Reddy VV, Gavini SK, Chandrakasan C. Impact of route of reconstruction of gastrojejunostomy on delayed gastric emptying after pancreaticoduodenectomy: a prospective randomized study. Ann Hepatobiliary Pancreat Surg. 2023;27:287-91.

15. Wellner UF, Kulemann B, Lapshyn H, Hoeppner J, Sick O, Makowiec F, et al. Postpancreatectomy hemorrhage--incidence, treatment, and risk factors in over 1,000 pancreatic resections. J Gastrointest Surg. 2014;18:464-75.

16. Wente MN, Veit JA, Bassi C, Dervenis C, Fingerhut A, Gouma DJ, et al. Postpancreatectomy hemorrhage (PPH): an international study group of pancreatic surgery (ISGPS) definition. Surgery. 2007;142:20-5.

17. Sewnath ME, Karsten TM, Prins MH, Rauws EJ, Obertop H, Gouma DJ. A meta-analysis on the efficacy of preoperative biliary drainage for tumors causing obstructive jaundice. Ann Surg. 2002;236:17-27.

18. El-Haddad HM, Sabry AA, Shehata GM. Endoscopic versus percutaneous biliary drainage for resectable pancreatic head cancer with hyperbilirubinemia and impact on pancreaticoduodenectomy: A randomized controlled study. Int J Surg. 2021;93:106043.

19. Gregg JA, De Girolami P, Carr-Locke DL. Effects of sphincteroplasty and endoscopic sphincterotomy on the bacteriologic characteristics of the common bile duct. Am J Surg. 1985;149:668-71.

20. Marshall JC, Bosco L, Adhikari NK, Connolly B, Diaz JV, Dorman T, et al. What is an intensive care unit? A report of the task force of the world federation of societies of intensive and critical care medicine. J Crit Care. 2017;37:270-6.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.91489
Nam Kem Med J 2025;13(3):269-275

The Role of Pan-Immune Inflammation Value in Predicting Saphenous Vein Graft Patency After Coronary Artery Bypass Surgery

Koroner Arter Bypass Cerrahisi Sonrası Safen Ven Greft Açıklığını Öngörmede Pan-İmmün Enflamasyon Değerinin Rolü

Cihan AYDIN, Aykut DEMİRKIRAN, Hüseyin ORTA

Tekirdağ Namık Kemal University Faculty of Medicine, Department of Cardiology, Tekirdağ, Türkiye

ABSTRACT

Aim: In this study, it was aimed to investigate the predictive value of pan-immune-inflammation value (PIV) in postoperative saphenous vein graft patency (SVG).

Materials and Methods: Data of 300 patients with coronary artery bypass grafting (CABG) who underwent angiography between January 2022 and January 2024 were retrospectively analyzed. These patients were divided into two groups according to the presence of 50% or more stenosis in SVG (Group 1: SVG patent, Group 2: SVG not patent). We investigated PIV, neutrophil to lymphocyte ratio, platelet to lymphocyte ratio (PLR), lymphocyte to monocyte ratio, systemic inflammation response index, and systemic inflammation index. The study aimed to investigate whether there was a difference in inflammation indexes between the groups.

Results: The cut-off value of PIV >444 was associated with 72% sensitivity and 62% specificity to predict SVG disease (SVGD) in patients with CABG. Multivariate logistic regression analysis showed that PLR and PIV levels were independent predictors of SVGD, respectively [odds ratio (OR): 1.025; 95% confidence interval (CI): 1.008-1.042; p=0.003)] and (OR: 1.012; 95% CI: 1.000-1.015; p=0.004).

Conclusion: PIV and PLR may be useful predictors of SVGD, which can be easily estimated from blood tests in routine practice. As a novel inflammatory biomarker, PIV may serve as a valuable tool for risk stratification and long-term follow up in this patient population.

Keywords: Inflammation, coronary artery disease, coronary artery bypass grafting, saphenous vein graft, pan-immune-inflammation value

ÖZ

Amaç: Bu çalışmada, postoperatif safen ven greft (SVG) açıklığında pan-immün enflamasyon değerinin (PIV) öngördürücü değerini araştırmayı amaçladık.

Gereç ve Yöntem: Ocak 2022-Ocak 2024 tarihleri arasında koroner arter baypass greftleme (KABG) uygulanan ve anjiyografi yapılan 300 hastanın verileri retrospektif olarak incelendi. Bu hastalar SVG %50 veya daha fazla darlık olup olmamasına göre iki gruba ayrıldı (Grup 1: SVG açık, Grup 2: SVG açık değil). Çalışmada PIV, nötrofil-lenfosit oranı, trombosit lenfosit oranı (PLR), lenfosit monosit oranı, sistemik enflamasyon yanıt indeksi, sistemik enflamasyon indeksi araştırıldı. Çalışmanın amacı gruplar arasında enflamasyon indeksleri açısından fark olup olmadığını araştırmaktı.

Bulgular: PIV >444 kesme değeri, KABG öngörmek için %72 duyarlılık ve %62 özgüllük ile ilişkilendirilmiştir. Çok değişkenli lojistik regresyon analizi, PLR ve PIV düzeylerinin sırasıyla SVG hastalığı'nın (SVGD) bağımsız öngörücüleri olduğunu göstermiştir [olasılık oranı (OR): 1,025; %95 güven aralığı (GA): 1,008-1,042; p=0,003)] ve (OR: 1,012; %95 GA: 1,000-1,015; p=0,004).

Address for Correspondence: Cihan AYDIN MD, Tekirdağ Namık Kemal University Faculty of Medicine, Department of Cardiology, Tekirdağ, Türkiye
E-mail: drcihanaydin@hotmail.com **ORCID ID:** orcid.org/0000-0002-1401-5727
Received: 18.04.2025 **Accepted:** 28.05.2025 **Publication Date:** 07.10.2025

Cite this article as: Aydın C, Demirkıran A, Orta H. The role of pan-immune inflammation value in predicting saphenous vein graft patency after coronary artery by-pass surgery. Nam Kem Med J. 2025;13(3):269-275

Sonuç: PIV ve PLR, rutin uygulamada kan testlerinden kolayca tahmin edilebilen SVGD'nin yararlı öngörücüleri olabilir. Yeni bir enflamatuvar biyobelirteç olarak PIV, bu hasta popülasyonunda risk sınıflandırması ve uzun vadeli takip için değerli bir araç olarak hizmet edebilir.

Anahtar Kelimeler: Enflamasyon, koroner arter hastalığı, koroner arter bypass greftleme, safen ven greft, pan-immün-enflamatuvar indeks

INTRODUCTION

Percutaneous coronary intervention (PCI) and coronary artery bypass grafting (CABG) are considered revascularization procedures, but CABG alone is the most effective treatment in some patient groups. PCI is only intended to treat flow-limiting lesions, even though non-flow-limiting stenoses cause many infarcts. PCI cannot thus be expected to reduce the number of new infarcts considerably; but, by supplying flow distal to artery occlusions, CABG may be able to accomplish so. Long-term clinical outcomes and recurring symptoms following CABG surgery are contingent upon the patency of the bypass graft and the rate of native coronary artery disease development. The most important problem after CABG is the faster progression of atherothrombotic occlusive disease in vein grafts. Inflammatory parameters are of great importance in predicting this early atherosclerosis in saphenous vein grafts (SVG). Thus, it has been discovered in several studies that choosing an arterial graft improves long-term survival and lowers the frequency of coronary angiographic procedures[1]. In the first month following bypass surgery, thrombotic occlusion is the cause of occlusion in vein grafts; however, atherosclerosis and neointimal hyperplasia are the causes in later stages. Atherosclerosis is a chronic inflammatory vascular disease whose pathogenesis is caused by traditional and non-traditional risk factors. Genome studies have shown that innate and adaptive immune responses can promote or suppress atherosclerosis. Russell Ross originally put up the theory that atherosclerosis is an inflammatory disease in 1999 based on data showing that circulating monocytes penetrate the fatty streak as it develops[2]. Many simple markers have been studied to predict cardiovascular mortality and stent re-stenosis, which can be obtained from biochemical parameters[3,4].

The increase in inflammation risk markers such as neutrophil-tolymphocyte ratio (NLR) after CABG or PCI is useful in predicting cardiovascular mortality or in-stent restenosis[5]. Many studies have demonstrated that high NLR and platelet to lymphocyte ratio (PLR) levels correlate with the severity of coronary artery disease[6,7]. Pan-immune-inflammation value (PIV) includes more comprehensive blood parameters, which makes it a better predictor in coronary artery disease than other inflamatuar indexes. In this study, it was aimed to predict saphenous vein patency with these new inflammation markers in patients with CABG.

MATERIALS AND METHODS

Data of 300 patients with CABG who underwent angiography between January 2022 and January 2024 were retrospectively analyzed (in flow chart). A 50% or less coronary artery stenosis was considered non-critical. Patients who had previously undergone CABG were divided into two groups according to whether there was 50% or more stenosis in the SVG. Patients' medical records were reviewed for basic demographic information, coronary angiography reports, clinical history, prescription information, and blood chemistry test findings (Table 1). The following systemic inflammation indexes were calculated from whole blood assays: [systemic inflammation response index (SIRI) (neutrophils × monocytes / lymphocytes), systemic immune inflammation index (SII) (neutrophils × platelets / lymphocytes), PIV (neutrophils × monocytes × platelets) / lymphocytes), PLR (platelets / lymphocytes ratio), lymphocyte to monocyte ratio (LMR) (lymphocytes /monocytes ratio), and NLR (neutrophils / lymphocytes ratio)]. The numbers of all blood parameters were multiplied by ($\times 10^3/\mu L$). Patients with a history of acute coronary syndrome in the last 3 months and those with a recent history of PCI were excluded from the study. Blood samples of the patients were taken from the antecubital vein at first hospital admission after 12 hours of fasting before the angiography procedure. Patients with active infectious disease [white blood cell count (WBC)> 11 $\times 10^3/\mu L$ or C-reactive protein (CRP)> 5mg/dL], chronic inflammatory disease (CRP> 5mg/dL or sedimentation> 20mm/hour), or clinical evidence of cancer, severe renal disease (estimated glomerular filtration rate <30 mL/min/1.73 m^2), and hematological diseases were excluded from the study. Patients who underwent valve surgery together with CABG surgery were not included in the study. The study was authorized by Tekirdağ Namık Kemal University Local Ethics Committee (decision no: 2021.284.12.07, date: 28.12.2021) and was carried out by the Helsinki Declaration.

Angiographic Analysis

After getting each patient's informed consent, the Judkins technique was used to perform coronary angiography in normal standard projections with the required catheters. Two independent cardiologists who were blind to the patient data analyzed the coronary angiograms of 300 patients. Bypass grafts with visual stenosis of 50% or more were considered significant. These patients were divided into two groups according to the presence of 50% or more stenosis in SVG. In

Group 1, there were 150 patients with less than 50% stenosis in their bypass grafts. In Group 2, there were 150 patients with significant stenosis of over 50% in their bypass grafts.

Statistical Analysis

SPSS 22.0 statistical software (SPSS Inc., Chicago, IL) was used to analyze all of the study's data. The Kolmogorov-Smirnov test was used to evaluate data distribution. Categorical variables were reported as percentages and compared using chi-square test or Fischer's exact test while continuous variables were expressed as mean ± standard deviation or median (minimum-maximum). Continuous data conforming to normal distribution was evaluated with the Student's t-test, and data not compatible with normal distribution was evaluated with the Mann-Whitney U test. ROC curve analysis was used to determine the cut-off values for PVI and PLR to predict SVG patency. Effects of different variables on SVG patency were evaluated with univariate and Multivariate logistic regression tests. P-values <0.05 were considered statistically significant.

RESULTS

Baseline characteristics and laboratory results of the patients are summarized in Tables 1 and 2. The groups were similar in terms of demographic data except for the number of heart failures. The frequency of heart failure was higher in Group 2 (p=0.032).

Beta-blocker and angiotensin-converting enzyme inhibitors/angiotensin receptor blockers use was higher in Group 2. In terms of biochemical parameters; fasting glucose, serum creatinine, hemoglobin, hematocrit, WBC, high sensitivity CRP, neutrophil and monocyte counts, cholesterol levels were similar between the groups. While platelet levels were higher in Group 2, lymphocyte levels were found to be lower in this group. While inflammation indexes (PIV, NLR, SII, SIRI, PLR) were statistically higher in Group 2, LMR was higher in Group 1 (Table 2). In Multivariate logistic regression analysis of the independent predictors, PLR and PIV were all significantly associated with SVG disease [odds ratio (OR): 1.025; 95% confidence interval (CI): 1.008-1.042; p=0.003, OR: 1.012; 95% CI: 1.000-1.015; p=0.004, respectively] (Table 3). ROC curve analysis also showed that PIV had a sensitivity of 76% and specificity of 72% for SVG disease when the cut–off value of PIV was >444 (p<0.001). Area under the curve (AUC) (95% CI); 0.793(0.700-0.885) (Figure 1). ROC curve analysis also showed that PLR had a sensitivity of 72% and specificity of 61.5% for SVG disease when the cut–off value of PLR was >151 (p<0.001). AUC (95% CI); 0.722 (0.613-0.831) (Figure 1).

DISCUSSION

In this study, inflammatory markers such as PIV, PLR, SII, LMR, SIRI, and NLR were detected at higher levels in patients with SVG disease. Among these inflammation markers, higher PIV and PLR better predicted the development of SVG disease. PIV includes more comprehensive blood parameters, which makes it a better predictor than PLR (AUC 0.793 vs. 0.722). This suggests that PIV may be a more accurate and comprehensive index in predicting immunological and inflammatory/anti-inflammatory conditions. These findings suggest that systemic inflammation plays a central role in the pathophysiology of

Table 1. Baseline characteristics of the groups

Variables	Group 1 (n=150)	Group 2 (n=150)	p-value
Age (years)	66.4±9.6	67.4±8.1	0.582
Male, n (%)	69 (46)	70 (46.6)	0.986
Heart failure, n (%)	10 (6.6)	30 (20)	**0.032**
COPD, n (%)	12 (8)	10 (6.6)	0.232
Stroke, n (%)	6 (4)	4 (2.6)	0.123
Hypertension, % n (%)	50 (33.3)	48 (32)	0.853
Diabetes mellitus, n (%)	22 (14.6)	25 (16.6)	0.548
Medical treatment			
Beta blocker, n (%)	35 (23.3)	70 (46.6)	**0.002**
Ca-channel blocker, n (%)	15 (10)	20 (13.3)	0.724
ACE-I/ARB, n (%)	60 (40)	100 (66.6)	**0.001**
Diuretic, n (%)	26 (17.3)	36 (24)	0.051
OAD n, (%)	15 (10)	18 (12)	0.825
Insulin, n (%)	20 (13.3)	18 (12)	0.659

OAD: Oral anti-diabetic drugs, ACE-I/ARB: Angiotensin-converting enzyme inhibitors/angiotensin receptor blockers, COPD: Chronic obstructive pulmonary disease, Group 1: Saphenous vein grafts patent, Group 2: Saphenous vein grafts not patent

Table 2. Laboratory parameters of the groups

Variables	Group 1 (n=150)	Group 2 (n=150)	p-value
Glucose (mg/dL)	130 (77-291)	122 (85-349)	0.432
Hemoglobin (g/dL)	12.6±1.7	11.9±2	0.087
Hematocrit %	40.1±5	40.5±4	0.234
Platelet count (×10³/µL)	220±51	243±54	**0.047**
Serum creatinine (mg/dL)	0.81 (0.47-1.4)	0.96 (0.54-1.3)	0.894
Total cholesterol (mg/dL)	178.7±55	215.5±50	0.370
High density lypoprotein-cholesterol (mg/dL)	41.7±14	42.8±11.3	0.698
Low density lypoprotein-cholesterol (mg/dL)	102.1±44.8	103.8±54.6	0.870
Triglyceride (mg/dL)	151 (40-521)	141.9 (45-492)	0.814
High sensitivity C-reactive protein (mg/dL)	4.1 (0.2-152)	7.1 (0.32-160)	0.422
White blood cell count(×10³/µL)	7.4 (5.2-16.2)	7.2 (4.7-15.7)	0.264
Neutrophil count (×10³/µL)	4.7 (3-12.4)	5.1 (3.2-13.9)	0.631
Lymphocyte count (×10³/µL)	1.8 (0.75-3.2)	1.1 (0.54-2.5)	**<0.001**
Monocyte count (×10³/µL)	0.5 (0.3-1.2)	0.6 (0.2-1.4)	0.213
PIV (×10⁶/µL)	403.5 (98.5-2185)	631.2 (244-7603)	**<0.001**
NLR	2.2 (1.3-12.1)	2.8 (0.7-25)	**0.027**
SII (×10³/µL)	641 (300-2335)	1066 (458-6980)	**<0.001**
SIRI (×10³/µL)	1.8 (0.58-10.4)	2.7 (1.1-30)	**0.002**
PLR	119 (51-280)	220 (79-501)	**<0.001**
LMR	3.9 (1.1-6.1)	1.9 (0.5-4.9)	**<0.001**

PIV: Pan-immune-inflammation value, NLR: Neutrophil-tolymphocyte ratio, SII: Systemic immune inflammation index, SIRI: Systemic inflammation response index, PLR: Platelet to lymphocyte ratio, LMR: Lymphocyte to monocyte ratio, Group 1: Saphenous vein grafts patent, Group 2: Saphenous vein grafts not patent

Table 3. Univariate and multivariate logistic regression analysis of the independent predictors of late saphenous vein graft disease

	Univariate analysis			Multivariate analysis		
Variables	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
Hypertension	1.027	1.014-1.041	0.460	-	-	-
Diabetes mellitus	0.773	0.333-1.792	0.548	-	-	-
SII	1.001	1.000-1.002	**0.007**	0.998	0.994-1.001	0.998
SIRI	1.226	0.981-1.533	0.073	-	-	-
PLR	1.017	1.009-1.025	**<0.001**	1.025	1.008-1.042	**0.003**
LMR	0.433	0.274-0.685	**<0.001**	0.772	0.410-1.455	0.424
NLR	1.001	0.884-1.134	0.982	-	-	-
PIV	1.001	1.000-1.002	**0.030**	1.012	1.000-1.015	**0.004**

PIV: Pan-immune-inflammation value, NLR: Neutrophil-tolymphocyte ratio, SII: Systemic immune inflammation index, SIRI: Systemic inflammation response index, PLR: Platelet to lymphocyte ratio, LMR: Lymphocyte to monocyte ratio, OR: Odds ratio, CI: Confidence interval

SVG failure and that PIV could serve as a novel biomarker for assessing long-term graft patency. Today, CABG is the first option for left main coronary artery lesions and multivessel coronary artery diseases. However, the rapid progression of the atherothrombotic process in vein grafts compared to arterial grafts is the most important problem after CABG. Conditions such as hypertension, diabetes, and obesity cause damage to the SVG wall, resulting in increased expression and secretion of proinflammatory cytokines that promote the formation of atheromatous plaques. It is known that the use of arterial grafts increases the patency rate of the grafts and reduces the frequency of angina and re-intervention rates after CABG. Therefore, the use of the left internal mammary artery (LIMA) is the gold standard in left anterior descending artery revascularization. Venous grafts are easily accessible and do not spasm. However, it also has disadvantages such as

having valves, limited adaptation to arterial pressure, diameter incompatibility, and rapidly progressing atherosclerosis. Early graft occlusion can be caused by a variety of factors, including the choice of grafts, the site of the anastomosis, severe graft stretching, and inflammation.

While LIMA has an 85% 10-year patency rate, SVG has a 61% rate[8]. When we examine the mechanisms responsible for SVG occlusion, thrombosis in the first month after surgery, neointimal hyperplasia between 1-12 months, and atherosclerosis after the 12th month is at the forefront[9]. Inflammation is known to be involved in all stages of atherosclerotic diseases[10]. The process is mostly initiated by monocytes and lymphocytes secreting growth hormones and cytokines including platelet derived growth factor, interleukin-6, and interleukin-1. After entering the subendothelial layer to phagocytose oxidized low density lipoprotein particles, monocytes undergo a metamorphosis into foamy cells and contribute to the creation of the central region of atheroma plaque. High concentrations of monocytes and WBC have been related to an increased risk of coronary artery disease, according to Olivares et al.[11]. Additionally, in cases of acute coronary syndrome, de LMR creased lymphocyte counts have been related to poor cardiovascular outcomes[12]. We demonstrated, in our study, that two indicators that may be utilized to predict the patency of SVG are low levels of LMR and high levels of NLR. Supporting our study, previous studies have found that low LMR levels are associated with poor endpoints in both critical limb ischemia and in-stent restenosis[13,14]. PIV was related to long-term mortality in ST-elevation myocardial infarction, according to research by Murat et al[15]. A relationship was found between PIV and no-reflow phenomenon in patients who underwent PCI after ST-segment elevation myocardial infarction (STEMI)[16]. In different studies, PIV has been found to be a good inflammatory index in predicting cardiovascular adverse events after STEMI and prognosis in heart failure[17,18]. To our knowledge, this is one of the few studies to investigate PIV in the context of late SVG patency. Our results complement prior studies, which have demonstrated associations between systemic inflammation and vein graft failure, though relying on simpler markers like CRP or WBC count[19,20] By contrast, PIV offers a more nuanced measure that reflects both innate and adaptive immune activity. This could explain its superior predictive value for SVG stenosis, as inflammation is involved in every phase of vein graft atherosclerosis from early thrombus formation to late plaque progression. PIV has also been associated with poor coronary collateral circulation[21]. It has been demonstrated that some inflammatory markers are associated with major amputation in peripheral arterial diseases[22]. Furthermore, our findings may have clinical implications. Identifying patients with elevated PIV or PLR prior to or after CABG could help in risk stratification, enabling more intensive monitoring, pharmacological intervention, or consideration of arterial graft alternatives. Statin therapy, for instance, has been shown to reduce inflammatory markers and improve SVG outcomes[23].

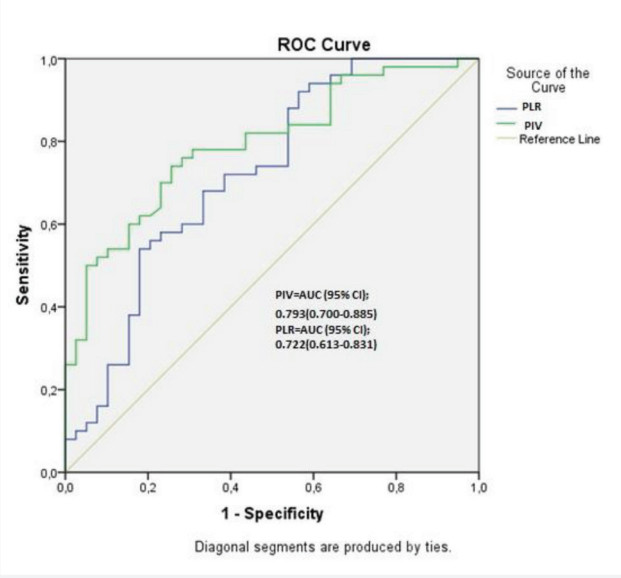


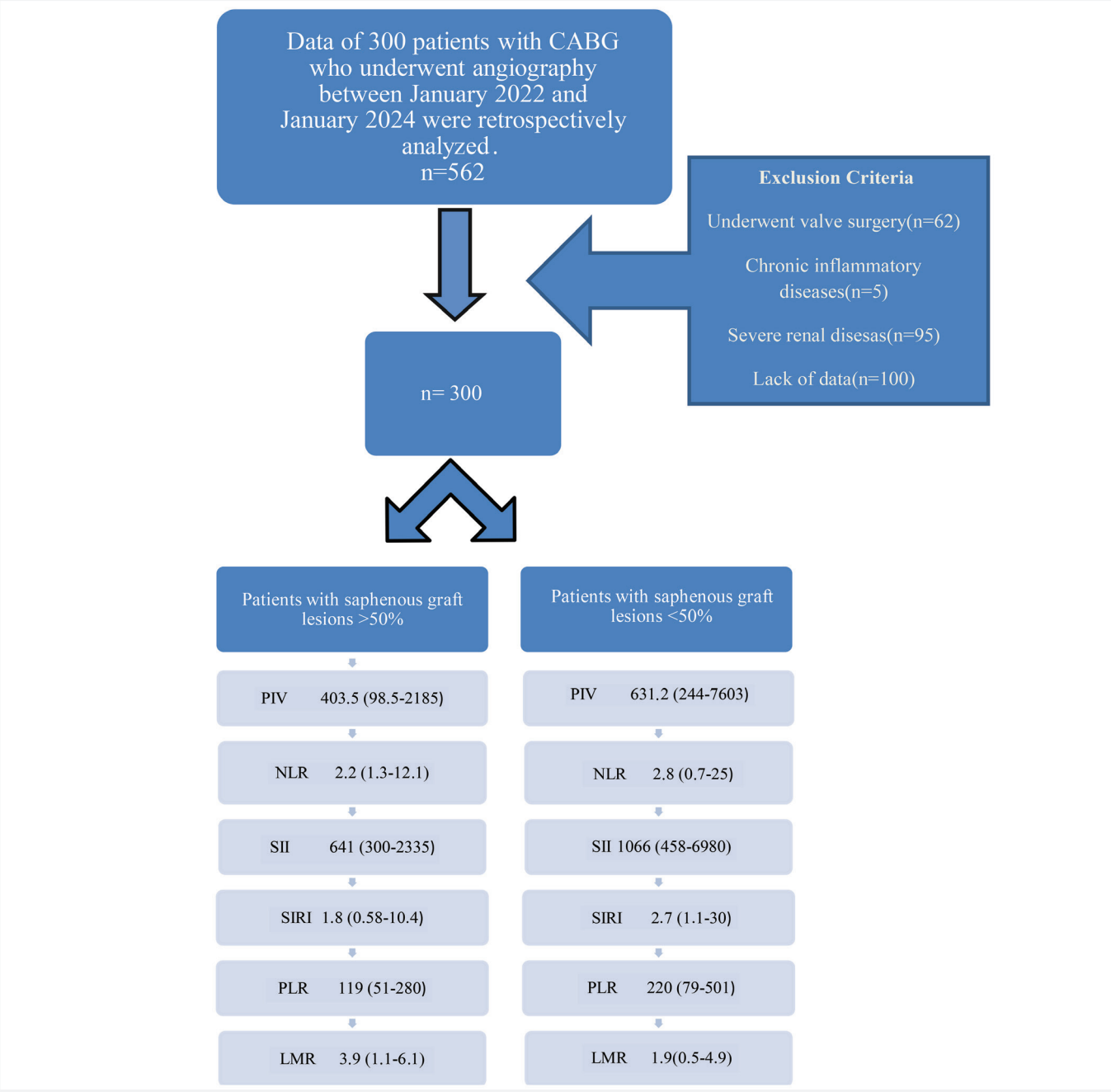
Figure 1. Receiver-operating characteristic curve analysis of platelet to lymphocyte ratio lymphocyte ratio and pan-immune-inflammation value for the prediction of saphenous vein graft disease after surgery

PLR: Platelet to lymphocyte ratio, PIV: Pan-immune-inflammation value, AUC: Area under the curve, CI: Confidence interval

In the present study, we investigated many indexes that may be used to predict late SVG occlusion. In Multivariate logistic regression analysis, we found that PLR and PIV were independent predictors of SVG. Whether PIV-guided anti-inflammatory strategies can further enhance graft patency warrants prospective evaluation.

Study Limitations

This study has several limitations that should be acknowledged. First, it is a single-center, retrospective study, which may limit the generalizability of the results to broader populations. We were unable to obtain detailed data on statin use and smoking. Second, although we attempted to exclude patients with active infections, malignancies, or chronic inflammatory conditions, subclinical inflammation or undiagnosed conditions may have influenced systemic inflammatory markers. Third, the observational nature of the study precludes establishing a causal relationship between inflammatory markers such as PIV and PLR and SVG disease. Fourth, the study population size, while adequate for preliminary analysis, may still be insufficient to detect small but clinically significant associations. Additionally, the inflammatory markers were measured only once before angiography; serial measurements over time could



Flow Chart

CAGB: Coronary artery bypass grafting, PIV: Pan-immune-inflammation value, NLR: Neutrophil-tolymphocyte ratio, SII: Systemic immune inflammation index, SIRI: Systemic inflammation response index, PLR: Platelet to lymphocyte ratio, LMR: Lymphocyte to monocyte ratio

provide a more dynamic and accurate understanding of the inflammatory state. Although angiographic assessment was performed by two blinded cardiologists, visual estimation of stenosis can still be subject to interobserver variability, and no intravascular imaging modalities were used to confirm graft patency or characterize plaque morphology. Another limitation of our study is that it is not precisely stated when saphenous graft disease occurs since these patients were not followed regularly after CABG surgery.

CONCLUSION

We showed in this study that two of the novel inflammatory indexes, PIV and PLR, may be significant predictors of saphenous graft patency following coronary by-pass surgery.

SVG disease can be predicted from routine blood tests that can be calculated simply and practically. In clinical practice, these indexes may help specialists in predicting saphenous graft stenosis in patients with CABG.

Ethics

Ethics Committee Approval: The study was authorized by Tekirdağ Namık Kemal University Local Ethics Committee (decision no: 2021.284.12.07, date: 28.12.2021) and was carried out by the Helsinki Declaration.

Informed Consent: It is a single-center, retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: C.A., Concept: A.D., Design: H.O., Data Collection or Processing: C.A., Analysis or Interpretation: A.D., Literature Search: H.O., Writing: C.A.

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

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

REFERENCES

1. Nwaejike N, Tennyson C, Mosca R, Venkateswaran R. Reusing the patent internal mammary artery as a conduit in redo coronary artery bypass surgery. Interact Cardiovasc Thorac Surg. 2016;22:346-50.
2. Ross R. Atherosclerosis--an inflammatory disease. N Engl J Med. 1999;340:115-26.
3. Karadeniz M, Aydın C, Demirkıran A, Alp Ç. Predictors of severity of coronary artery disease in patients with acute st-elevation myocardial ınfarction. Namik Kemal Med J. 2024;12:171-5.
4. Bagger JP, Zindrou D, Taylor KM. Leukocyte count: a risk factor for coronary artery bypass graft mortality. Am J Med. 2003;115:660-3.
5. Turak O, Ozcan F, Isleyen A, Tok D, Sokmen E, Buyukkaya E, et al. Usefulness of the neutrophil-to-lymphocyte ratio to predict bare-metal stent restenosis. Am J Cardiol. 2012;110:1405-10.
6. Aydın C, Uyan U, Karadeniz M, Demirkıran A. Role of simple inflammatory parameters in predicting the severity of coronary artery disease. Rev Assoc Med Bras (1992). 2023;69:e20230518.
7. Kurtul S, Sarli B, Baktir AO, Demirbas M, Saglam H, Doğan Y, et al. Neutrophil to lymphocyte ratio predicts SYNTAX score in patients with non-ST segment elevation myocardial infarction. Int Heart J. 2015;56:18-21.
8. Goldman S, Zadina K, Moritz T, Ovitt T, Sethi G, Copeland JG, et al. Long-term patency of saphenous vein and left internal mammary artery grafts after coronary artery bypass surgery: results from a department of veterans affairs cooperative study. J Am Coll Cardiol. 2004;44:2149-56.
9. Parang P, Arora R. Coronary vein graft disease: pathogenesis and prevention. Can J Cardiol. 2009;25:e57-62.
10. 10. Oude Nijhuis MM, van Keulen JK, Pasterkamp G, Quax PH, de Kleijn DP. Activation of the innate immune system in atherosclerotic disease. Curr Pharm Des. 2007;13:983-94.
11. Olivares R, Ducimetière P, Claude JR. Monocyte count: a risk factor for coronary heart disease? Am J Epidemiol. 1993;137:49-53.
12. Bian C, Wu Y, Shi Y, Xu G, Wang J, Xiang M, et al. Predictive value of the relative lymphocyte count in coronary heart disease. Heart Vessels. 2010;25:469-73.
13. Gary T, Pichler M, Belaj K, Eller P, Hafner F, Gerger A, et al. Lymphocyte-to-monocyte ratio: a novel marker for critical limb ischemia in PAOD patients. Int J Clin Pract. 2014;68:1483-7.
14. Murat SN, Yarlioglues M, Celik IE, Kurtul A, Duran M, Kilic A, et al. The relationship between lymphocyte-to-monocyte ratio and bare-metal stent ın-stent restenosis in patients with stable coronary artery disease. Clin Appl Thromb Hemost. 2017;23:235-40.
15. Murat B, Murat S, Ozgeyik M, Bilgin M. Comparison of pan-immune-inflammation value with other inflammation markers of long-term survival after ST-segment elevation myocardial infarction. Eur J Clin Invest. 2023;53:e13872.
16. Bayramoğlu A, Hidayet Ş. Association between pan-immune-inflammation value and no-reflow in patients with ST elevation myocardial infarction undergoing percutaneous coronary intervention. Scand J Clin Lab Invest. 2023;83:384-9.
17. Liu Y, Liu J, Liu L, Cao S, Jin T, Chen L, et al. Association of systemic ınflammatory response ındex and pan-ımmune-ınflammation-value with long-term adverse cardiovascular events in ST-segment elevation myocardial ınfarction patients after primary percutaneous coronary ıntervention. J Inflamm Res. 2023;16:3437-54.
18. Inan D, Erdogan A, Pay L, Genc D, Demırtola AI, Yıldız U, et al. The prognostic impact of inflammation in patients with decompensated acute heart failure, as assessed using the pan-immune inflammation value (PIV). Scand J Clin Lab Invest. 2023;83:371-8.
19. Yayla C, Gayretli Yayla K. C-Reactive protein to albumin ratio in patients with saphenous vein graft disease. angiology. 2021;72:770-5.
20. Yavuz S, Engin M, Yazgan E, Demir OF, Turk T. Letter: potential predictors of saphenous vein graft disease after coronary artery bypass operations. Angiology. 2022;73:689-90.
21. Yilmaz Y, Kelesoglu S. The Importance of Pan-immune inflammation value (PIV) in predicting coronary collateral circulation in stable coronary artery patients. Angiology. 2024:33197241258529.
22. Çakmak AC, Vatan MB, Sarıbıyık Çakmak B, Erkin A, Tatlı E, Kocayigit I. The systemic immune-inflammation index is an independent predictive factor in predicting major amputation in chronic limb-threatening ischemia. Vascular. 2025;0
23. Serruys PW, de Feyter P, Macaya C, Kokott N, Puel J, Vrolix M, et al. Fluvastatin for prevention of cardiac events following successful first percutaneous coronary intervention: a randomized controlled trial. JAMA. 2002;287:3215-22.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.68926
Nam Kem Med J 2025;13(3):276-286

Video-Assisted Thoracoscopic Sleeve Lobectomy Versus Open Approach in Non-Small Cell Lung Cancer: A Single-Center Study with Propensity Score Matching

Küçük Hücreli Dışı Akciğer Kanserinde Video Yardımlı Torakoskopik Sleeve Lobektomi ve Açık Yaklaşım: Propensity Skor Analizi ile Tek Merkezli Çalışma

Ece Yasemin DEMİRKOL, Volkan ERDOĞU, Melike ÜLKER, Merve EKİNCİ FİDAN, Nisa YILDIZ, Ezgi KILIÇASLAN, Celal Buğra SEZEN, Muzaffer METİN

University of Health Sciences Türkiye, Yedikule Chest Diseases and Thoracic Surgery Training and Research Hospital, Clinic of Thoracic Surgery, İstanbul, Türkiye

ABSTRACT

Aim: It was aimed to compare the outcomes of sleeve lobectomy with video-assisted thoracoscopic surgery (VATS) and thoracotomy in patients with centrally located non-small cell lung cancer (NSCLC).

Materials and Methods: Between January 2020 and February 2024, 127 patients who underwent sleeve lobectomy for NSCLC were retrospectively analyzed. Thoracotomy was used in 105 (82.6%) sleeve lobectomy cases while VATS was used in 22 (17.4%) cases. Subgroups were created according to pathology stages using propensity score analysis. Both groups were compared in terms of perioperative and early postoperative complications.

Results: No significant differences were found between the thoracotomy and VATS groups in operation time or perioperative bleeding (4.3±0.9 hours vs 4.5±0.9, 467±385 mL vs 370±70 mL, p=0.474, 0.525, respectively). However, drainage time and hospital stay were significantly shorter in the VATS group (4.5±4 days vs 3.6±3.3 days, 7.1±7.9 days vs 5.1±3.4 days, respectively, (p=0.014, 0.005). In terms of oncological principles, there was no statistically significant difference between the groups regarding the number of sampled lymph nodes, pathological tumor sizes, pathological stages, and histopathological cell types (p=0.349, 0.106, 0.709, 0.066, respectively). There was no significant difference between the groups in terms of early postoperative complications (30.5% vs 40.9%, p: 0.341). After propensity score analysis, it was found that the VATS group had shorter drainage and hospital stay (p=0.023, 0.043, respectively).

Conclusion: In NSCLC cases, sleeve lobectomies performed with the VATS approach are superior to open surgery with shorter drainage times and hospital stays without compromising oncological principles.

Keywords: Video-assisted thoracoscopic surgery, lung cancer, sleeve resection

ÖZ

Amaç: Santral yerleşimli küçük hücreli dışı akciğer kanseri (KHDAK) olgularında video yardımlı torakoskopik cerrahi (VATS) ile sleeve lobektomi ve torakotomi ile sleeve lobektomi yaklaşımında sonuçlarımızı karşılaştırdık.

Gereç ve Yöntem: Ocak 2020 ve Şubat 2024 yılları arasında, KHDAK tanısıyla sleeve lobektomi yapılan 127 olgu retrospektif olarak incelendi. Sleeve lobektomi yapılan olguların 105'ine (%82,6) torakotomi, 22 (%17,4) olguya VATS uygulandı. Alt gruplar, patoloji evrelerine göre propensity skor analizi ile oluşturuldu. Her iki grup perioperatif ve erken postoperatif komplikasyonlar açısından karşılaştırıldı.

Bulgular: Torakotomi grubu ve VATS grubu arasında, operasyon süresi ve peroperatif kanama miktarları açısından istatistiksel olarak anlamlı bir fark görülmedi (4,3±0,9 saat vs 4,5±0,9, 467±385 mL vs 370±70 mL, sırasıyla p=0,474, 0,525). Fakat, VATS grubunda, drenaj süresi ve hastanede

Address for Correspondence: Melike ÜLKER MD, University of Health Sciences Türkiye, Yedikule Chest Diseases and Thoracic Surgery Training and Research Hospital, Clinic of Thoracic Surgery, İstanbul, Türkiye
E-mail: melike.ulker@hotmail.com **ORCID ID:** orcid.org/0000-0003-1739-8329
Received: 27.03.2025 **Accepted:** 10.06.2025 **Publication Date:** 07.10.2025

Cite this article as: Demirkol EY, Erdoğu V, Ülker M, Ekinci Fidan M, Yıldız N, Kılıçaslan E, et al. Video-assisted thoracoscopic sleeve lobectomy versus open approach in non-small cell lung cancer: a single-center study with propensity score matching. Nam Kem Med J. 2025;13(3):276-286

kalış süresi istatistiksel olarak anlamlı şekilde daha kısaydı (4,5±4 gün vs 3,6±3,3 gün, 7,1±7,9 gün vs 5,1±3,4 gün, sırasıyla p=0,014, 0,005). Onkolojik prensipler açısından gruplar arasında; örneklenen lenf nodu sayısı, patoloji tümör boyutları, patolojik evreler ve histopatolojik hücre tipleri arasında, gruplar arası istatistiksel anlamlı fark saptanmadı (sırasıyla p=0,349, 0,106, 0,709, 0,066). Gruplar arasında, postoperatif erken dönem komplikasyonlar açısından anlamlı fark yoktu (%30,5 vs % 40,9, p: 0,341). Propensity skor analizi sonrası gruplar arasında VATS grubunda daha kısa süre drenaj ve hastane yatışı olduğu saptandı (p=0,023, 0,043).

Sonuç: KHDAK olgularında VATS yaklaşımı ile yapılan sleeve lobektomiler, onkolojik prensiplerden ödün vermeden, daha kısa drenaj süreleri ve daha kısa hastanede kalış süreleri ile açık cerrahiye üstünlük göstermektedir.

Anahtar Kelimeler: Video yardımlı torakoskopik cerrahi, akciğer kanseri, sleeve rezeksiyon

INTRODUCTION

Surgery is the accepted treatment modality for early and locally advanced non-small cell lung cancer (NSCLC)[1]. The standard approach in centrally located tumors has been pneumonectomy for years[2]. Since 1956, when Dr. Thomas first described sleeve resection in a case of lung cancer, this procedure has been accepted as an alternative to pneumonectomy for centrally located tumors, preserving more lung parenchymal tissue without compromising oncological principles[3]. Bronchoplastic resections are required in 3%-19% of NSCLC cases[4]. Sleeve resections have been performed with open surgery in experienced centers for years since they are technically challenging and require significant experience in operations. As a result of the experience developed in minimally invasive surgery in recent years, complex and extended surgeries such as sleeve resection can be performed by experienced centers and physicians with video-assisted thoracoscopic surgery (VATS) approaches[5].

In our study, we compared the perioperative and early postoperative results of patients who underwent VATS sleeve lobectomy for NSCLC with the open surgical method in terms of the safety of the method and its compliance with oncological principles.

MATERIALS AND METHODS

Patient Selection

We retrospectively analyzed 127 patients who underwent sleeve lobectomy for lung cancer between January 2020 and February 2024. To determine the compliance of VATS sleeve resections with oncological principles, patients who underwent sleeve resection for non-tumor reasons were not included in the study. The study was approved by the ethics/scientific committee of University of Health Sciences Türkiye, Yedikule Chest Diseases and Thoracic Surgery Training and Research Hospital and (decision no: 2023-462.28, date: 12.2023) was conducted by the principles of the Declaration of Helsinki.

Surgical Technique

All cases were evaluated with positron emission-computed tomography (PET-CT) in the preoperative period. Mediastinoscopy or endobronchial ultrasonography was performed for mediastinal staging in cases with no distant metastasis detected on PET-CT and suspicious mediastinal lymph node invasion findings. In patients with single and non-bulky N2, 3 cycles of platinum-based chemotherapy (3 cycles of nivolumab were added to neoadjuvant CT in 3 patients in the VATS group) and surgery was performed in patients with downstaging in terms of N2. Direct surgery was performed in cases without suspicion of mediastinal lymph node invasion. In cases with multiple N2, bulky N2, or no downstage after neoadjuvant treatment, surgery was not considered, and these cases were referred to oncology clinics for definitive treatment. T0 cases reflect complete pathological response post-neoadjuvant the therapy.

Indications for sleeve resection were often established preoperatively by radiological and bronchoscopic evaluations. However, in some cases, it was also decided based on perioperative findings during the operation. Indications for sleeve resection were tumor extension from the lobe bronchus to the secondary carina, progression to the main bronchus and/or invasion of the secondary carina outside the lobe bronchus by a metastatic interlobar (#11) lymph node.

Bronchovascular (double sleeve) lobectomy was performed in cases with simultaneous pulmonary artery invasion and carinal sleeve lobectomy in cases invading the carina. These cases were defined as extended sleeve resections in this study.

Bronchopleural fistulas (BPF) that developed after sleeve anastomosis were defined as early BPF if they developed within the first 7 days, intermediate BPF if they developed between 7-30 days, and late BPF if they developed >30 days later. Postoperative complications were graded according to the extended Clavien-Dindo classification of surgical complications established by the Japan Clinical Oncology Group.

In this study, VATS sleeve resections performed between 2020-2022 were defined as the first period while VATS sleeve resections performed after January 2023 were defined as the second period. This process was carried out by a single surgical team, utilizing an established and experienced team for VATS lobectomy and VATS segmentectomy procedures. The initiation of extended surgical procedures was used to divide the timeline into two periods, and a retrospective evaluation was conducted to compare the first and second periods.

Technical Consideration in VATS Sleeve Lobectomy

To provide more space for manipulation and to avoid unnecessary traction of the anastomosis, a systematic lymphadenectomy was routinely performed before the anastomosis. During anastomosis, care should be taken to ensure that the cartilage and membranous faces are opposite. A continuous suture technique with a 3/0 prolene needle is used in our clinic. The frequently preferred anastomosis technique is to start suturing from the proximal bronchus from the inside out at the junction of the cartilage and membranous structure. To avoid tangling the sutures during the procedure, it is beneficial to pass this first suture through the parietal pleura of the chest wall and out of the thorax and continue the anastomosis with the other needle tip. After anastomosis, air leakage is checked, and the surgeon may place additional sutures if deemed necessary. Especially in cases where sleeve resection is performed after neoadjuvant treatment, the anastomosis line can be supported with a live flap (parietal pleura, pericardial fatty tissue) if the surgeon deems it necessary. Images of the patient on whom we performed VATS right upper sleeve lobectomy are shown in Figure 1.

Statistical Analysis

While the data were analyzed retrospectively through patient files, there was no missing data since the patients in the study belonged to the last years. Windows Office Excel 2020 and Word 2019 versions were used to create the database. IBM SPSS Statistics Version 26 program was used for statistical calculations. The descriptive results of the study are presented together with the corresponding percentages in the case of nominal or ordinal variables. Continuous variables are presented with mean and standard deviation values. Pearson's chi-square test and Fisher's exact test were used for categorical variables, and the Mann-Whitney U test was used for non-categorical variables. Propensity score analysis was used to make peer groups according to pathological staging. The data were matched in a 1:1 ratio using logistic regression with the nearest method. "p" value below 0.05 was considered significant.

RESULTS

In 105 (82.6%) sleeve resection cases, posterolateral/anterolateral thoracotomy was performed (Thoracotomy group), while sleeve resection was performed by VATS in 22 (17.4%) cases (VATS group). While 22.7% (n=5) of VATS cases were performed uniportal, 77.3% (n=17) cases were performed with the biportal technique. Among 32 cases in which surgery was initiated with a VATS approach, conversion to thoracotomy was required in 10 cases and the conversion rate was 10/32 (31.2%). The reasons for conversion to thoracotomy were technical difficulty in five cases, severe adhesion in two cases and vascular hemorrhage in three cases.

Mean age of all patients was 58.6±11.6 years (range: 16-78), and the majority of the patients were male (n=106, 83.5%). There were no statistically significant differences between the groups in age, gender, smoking and forced vital capacity in 1st second (p=0.339, 0.390, 0.894, 0.087, respectively).

The operation time was 4.3±0.9 (range: 2.5-6) hours in the Thoracotomy group and 4.5±0.9 (range: 3-6) hours in the VATS group. Perioperative bleeding was 467±385 mL (range: 350-3000 mL) in the Thoracotomy group and 370±70 mL (range: 350-650 mL) in the VATS group. No statistically significant difference was found between the groups (p: 0.474, 0.525, respectively). Perioperative bleeding exceeding 2 liters was observed in a single case and was attributed to vascular hemorrhage in the Thoracotomy group. In the Thoracotomy group, drainage time was 4.5±4 days (range: 1-30 days) and hospital stay was 7.1±7.9 days (range: 2-68 days), while in the VATS group, drainage time was 3.6±3.3 days (range: 2-18 days) and hospital stay was 5.1±3.4 days (range: 3-19 days). These differences between the groups were statistically significant (p: 0.014, 0.005, respectively) (Figure 2). The demographic characteristics of the patients and details of the preoperative and postoperative processes are summarized in Table 1.

There was no statistically significant difference between the groups in the number of sampled lymph nodes, pathological tumor sizes, pathological stages, and histopathological cell types (p=0.349, 0.106, 0.709, 0.066, respectively). The oncological features of the cases are provided in detail in Table 2. In both the Thoracotomy group (n=1, 0.9%) and VATS Group (n=1, 4.5%), the bronchial surgical margin was reported as R1 in two cases. In the Thoracotomy group, the bronchial surgical margin was identified as R1 in one case, and due to the detection of perioperative N2 positivity, further extension of the surgical margin was not performed. In the VATS group, the R1 margin was accepted owing to the patient's history of contralateral thoracic surgery, which limited further resection.

Between January 2020 and December 2022, 6 (50%) of the 12 patients started with VATS in the period we defined as the first period were switched to open. BPF was developed in one case (16.7%). VATS was not performed after neoadjuvant treatment in this period. In the first period, only 6.1% (n=6) of the patients who underwent sleeve resection in our clinic were performed via VATS.

In the second period after January 2023, when our experience developed, 20 cases were started with VATS and only 4 cases (20%) were switched to open. Out of 16 cases completed with VATS, BPF developed in only 2 cases (12.5%). In 3 cases (18.7%), VATS sleeve resection was performed after neoadjuvant treatment. With increasing experience, 3 sleeve lower bilobectomies and one case of broncho-vascular sleeve resection were performed in this period. In 55% (n=16) of the

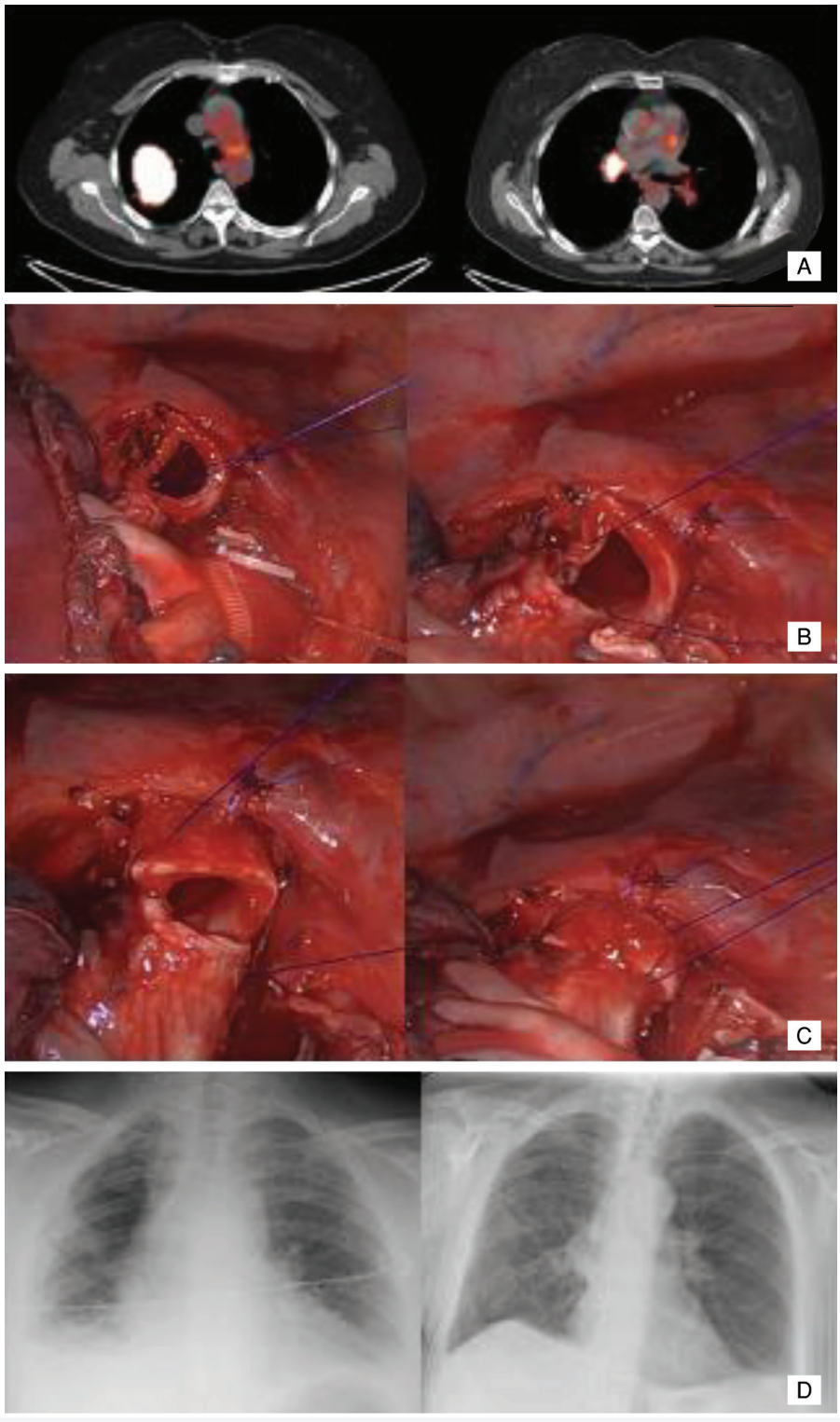


Figure 1. A) The patient, who was evaluated due to a complaint of cough, was found to have a right upper lobe mass. Subsequently, a PET-CT scan was performed, and the sections demonstrating the involvement of the mass and the right hilar lymph node are presented. B, C) After transection of the right upper lobe bronchus and removal of the specimen, the perioperative images demonstrate the anastomosis of the right main bronchus to the interlobar bronchus, initiated from the posterior membranous wall. D) Postoperative first-day and postoperative fifth-day chest radiographs are shown in the images

PET-CT: Positron emission-computed tomography

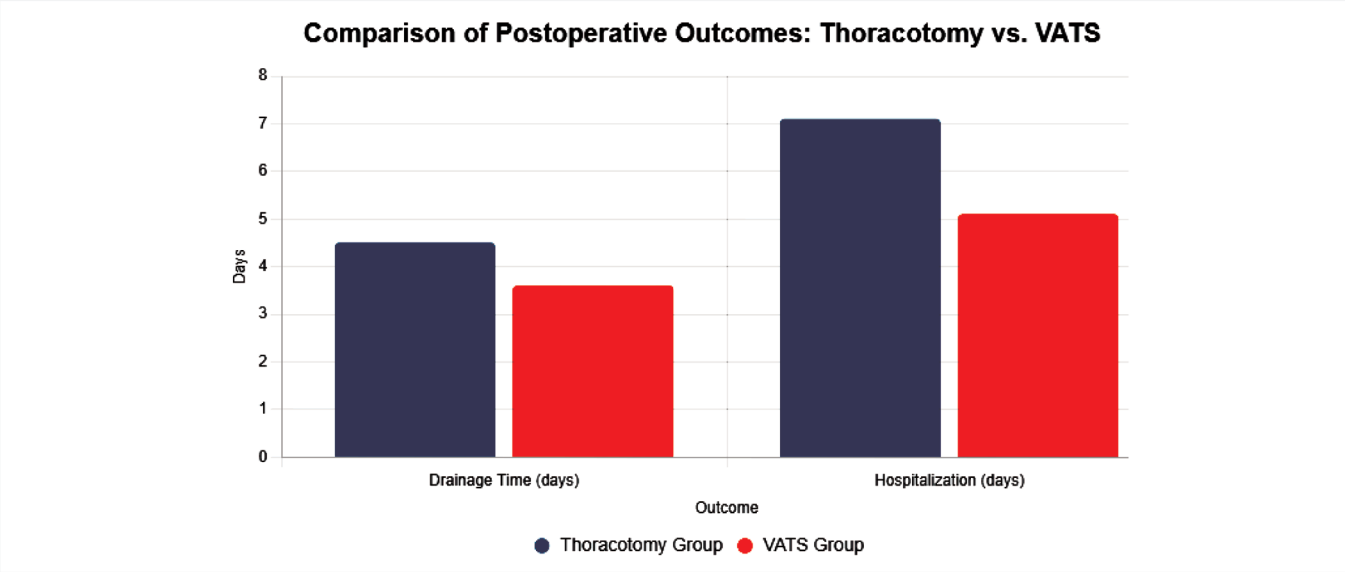


Figure 2. Comparison of postoperative outcomes

In the thoracotomy group, drainage time was 4.5 days, while in the VATS group, drainage time was 3.6 days (p=0.014). In the thoracotomy group, hospital stay was 7.1 days, while in the VATS group, hospital stay was 5.1 days (p=0.005)

VATS: Video-assisted thoracoscopic surgery

Table 1. Demographic characteristics of the patients, preoperative and postoperative process details

Variables	Units	All patients (n=127)	Thoracotomy group (n=105)	VATS group (n=22)	p-value
Age ± SD	Years	58.6±11.6 (R: 16-78)	58.2±11.5 (R: 16-77)	60.7±11.9 (R: 27-78)	p=0.339
Male (n/%) Female (n/%)	Sex	106 (83.5%) 21 (16.5%)	89 (84.8%) 16 (15.2%)	17 (77.3%) 5 (22.7%)	p=0.390
Smoking ± SD	Pack/year	30.7±16.5 (R: 0-50)	30.9±16.4 (R: 0-50)	30±17.4 (R: 0-50)	p=0.894
FEV1 (n/%) <50 50-60 60-80 >80	%	 7 (5.5%) 18 (14.2%) 59 (46.5%) 43 (33.9%)	 5 (4.8%) 16 (15.2%) 53 (50.5%) 31 (29.5%)	 2 (9.1%) 2 (9.1%) 6 (27.3%) 12 (54.5%)	p=0.087
Operation time ± SD	Hours	4.4±0.9 (R: 2.5-6)	4.3±0.9 (2.5-6)	4.5±0.9 (3-6)	p=0.474
Amount of perioperative bleeding ± SD	mL	450±353 (R: 350-3000)	467±385 (350-3000*)	370±70 (350-650)	p=0.525
Drainage time ± SD	Days	4.4±3.9 (R: 1-30)	4.5±4 (1-30)	3.6±3.3 (2-18)	**p=0.014**
Duration of hospitalization ± SD	Days	6.7±7.3 (R: 2-68)	7.1±7.9 (2-68)	5.1±3.4 (3-19)	**p=0.005**

*Perioperative bleeding exceeding 2 liters was observed in a single case and was attributed to vascular hemorrhage in the thoracotomy group

FEV1: 1. forced vital capacity per second, n: Number, R: Range, SD: Standard deviation, VATS: Video-assisted thoracoscopic surgery

patients who underwent sleeve resection in our clinic during the second period, the procedure was performed by VATS. The periodical development of our clinic in VATS sleeve resections is summarized in Table 3.

Complications were seen in 41 (32.3%) of all sleeve resection cases. Early postoperative complications were observed in 32 (30.5%) patients in the Thoracotomy group and 9 (40.9%) patients in the VATS group. No significant difference was observed between the groups in terms of complication development (p: 0.341). In the Thoracotomy group, 12 (11.4%) cases were revised early or late, while 4 (18.2%) cases were revised in the VATS group (p=0.477). Postoperative 30-day mortality was 2 (1.9%) in the Thoracotomy group, while there was no early mortality in the VATS group (p=0.682).

Complication details and 30-day mortality are summarized in Table 4. The characteristics of the patients after propensity score analysis are shown in Table 5 and the VATS group had shorter drainage and hospital stay (p=0.023, 0.043, respectively).

DISCUSSION

Sleeve resections have been performed as parenchyma-sparing surgery in lung cancer treatment with the open surgical method by experienced centers for years[6]. Protection of patients from pneumonectomy brings with it advantages such as better quality of life, better survival, lower mortality, morbidity, and recurrence rates[6-8].

The VATS approach in thoracic surgery has been shown to reduce morbidity, shorten drainage duration, and decrease the length of hospital stay. Additionally, it contributes to improved postoperative quality of life[8]. However, the place and

Table 2. Oncological features of the cases

Variables	All patients (127) n (%)	Thoracotomy (105) n (%)	VATS (22) n (%)	p-value
Resection side				
Right	84 (66.1%)	71 (67.6%)	13 (59.1%)	p=0.442
Left	43 (33.9%)	34 (32.4%)	9 (40.9%)	
Neoadjuvant treatment	17 (13.4%)	14 (13.3%)	3 (13.6%)	p=0.970
Extended surgery	12 (9.4%)	11 (10.4%)	1 (4.5%)	p=0.690
Double sleeve	9 (7.1%)	8 (7.6%)	1 (4.5%)	p=0.516
Carinal sleeve	3 (2.4%)	3 (2.9%)	--	p=1
Upper lobectomy	75 (59.1%)	60 (57.1%)	15 (68.2%)	
Upper bilobectomy	5 (3.9%)	5 (4.8%)	--	
Middle lobectomy	3 (2.4%)	3 (2.9%)	--	p=0.133
Lower lobectomy	34 (26.8%)	30 (28.6%)	4 (18.2%)	
Lower bilobectomy	10 (7.9%)	7 (6.7%)	3 (13.6%)	
Number of lymph node stations sampled	6.8±1 (R: 6-10)	6.8±0.9 (R: 6-10)	7±1.1 (R: 6-10)	p=0.349
Squamous cell carcinoma	85 (66.9%)	75 (71.4%)	10 (45.5%)	
Carcinoid tumor	21 (16.5%)	14 (13.3%)	7 (31.8%)	p=0.066
Adenocarcinoma	12 (9.4%)	10 (9.5%)	2 (9.1%)	
Other*	9 (7.1%)	6 (5.7%)	3 (13.6%)	
Pathology T size (cm)	3.6±1.9 (R: 0-15)	3.7±2 (R: 0-15)	3±1.5 (R: 0-6.5)	p=0.106
Pathology T staging				
T0	2 (1.6%)	1 (1%)	1 (4.5%)	
T1	43 (33.9%)	35 (33.3%)	8 (36.4%)	p=0.793
T2	52 (40.9%)	44 (41.9%)	8 (36.4%)	
T3	24 (18.9%)	20 (19%)	4 (18.2%)	
T4	6 (4.7%)	5 (4.8)	1 (4.5%)	
Pathology lymph node staging				
N0	77 (60.6%)	64 (61%)	13 (59.1%)	p=0.363
N1	37 (29.1%)	32 (30.5%)	5 (22.7%)	
N2	13 (10.2%)	9 (8.6%)	4 (18.2%)	
Pathology staging				
Stage 0	2 (1.6%)	1 (1%)	1 (4.5 %)	
Stage 1A	33 (26%)	28 (26.7%)	5 (22.7 %)	
Stage 1B	24 (18.9%)	20 (19%)	4 (18.2 %)	p=0.709
Stage 2B	46 (36.2%)	38 (36.2%)	8 (36.4 %)	
Stage 3A	16 (12.6%)	14 (13.3%)	2 (9.1 %)	
Stage 3B	6 (4.7%)	4 (3.8%)	2 (9.1 %)	
R0 resection	125 (98.4%)	104 (99%)	21 (%95.5)	p=0.318
R1 resection	2 (1.6%)	1 (1%)	1 (%4.5)	

*Adenosquamous carcinoma, mixed neuroendocrine tumor, pleomorphic carcinoma, inflammatory myofibroblastic tumor and unspecified type

VATS: Video-assisted thoracoscopic surgery, T: Tumor, n: Number, R: Range

Table 3. Periodical development in VATS sleeve resections

	First period (year 2020-2022) n=6	Second period (after January 2023) n=16
Conversion rate (VATS /initiated with VATS)	6/12 (50%)	4/20 (20%)
After neoadjuvant therapy n (%)	--	3 (100%)
Complication rate	2 (33.3%)	7 (43.8%)
BPF n (%)	1 (16.7%)	2 (12.5%)

BPF: Bronchopleural fistula, VATS: Video-assisted thoracoscopic surgery, n: Number

Table 4. Early postoperative morbidity and 30-day mortality

Variables	All patients (127) n (%)	Thoracotomy group (105) n (%)	VATS group (22) n (%)	p-value
Complications	41 (32.3%)	32 (30.5%)	9 (40.9%)	p=0.475
BPF	9 (7%)	6 (5.7%)	3 (13.6%)	
Early (1-7 days)	2	2	--	
Medium (7-30 days)	6	3	3	p=0.188
Late (>30 days)	1	1	--	
BPF Treatment				
Complementary pneumonectomy	6 (66.6%)	5 (83.3%)	1 (33.3%)	
- Omentum support available	2	2	--	
- Omentum support not available	4	3	1	
- Addition of middle lobectomy	1	--	1	
- Primary repair	2	1	1	
Revised cases	16 (12.6%)	12 (11.4%)	4 (18.2%)	
- BPF	9	6	3	
- Hemorrhage/hematoma	1	1	--	p=0.477
- Middle lobe syndrome	1	1	--	
- Wound site infection	5	4	1	
Grade II	15 (36.6%)	13 (40.6%)	2 (22.2%)	
AF/cardiac problems	5	4	1	
Metabolic problems	2	1	1	
Pneumonia	2	2	--	
PAL	6	6	--	
Grade IIIA*	6 (14.6%)	5 (15.6%)	1 (11.1%)	
Empyema	1	1	--	
Secretion retention	3	2	1	
PAL (requiring revision)	2	2	--	
Grade IIIB**	16 (39%)	12 (37.5%)	4 (44.4%)	
BPF	9	6	3	
Hemorrhage/hematoma	1	1	--	
Middle lobe syndrome	1	1	--	
Wound infection	5	4	1	
Grade IV	2 (4.8%)	1 (3.1%)	1 (11.1%)	
Pulmonary embolism	1	--	1	
Pneumonia (tracheotomy opened)	1	1	--	
Grade V	2 (1.57%)	2 (1.9%)	--	p=0.682
Mortality (30-days)	2	2	--	

*: Grade, IIIA: Intervention under local anesthesia, **: Grade,VATS: Video-assisted thoracoscopic surgery, BPF: Bronchopleural fistula, AF: Atrial fibrillation, PAL: Prolonged air leak, IIIB: Intervention under general anesthesia, n: Number

benefits of the VATS approach in more extended and complex cases such as sleeve resections is a controversial issue.

As a result of the developments in the field of minimally invasive surgery, in 2002, Santambrogio et al.[9] performed sleeve resection with VATS for the first time in a 15-year-old patient diagnosed with mucoepidermoid carcinoma and in the following years, sleeve resection with VATS was performed in experienced centers with large series. VATS sleeve resection operations, which have started in our clinic in recent years, have increased over the years and today most cases requiring sleeve anastomosis can be performed by VATS. In a study by Huang et al.[10] in 2016, 118 VATS bronchial sleeve cases were approached with 3 ports. Similarly, Acar and Ceylan[11] defined the technique as 3 ports in their VATS sleeve resection series. On the other hand, Gonzalez-Rivas et al.[12] published their first case of uniportal VATS sleeve lobectomy in 2013 and published a large series in the following years[13]. In our clinic, the three-port approach was not used in any case. Although biportal approach was the preferred method, 22.7% of the cases were performed by uniportal approach.

Zhang et al.[14] underlined that at least 100 standard lobectomies by VATS and at least 10 sleeve lobectomies by thoracotomy are required to gain sufficient experience and expertise in VATS sleeve resections. Imai and Weksler[5] also emphasized that applying VATS in such complex procedures requires meticulous planning and considerable technical expertise. In a series of 201 cases of sleeve resection between 2010-15 using the National Cancer Data Base (NCDB), the rate of patients who underwent VATS was reported to be 21%[15]. Although this rate was 17.4% in our study, this rate has reached 55% with increasing experience in recent years. However, despite all this experience, surgeons should always keep in mind the decision criteria and timing for conversion to thoracotomy[16].

When the literature is analyzed, conversion to thoracotomy rates, European Society of Thoracic Surgeons (ESTS) database and the US NCDB, these values were 24.5% and 20.5%, respectively[17-19]. Despite these high rates[19], conversion to thoracotomy rates have been reported to be quite low in Chinese studies, ranging from 2.9% to 4.5%[19]. Although our conversion rate was as high as 31.2% in our study, this rate decreased to 20% in the second period with the experience we gained, according to the ESTS database. In the initial period, a higher rate of conversion to thoracotomy was observed due to the preference for initiating surgery with VATS in cases where preoperative thoracic CT imaging did not allow for a definitive decision. With the accumulation of experience, the conversion rate has subsequently decreased. As a result of increasing experience, more extended sleeve resections can be performed in experienced centers. In our clinic, broncho-vascular sleeve resection in one case, sleeve bilobectomy in 3 cases and sleeve resection after neoadjuvant treatment in 3 cases were performed in the advanced period.

Table 5. After propensity score matching

	Thoracotomy group n=22 (%)	VATS group n=22 (%)	p-value
Age ± SD	59.3±11.7	60.7±11.9	p=0.707
Male Female	20 (90.9) 2 (9.1)	17 (77.3) 5 (22.7)	p=0.412
Smoking (pack/year) ± SD	32±15.7	30±17.4	p=0.774
Neoadjuvant treatment	4 (18.2)	3 (13.6)	p=1
Right zone Left zone	14 (63.6) 8 (36.4)	13 (59.1) 9 (40.9)	p=0.757
Extended surgery (double sleeve)	1 (4.5)	1 (4.5)	p=1
Operation time (hours) ± SD	4.5±1 (R: 3-6)	4.5±0.9 (R:3-6)	p=0.962
Amount of perioperative bleeding (mL) ± SD	459±356 (R: 350-1700)	370±70 (R: 350-650)	p=0.925
Drainage time (days) ± SD	5.3±5.8 (R: 2-29)	3.6±3.3 (R:2-18)	p=0.023
Duration of hospitalization (days) ± SD	7.7±7.7 (R: 3-34)	5.1±3.4 (R:3-19)	p=0.043
Complication Minor Major	8 (36.4) 1 (4.5) 7 (31.8)	9 (40.9) 2 (9.1) 7 (31.8)	p=0.951
Bronchopleural fistula	2 (9.1)	3 (13.6)	p=1
Mortality (30-days)	--	--	p=1

VATS: Video-assisted thoracoscopic surgery, SD: Standard deviation, n: Number, R: Range

Although the VATS sleeve resection technique provides the advantages of a minimally invasive approach, it is an important issue whether the method provides oncological principles compared to open surgery. In their meta-analysis, Deng et al.[19] have reported that VATS has similar oncological outcomes to open surgery. Yang et al.[20] have also demonstrated comparable short- and long-term outcomes in their propensity score matched analysis of VATS versus open thoracotomy sleeve lobectomy cases. In the present study, we did not observe any difference between the groups in terms of oncological results from the literature.

In the NCDB, the duration of hospital stay after open surgery and VATS sleeve resection was reported to be 6 days in both groups[16]. In their evaluation using the ESTS database, Gonzalez et al.[17] evaluated 1652 sleeve lobectomy patients performed by 270 thoracic surgery units from 25 different European countries between 2007 and 2021 and showed a significant difference in hospital stay between open surgery and VATS patient groups (5 vs 8 days). In the same study, VATS sleeve resection was associated with significantly decreased overall morbidity (30.4% vs 41.7%, p=0.006). In our case series, there was no statistically significant difference between the groups in terms of postoperative complications. Similar to the Gonzalez review, we observed that patients were discharged faster in the VATS group and this difference was statistically significant (7.1 vs 5.1 days).

Xie et al.[18] reported that thoracic drains were terminated more rapidly in patients undergoing sleeve resection by VATS. (6 vs 5 days). In our study, we found statistically significantly shorter drainage times in the VATS group (4.5 vs 3.6).

According to the meta-analysis results of 5 studies with a total of 436 patients comparing the preoperative findings of both groups, less blood loss and longer operation time were observed in the VATS sleeve group[21]. Geropoulos reported this time as 45 minutes longer in the VATS group in his case series[22]. In our study, we observed less blood loss in the VATS sleeve group, although not statistically significant. Also, we did not observe a significant difference between the groups in terms of operative times.

The most important problems that may develop in sleeve resections are those related to anastomosis lines. Tapias et al.[23] reported this rate as 4.3% in VATS sleeve resection cases. The rate of BPF in our patient cohort is higher than in the literature. However, with increasing experience, this rate has reached levels compatible with the literature in the advanced period (6.2%). In the literature, many studies, mainly from China, have been published in recent years on the results of VATS sleeve resections and/or comparisons with open surgery (Table 6)[15,17,18,20,21,23,24,25].

Study Limitations

There are some limitations to the study. First, our study is retrospective. Although we are a high-volume hospital and have experience in sleeve resections in open surgery, our VATS experience in these cases has increased in recent years. Therefore, our number of patients is considerably lower than similar studies in the literature and our outcomes have reached levels compatible with the literature only in recent years. Another limitation of our study is that we cannot state whether the method provides a survival advantage over open surgery because the patient results have not yet reached sufficient time for survival analyses. One of the biggest advantages of minimally invasive surgery compared to open surgery is its effects on early pain and quality of life. Unfortunately, no evaluation of these parameters was made in our study. However, our study is the first study in our country in which VATS sleeve resections were compared to open surgery. The importance of surgical education and transmission of advanced minimally invasive techniques to the next generation has been underlined in recent literature as well[16].

In comparative studies, it is generally expected that the number of cases in both groups be similar. However, VATS sleeve lobectomy is performed in a limited number of centers, and the case numbers remain lower compared to other surgical approaches. This is a key limitation; however, due to the limited literature, the study may still offer valuable contributions. The initial analysis aimed to compare the general features of the two techniques, and thus, downstaging data after neoadjuvant therapy were not included.

CONCLUSION

VATS sleeve resections in NSCLC cases provide advantages over open surgery with shorter hospital stays and drainage times without compromising oncological principles. VATS sleeve resections are surgical procedures that require experience and can be preferred as an alternative to open surgery by experienced centers and physicians.

Table 6. Studies on VATS sleeve resections

Author/country/year	VATS (n) VATS-Thr (%)	Conversion (%)	Complication rate VATS (%)-Thr (%) p-value	Hospitalization VATS-Thr (days) p-value	The mean operation time (minimum) p-value	Blood loss (mL) VATS-Thr p-value	The mean postoperative drainage duration VATS-Thr (day) p-value
Zhong Y, China 2020 (Meta-analysis of 6 studies)[21]	281 43.2%	2.9%-4.5%		SMD -0.24, 95% CI: -0.51 to 0.03, p=0.078	SMD 0.59, 95% CI: 0.14 to 1.03, p=0.010		
Mayne NR, USA, 2021[15]	44.21%	20%		6-6 p=0.36			
Zhu XY, China 2021[25]				10.5	247.8±73.1	300.4±321.8 mL	
Geropoulos G, Greece-London 2022 6 studies[22]	229.35%				45.85 minimum less in VATS, p=0.01	37 mL less in VATS p<0.001	
Xie D, China 2021[18]	112 30.8%			6-7		100-200 p<0.1	5-6 p<0.1
Yang Y, China 2020[20]	44 23.5%		20.3%-30.2% p=0.028				
Ceylan KC, Türkiye 2020[24]				6.4±1.9	288.2±77.1		5.5±1.9
Gonzalez M, Thoracic Surgery database 2021[17]	161 9.8%	21.1%	30.4%-41.7% p=0.006	5-8 p<0.001			
Demirkol EY. (this study)	22 17.3%	33%	40.9%-30.5% p=0.475	5.1±3.4 7.1±7 p=0.005	4.5 0.9 p=0.474	370±70-467±385 p=0.525	3.6±3.3-4.5±4 p=0.014

VATS: Video-assisted thoracoscopic surgery, Thr: Thoracotomy, SMD: Standardized mean difference, CI: Confidence interval

Ethics

Ethics Committee Approval: The study was approved by the ethics/scientific committee of University of Health Sciences Türkiye, Yedikule Chest Diseases and Thoracic Surgery Training and Research Hospital and (decision no: 2023-462.28, date:12.2023) was conducted by the principles of the Declaration of Helsinki.

Informed Consent: It is a retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: V.E., M.M., Concept: M.Ü., Design: V.E., M.Ü., Data Collection or Processing: N.Y., E.K., Analysis or Interpretation: M.E.F., C.B.S., Literature Search: V.E., M.Ü., Writing: E.Y.D.

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

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

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

REFERENCES

1. National Comprehensive Cancer Network (NCCN). Clinical practice guidelines in oncology: non-small cell lung cancer. Version 2. 2023.
2. Brewer LA 3rd. The first pneumonectomy. Historical notes. J Thorac Cardiovasc Surg. 1984;88:810-26.
3. THOMAS CP. Conservative resection of the bronchial tree. J R Coll Surg Edinb. 1956;1:169-86.
4. Yamamoto K, Miyamoto Y, Ohsumi A, Kojima F, Imanishi N, Matsuoka K, et al. Sleeve lung resection for lung cancer: analysis according to the type of procedure. J Thorac Cardiovasc Surg. 2008;136:1349-56.
5. Imai T, Weksler B. Commentary: open or VATS? Roll up your sleeve. J Thorac Cardiovasc Surg. 2021;161:414-5.
6. Chen J, Soultanis KM, Sun F, Gonzalez-Rivas D, Duan L, Wu L, et al. Outcomes of sleeve lobectomy versus pneumonectomy: a propensity score-matched study. J Thorac Cardiovasc Surg. 2021;162:1619-28.e4.

7. Stallard J, Loberg A, Dunning J, Dark J. Is a sleeve lobectomy significantly better than a pneumonectomy? Interact Cardiovasc Thorac Surg. 2010;11:660-6.

8. Erdogu V, Akin H, Sonmezoglu Y, Kutluk AC, Sezen CB, Dogru MV, et al. Comparison of the video-assisted thoracoscopic lobectomy versus open thoracotomy for primary non-small cell lung cancer: single cohort study with 269 Cases. Sisli Etfal Hastan Tip Bul. 2020;54:291-6.

9. Santambrogio L, Cioffi U, De Simone M, Rosso L, Ferrero S, Giunta A. Video-assisted sleeve lobectomy for mucoepidermoid carcinoma of the left lower lobar bronchus: a case report. Chest. 2002;121:635-6.

10. Huang J, Li S, Hao Z, Chen H, He J, Xu X, et al. Complete video-assisted thoracoscopic surgery (VATS) bronchial sleeve lobectomy. J Thorac Dis. 2016;8:553-74.

11. Acar A, Ceylan KC. VATS sleeve resections. Turk Gogus Kalp Damar Cerrahisi Derg. 2023;31:S45-53.

12. Gonzalez-Rivas D, Fernandez R, Fieira E, Rellan L. Uniportal video-assisted thoracoscopic bronchial sleeve lobectomy: first report. J Thorac Cardiovasc Surg. 2013;145:1676-7.

13. Soultanis KM, Gonzalez-Rivas D. Uniportal video-assisted sleeve resections: how to deal with specific challenges. J Thorac Dis. 2019;11:S1670-7.

14. Zhang C, Ma Y, Yu Z, Zhu B, Liu H. Comparison of efficacy and safety of hybrid video-assisted thoracoscopic surgery *vs.* thoracotomy sleeve lobectomy for non-small cell lung cancer: a propensity score matching study. J Thorac Dis. 2022;14:2635-44.

15. Mayne NR, Darling AJ, Raman V, Balderson S, Berry MF, Harpole DH Jr, et al. Perioperative outcomes and 5-year survival after open versus thoracoscopic sleeve resection for lung cancer. Semin Thorac Cardiovasc Surg. 2021;33:522-30.

16. Nakagawa K. Minimally invasive sleeve lobectomy: it is important to pass excellent techniques to the next generation. J Thorac Dis. 2023;15:946-9.

17. Gonzalez M, Chriqui LE, Décaluwé H, Aigner C, Rényi-Vámos F, Opitz I, et al. Sleeve lobectomy in patients with non-small-cell lung cancer: a report from the European Society of Thoracic Surgery database 2021. Eur J Cardiothorac Surg. 2022;62:ezac502.

18. Xie D, Deng J, Gonzalez-Rivas D, Zhu Y, Jiang L, Jiang G, et al. Comparison of video-assisted thoracoscopic surgery with thoracotomy in bronchial sleeve lobectomy for centrally located non-small cell lung cancer. J Thorac Cardiovasc Surg. 2021;161:403-13.

19. Deng HY, Qiu XM, Zhu DX, Tang X, Zhou Q. Video-assisted thoracoscopic sleeve lobectomy for centrally located non-small cell lung cancer: a meta-analysis. World J Surg. 2021;45:897-906.

20. Yang Y, Mei J, Lin F, Pu Q, Ma L, Liu C, et al. Comparison of the short- and long-term outcomes of video-assisted thoracoscopic surgery versus open thoracotomy bronchial sleeve lobectomy for central lung cancer: a retrospective propensity score matched cohort study. Ann Surg Oncol. 2020;27:4384-93.

21. Zhong Y, Wang Y, Hu X, Wang G, She Y, Deng J, et al. A systematic review and meta-analysis of thoracoscopic versus thoracotomy sleeve lobectomy. J Thorac Dis. 2020;12:5678-90.

22. Geropoulos G, Esagian SM, Skarentzos K, Ziogas IA, Katsaros I, Kosmidis D, et al. Video-assisted thoracoscopic versus open sleeve lobectomy for non-small cell lung cancer: a systematic review and meta-analysis from six comparative studies. Asian Cardiovasc Thorac Ann. 2022;30:881-93.

23. Tapias LF, Ott HC, Mathisen DJ. Complications following carinal resections and sleeve resections. Thorac Surg Clin. 2015;25:435-47.

24. Ceylan KC, Acar A, Örs Kaya Ş. The initial experience on thoracoscopic sleeve lobectomy: continuous suturing technique of the bronchial anastomosis in 12 cases. Surg Laparosc Endosc Percutan Tech. 2020;30:476-9.

25. Zhu XY, Ding C, Xu C, Chen J, Ju S, Pan S, et al. Techniques and outcomes of bronchoplastic and sleeve resection: an 8-year single-center experience. Transl Lung Cancer Res. 2021;10:4538-48.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.93041
Nam Kem Med J 2025;13(3):287-292

The Effect of 3D Models Created with Augmented Reality on Diagnosis and Orthopedic Resident Education of Tibial Plateau Fractures

Artırılmış Gerçeklik ile Oluşturulan 3D Modellerin Tibia Plato Kırıklarının Tanı ve Ortopedi Asistan Eğitimi Üzerindeki Etkisi

Alper DÜNKİ[1], Ömer POLAT[1], Melih GÜNEY[2]

[1]University of Health Sciences Türkiye, Ümraniye Training and Research Hospital, Clinic of Orthopaedics and Traumatology, İstanbul, Türkiye
[2]Dr. İsmail Fehmi Cumalıoğlu City Hospital, Clinic of Orthopaedics and Traumatology, Tekirdağ, Türkiye

ABSTRACT

Aim: Computed tomography (CT) is the gold standard imaging for the diagnosis of intra-articular fractures such as tibial plateau fractures that may show different fracture patterns. The aim of the study was to compare the diagnostic accuracy (DA) and response time (RT) between new technologies such as augmented reality (AR), CT and 3-dimensional CT (3D-CT): in the classification of tibial plateau fractures in orthopedic assistant education and daily practice.

Materials and Methods: The orthopedic residents receiving training in our clinic were divided into 2 groups according to their training period: 2.5 years and below and above 2.5 years. Nine separate tibial plateau fractures were selected according to the Schatzker and Luo classification. DA, RT and method confidence of each resident were measured with a double-blind questionnaire.

Results: DA averages of the participants were examined, and it was seen that DA for CT was 67.8%, for 3D-CT, it was 52.9% and for AR, it was 64%. When the correct RT were examined, the average RT-AR was 49.9 (±11.8) sec, RT-3D 58 (±16.7) sec and RT-CT 80 (±23.8) sec. When the RT values of the AR models were examined, the average RT-AR was 41.6 (±8.49) sec in the SEN group and 58.2 (±8.42) sec in the JUN RT-AR.

Conclusion: In this study, it was shown that the DA of AR in the diagnosis and treatment planning of tibial plateau fractures was similar to conventional CT and superior to 3D-CT. The DA rate of novice assistants was lower and the correct diagnosis time was longer, and similar results were obtained in all 3 groups.

Keywords: Tibia fracture, augmented reality, orthopaedic residents, computed tomography

ÖZ

Amaç: Bilgisayarlı tomografi (BT), tibial plato kırıkları gibi intraartiküler kırıkların tanısında altın standart görüntüleme yöntemidir ve farklı kırık paternleri gösterebilir. Çalışmanın amacı, ortopedi asistan eğitimi ve günlük uygulamada tibial plato kırıklarının sınıflandırılmasında artırılmış gerçeklik (AR), BT ve 3 boyutlu BT (3D-BT) gibi yeni teknolojilerin tanısal doğruluk (DA) ve yanıt süresi (RT) açısından karşılaştırılmasıdır.

Gereç ve Yöntem: Kliniğimizde eğitim gören ortopedi asistanları, eğitim sürelerine göre 2,5 yıl ve altı ile 2,5 yıl ve üstü olmak üzere 2 gruba ayrılmıştır. Schatzker ve Luo sınıflandırmasına göre 9 ayrı tibial plato kırığı seçilmiştir. Her asistanın DA, RT ve yöntem güvenilirliği çift kör bir anket ile ölçülmüştür.

Bulgular: Katılımcıların DA ortalamaları incelendi ve BT için %67,8, 3D-BT için %52,9 ve AR için %64 olduğu görüldü. Doğru yanıt süreleri incelendiğinde; ortalama RT-AR 49,9 (±11,8) saniye, RT-3D 58 (±16,7) saniye ve RT-BT 80 (±23,8) saniye olarak bulunmuştur. AR modellerinin RT değerleri incelendiğinde; ortalama RT-AR, SEN grubunda 41,6 (±8,49) saniye ve JUN RT-AR'da 58,2 (±8,42) saniye olmuştur.

Address for Correspondence: Alper DÜNKİ MD, University of Health Sciences Türkiye, Ümraniye Training and Research Hospital, Clinic of Orthopaedics and Traumatology, İstanbul, Türkiye
E-mail: alperdunki93@yahoo.com **ORCID ID:** orcid.org/0000-0001-6577-3491
Received: 07.04.2025 **Accepted:** 12.06.2025 **Publication Date:** 07.10.2025

Cite this article as: Dünki A, Polat Ö, Güney M. The effect of 3D models created with augmented reality on diagnosis and orthopedic resident education of tibial plateau fractures. Nam Kem Med J. 2025;13(3):287-292

Sonuç: Bu çalışmada, tibial plato kırıklarının tanı ve tedavi planlamasında AR'nin DA konvansiyonel BT ile benzer ve 3D-BT'den üstün olduğu gösterilmiştir. Acemi asistanların DA oranı daha düşük ve doğru tanı süresi daha uzundu ve 3 grubun tümünde benzer bulgular elde edilmiştir.

Anahtar Kelimeler: Tibia kırığı, artırılmış gerçeklik, ortopedi asistanları, bilgisayarlı tomografi

INTRODUCTION

Tibial plateau fractures constitute 1-2% of the fractures seen in the extremities. These fractures occur in patients aged under 65 with high-energy trauma and in those over 65 with low-energy trauma[1]. Schatzker and Luo classifications are widely used today for the diagnosis and treatment of these fractures, which can be treated with surgical or conservative methods[2,3].

Anteroposterior (AP) and lateral radiographs alone are not sufficient to confirm the diagnosis and to plan surgery in tibial plateau fractures that are intra-articular and can show different fracture types. Computed tomography (CT) is the gold standard imaging for the diagnosis of intra-articular fractures that can show different fracture patterns, such as tibial plateau fractures[4]. Today, 3-dimensional CT (3D-CT) images that combine 2-D images in axial, sagittal and coronal planes are frequently used. With this imaging, which provides visual integrity, diagnosis and treatment planning in multi-part and complicated fractures can be made much more accurately and quickly.

In the last few years, virtual reality applications, which started with virtual reality glasses, have begun to be used in surgical planning and diagnosis with 3D imaging[5,6]. Augmented reality (AR) allows viewing images in a real-world environment[7,8]. Thanks to these developments, more understandable and easily accessible AR images can be utilized in student education. Thus, students can receive education with 3D images instead of the monotonous 2D images they only see in books, and they can learn faster and with more fun by keeping their excitement fresh[9].

Trauma is unpredictable by nature, and diagnosis and treatment may vary depending on the trauma energy and type of trauma for many different fracture types. When resident physicians see a fracture, ways are sought to facilitate planning its diagnosis and treatment and the ways to obtain the necessary information.

The aim of our study was to compare the diagnostic accuracy (DA) (and correct diagnosis speed between new technologies, AR, and the gold standard conventional CT and 3D-CT in the classification of tibial plateau fractures in orthopedic assistant training and daily practice. In addition, the relationship between the duration of orthopedics and traumatology training and the level of trust in these new technologies was also investigated.

MATERIALS AND METHODS

Ethics committee approval was received from the University of Health Sciences Türkiye, Ümraniye Training and Research Hospital with (decision no: E-54132726-000-271399262, date: 13.03.2025). The study was conducted in accordance with the principles of the Declaration of Helsinki.

In our hospital, 20 resident physicians who continue their postgraduate education in orthopedics and traumatology were divided into 2 groups as 2.5 years and below junior (JUN) and 2.5 years and above senior (SEN).

Patients between the ages of 18-65 who applied to our hospital in 2024 and were diagnosed with tibial plateau fractures were included in the study. Patients with pathological fractures, multiple traumas, pseudoarthrosis cases and patients without 3D-CT imaging were excluded from the study. Informed consent was obtained from the participants included in the study.

The tibial plateau fractures used in the study were selected and classified by 2 SEN surgeons with over 5 years of experience in the field of Orthopedics and Traumatology. Based on the Schatzker and Luo classifications, 9 different fracture models were numbered in 3 groups as conventional CT, 3D-CT and AR images. The final evaluation of the fracture type was completed in open surgery.

In conventional CT images, axial, coronal and sagittal images in the bone window were selected. DICOM format images were uploaded to the AR method "Object viewer'' application via the website. The images were uploaded to the cube as "format. stl'' files via the ''object viewer'' application downloaded from the application website (MergeEDU©, MergeLabsInc., SanAntonio, TX, USA). The cube was shown to the camera of a tablet. Thus, a 3D image that can be rotated in every axis was achieved on the tablet screen (Figure 1).

In addition to the two SEN surgeons who selected the fracture types, a third surgeon who was blind to the fracture types was shown the fracture patterns in a mixed order and only numbered to 20 orthopedic residents. The third surgeon had no previous involvement in the fracture selections and was the only one present during the survey, reducing the possibility of bias. They were asked to mark their answers using the prepared survey form and multiple-choice answers. Each candidate filled out the survey individually, and one of the study directors was with them during this time. The time it took them to reach the answer was recorded in seconds by the directors using a stopwatch.

They were asked to mark the method that came first when they ranked the three methods from the most reliable to the least reliable in terms of helping with diagnosis. Seven questions were asked as a survey about whether this method alone was sufficient for diagnosis, whether they would continue to use it in daily practice, and their views on the contribution of the AR method to resident education. The options given to the question of which method they felt most reliable were "Conventional CT", "AR", and "3D-CT". The answers to the other 6 questions were; "Absolutely yes", "Yes", "Undecided", and "No" (Figure 2).

The fractures were selected according to the Schatzker and Luo classifications, which are the two most commonly used classifications in the diagnosis and treatment of tibial plateau fractures. Nine different fracture patterns were selected, including Schatzker type "1,2,3,4,5 and 6" and Luo "medial, lateral and posterior column" types. The resident physicians who participated in the test were shown how to use the AR application. None of the participants had ever encountered AR technology before. Data: DA was recorded as "percentage", response time (RT) as "seconds" and survey responses were evaluated by groups and overall.

Statistical Analysis

Statistical analysis of the study was performed using SPSS program, version 29.0 (SPSS, Inc, an IBM Company, Chicago, IL). The conformity of the values to normal distribution was assessed using the Shapiro-Wilk test. Variables without normal distribution were analyzed using the Mann-Whitney U test. Independent Student's t-tests were used to compare variables with normal distribution. Categorical data were statistically analyzed using the chi-square test or Fisher's exact test. The statistical significance level (p) was determined as 0.05.

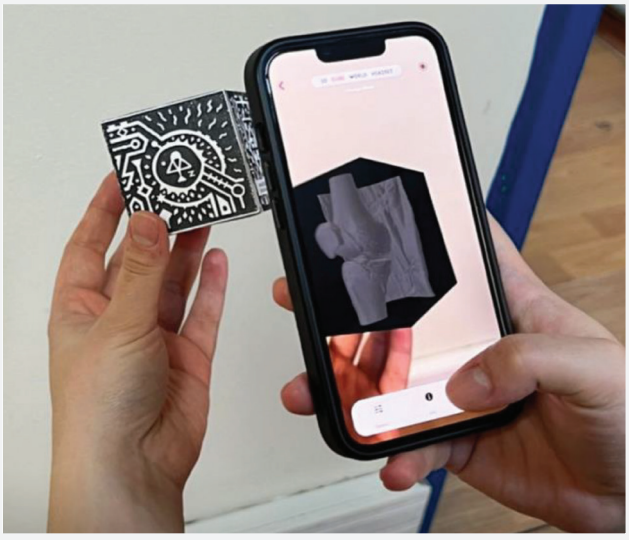


Figure 1. Screenshot of the AR model in the "Object Viewer" app

AR: Augmented reality

RESULTS

In our institution, 20 resident physicians who continue their postgraduate Orthopedics and Traumatology education were divided into two groups according to their education period as 2.5 years and below JUN (n: 10) and above 2.5 years SEN (n: 10). Conventional CT, 3D-CT and AR images of 9 different fractures that fit the Schatzker and Luo classification were given to the resident physicians separately in accordance with the double-blind study model and they were asked to answer.

When DA means were examined, it was seen that it was 67.8% for conventional CT, 52.9% for 3D-CT and 64% for AR. When the relationship between DA rates was examined, no statistically significant difference was observed between conventional CT and AR (p: 0.243). P<0.05 was observed between conventional CT and 3D-CT and between AR and 3D-CT. Accordingly, AR and conventional CT were found to be statistically superior to 3D-CT in terms of DA. However, no statistically significant difference was observed between conventional CT and AR.

When DA values were examined according to SEN and JUN groups; while the average DA-AR in the SEN group was 79%, it was found to be 49% in JUN DA-AR (Figure 3). In the SEN group, the average DA-3D was 66%, JUN DA-3D was 40%, and the average DA-CT in the SEN group was 83%, while it was found to be 52% in JUN DA-CT. Between the SEN and JUN groups, the p value for DA-CT, DA-3D and DA-AR was below 0.05, and it was seen that the correct response rate of the SEN group was higher than the JUN group.

When the correct RT were examined; mean RT-AR was 49.9 (±11.8) sec, RT-3D was 58 (±16.7) sec, and RT-CT was 80 (±23.8) sec. When the relationship between the correct RT of the groups was examined, the conventional CT group had the longest RT, and this difference was statistically significant (Table 1).

When the RT values were examined according to the SEN and JUN groups; in the SEN group, mean RT-AR was 41.6 (±8.49) sec, and in the JUN RT-AR, it was 58.2(±8.42) sec (Figure 4). In the SEN group, mean RT-3D was 46.2 (±11.0) sec, in the JUN RT-3D, it was 69.9 (±12.3) sec, and in the SEN group, mean RT-CT was 62.7 (±11.8) sec, and in the JUN RT-CT, it was 97.2 (±19.9) sec. The p-value was less than 0.05 in all 3 groups, and the mean correct RT of the SEN group had significantly faster RT than the JUN group. According to the answers given by the resident physicians participating in the study to the question of which method they find more reliable (Figure 5); according to the diagnostic method confidence ranking prepared according to the Likert scale, the JUN group chose AR first and found it more reliable, while the SEN group found conventional CT more reliable.

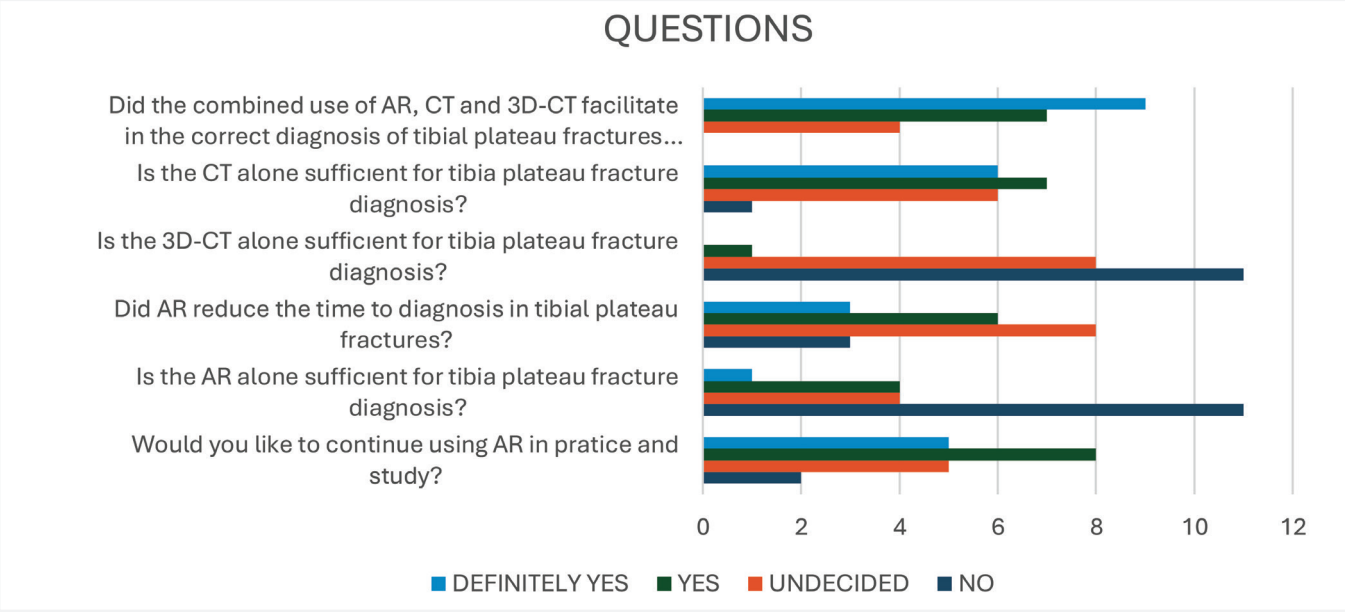


Figure 2. Questions asked to the participants

3D-CT: 3-dimensional computed tomography, AR: Augmented reality, CT: Computed tomography

Table 1. Summarization of DA and RT results for AR, 3D-CT and conventional CT between the JUN and SEN groups

	SEN	JUN
DA-AR	79%	49%
DA-3D-CT	66%	40%
DA-CT	83%	52%
RT-AR (sec)	41.6	58.2
RT-3D-CT (sec)	46.2	69.9
RT-CT (sec)	62.7	83.5

3D-CT: 3-dimensional computed tomography, AR: Augmented reality, DA: Diagnosis accuracy, JUN: Junior residents, Sec: Second, SEN: Senior residents, RT: Reponse time

DISCUSSION

Conventional CT imaging is superior to AP and lateral bidirectional radiography in the diagnosis and treatment of complex intra-articular fractures with 3 different image planes in axial, coronal and sagittal planes[10]. With the development of technology, 3D images have been developed from conventional CT images, and it has been proven that diagnosis and treatment planning in complex fractures can be made much faster and more accurately[11].

Tibial plateau fractures are also fractures where different fracture patterns and surgical planning change according to the fracture pattern[11,12]. Therefore, speed and accuracy in diagnosis are necessary for successful surgical results.

In recent years, the emergence of new technologies, 3D imaging and AR, has become increasingly important in student learning and practice[13]. Plates and cutting guides produced with 3D printers and imaging have been described in various studies, and articles showing technological developments not only in diagnosis but also in treatment are available in the literature[14,15].

Colcuc et al.[16] compared AR and conventional CT in surgical planning of tibial plateau fractures and although the planning time was longer in the AR group, the planned operation time was lower. In our study, we investigated the effect of AR on diagnosis and education. As a result, similar results were obtained with the literature.

Shen et al.[17] divided the patients into 2 groups in 42 complex tibial plateau fracture cases and performed the diagnosis in one group with conventional CT and in the other group with 3D-CT. As a result, the operation time, bleeding amount and fluoroscopy time were found to be significantly lower in the 3D-CT group. In our study, in accordance with the literature, the correct diagnosis time was found to be significantly lower in the 3D-CT group together with the AR group, but in terms of DA, conventional CT was found to be higher.

Montemagno et al.[18] compared AR, 3D printed models and conventional CT in the diagnosis and resident training of acetabular fractures. They divided 20 residents into 2 groups according to their education period and compared their correct diagnosis rate, diagnosis time and confidence in the methods of 5 different acetabular fractures. When the DA

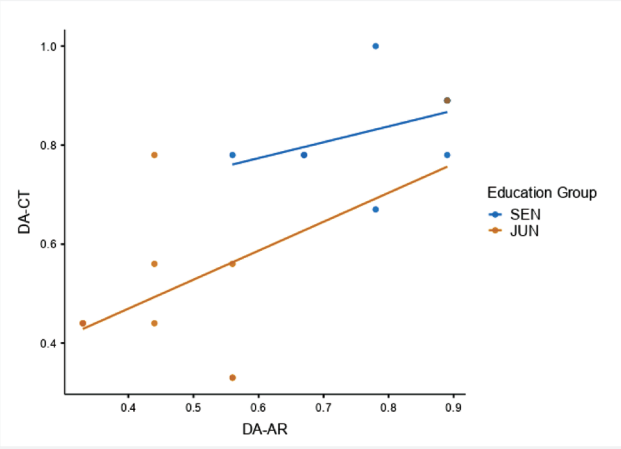


Figure 3. Scatter plot comparing diagnostic accuracy between JUN and SEN groups for conventional CT, 3D-CT, and AR

3D-CT: 3-dimensional computed tomography, JUN: Junior residents, SEN: Senior residents, AR: Augmented reality, DA: Diagnosis accuracy

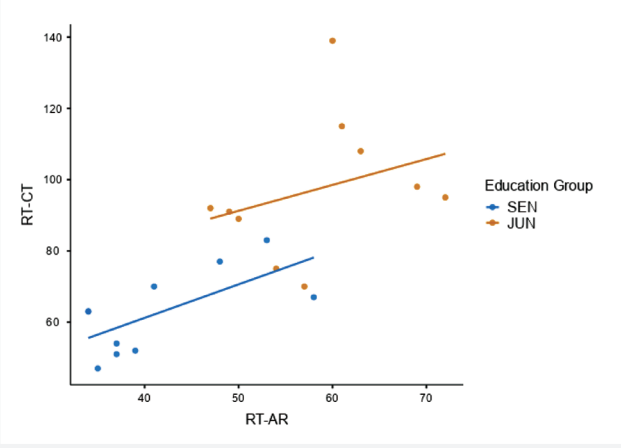


Figure 4. Scatter plot of education groups' response time between JUN and SEN groups for conventional CT, 3D-CT, and AR

3D-CT: 3-dimensional computed tomography, JUN: Junior residents, SEN: Senior residents, AR: Augmented reality, RT: Response time

rates were examined, AR and conventional CT were found to be better than 3D printed models. In our study, in accordance with the literature, the highest DA was observed in AR, which was close to conventional CT, and no statistically significant difference was observed between them. When the seniority period was examined, no difference was observed between the SEN and JUN groups in the AR group, and when the other groups and the SEN group were examined in general, the correct response rate was statistically higher[18]. In our study,

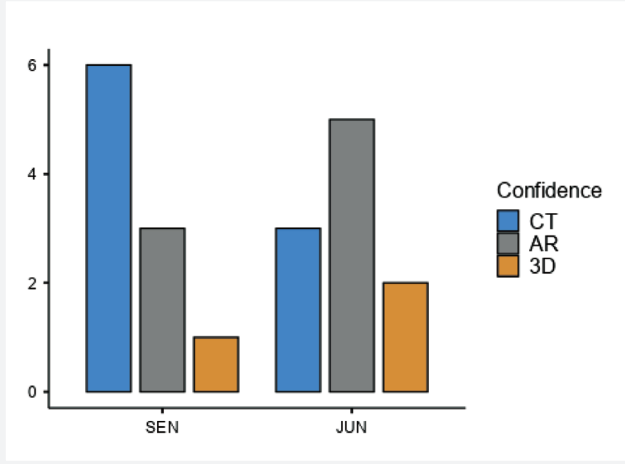


Figure 5. Confidence in diagnostic methods between JUN and SEN groups for conventional CT, 3D-CT, and AR

3D-CT: 3-dimensional computed tomography, JUN: Junior residents, SEN: Senior residents, AR: Augmented reality

all groups and the SEN group in general had a higher correct response rate, and these results were found to be consistent with the literature.

When the correct diagnosis time was examined, Montemagno et al.[18] found the diagnosis time to be shorter in the AR group compared to conventional CT. In their study, the diagnosis time of conventional CT was found to be higher in the SEN group. In our study, when conventional CT was used, the correct diagnosis time was generally found to be longer than in the AR and 3D-CT groups, and the SEN group made a statistically faster diagnosis than the JUN group.

In the Montemagno et al.[18] study, the JUN group found AR safer and the most reliable compared to the SEN group, and both groups placed AR ahead of conventional CT in terms of reliability. In our study, the JUN group found AR more reliable, but the SEN group found conventional CT more reliable than AR. We think that this difference is due to the SEN group's conventional CT usage habits for more than 2.5 years.

Montgomery et al.[19] investigated the role of 3D printed models in the diagnosis and treatment of calcaneus complex fractures with 16 residents and 5 specialist orthopedists and found that the correct diagnosis rate and time of residents were lower than those of specialists. However, this difference was observed to be closer in 3D printed models. In our study, the diagnosis time was faster and correct diagnosis rate of the SEN group was higher than the novice group. However, in our study, the novice group positioned AR ahead of 3D models. We assume that this difference is due to the use of 3D-CT images instead of 3D printed models in our study.

When the survey results of the study are examined, 60% of the participants stated that they could use AR again in their

daily practices. However, 55% of the participants stated that AR alone was not sufficient and 20% were undecided. When asked about the contribution of combining 3D imaging with AR to resident education, 80% of the participants answered "Definitely yes and yes". The JUN group has positioned AR ahead of conventional CT in terms of trust.

Study Limitations

In our study, 20 resident physicians working in our clinic were included in the study. However, the small sample size can be considered a limitation of the study. The lack of previous AR experience among the participants may have affected the results. More research and larger scale studies are needed to introduce AR technology to daily joint traumas.

CONCLUSION

In this study, AR's DA was comparable to conventional CT and superior to 3D-CT. When the correct diagnosis time was examined, AR and 3D-CT were found to be lower than conventional CT. The DA rate of novice residents was lower, and the correct diagnosis time was longer, and similar results were obtained in all 3 groups. However, although the difference was closer in AR, the SEN group was found to be statistically superior. Although 50% of JUN participants supporting AR in education, AR is not capable of replacing conventional CT, which is considered the gold standard in the diagnosis of complex tibial plateau fractures, but the existence of developing technology in orthopedics and traumatology resident education is encouraging in diagnosis and learning.

Ethics

Ethics Committee Approval: Ethics committee approval was received from the University of Health Sciences Türkiye, Ümraniye Training and Research Hospital with (decision no: E-54132726-000-271399262, date: 13.03.2025).

Informed Consent: Informed consent was obtained from the participants included in the study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.D., Ö.P., Concept: A.D., Design: A.D., M.G., Data Collection or Processing: A.D., Analysis or Interpretation: A.D., Ö.P., Literature Search: A.D., Ö.P., Writing: A.D., Ö.P.

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

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

REFERENCES

1. Larsen P, Elsoe R, Hansen SH, Graven-Nielsen T, Laessoe U, Rasmussen S. Incidence and epidemiology of tibial shaft fractures. Injury. 2015;46:746-50.
2. Millar SC, Arnold JB, Thewlis D, Fraysse F, Solomon LB. A systematic literature review of tibial plateau fractures: what classifications are used and how reliable and useful are they? Injury. 2018;49:473-90.
3. Zhu Y, Yang G, Luo CF, Smith WR, Hu CF, Gao H, et al. Computed tomography-based three-column classification in tibial plateau fractures: introduction of its utility and assessment of its reproducibility. J Trauma Acute Care Surg. 2012;73:731-7.
4. Chan PS, Klimkiewicz JJ, Luchetti WT, Esterhai JL, Kneeland JB, Dalinka MK, et al. Impact of CT scan on treatment plan and fracture classification of tibial plateau fractures. J Orthop Trauma. 1997;11:484-9.
5. Salta K, Paschalidou K, Tsetseri M, Koulougliotis D. Shift from a traditional to a distance learning environment during the COVID-19 pandemic: University Students' Engagement and Interactions. Sci Educ (Dordr). 2022;31:93-122.
6. Farid AR, Comtesse S, Sagi HC, Frosch KH, Weaver MJ, Yoon RS, et al. Enabling technology in fracture surgery: state of the art. J Bone Joint Surg Am. 2025;107:1636-47.
7. Sun P, Zhao Y, Men J, Ma ZR, Jiang HZ, Liu CY, et al. Application of virtual and augmented reality technology in hip surgery: systematic review. J Med Internet Res. 2023;25:e37599.
8. Dekhne MS, Stenquist D, Suneja N, Weaver MJ, Petersen MM, Singh UM, et al. Optimizing outcomes after operative treatment bicondylar tibial plateau fractures - time for ınnovation? Arch Bone Jt Surg. 2024;12:80-91.
9. Dhar P, Rocks T, Samarasinghe RM, Stephenson G, Smith C. Augmented reality in medical education: students' experiences and learning outcomes. Med Educ Online. 2021;26:1953953.
10. Puglisi G, Montemagno M, Denaro R, Condorelli G, Caruso VF, Vescio A, et al. 3D-printed models versus CT Scan and X-rays ımaging in the diagnostic evaluation of proximal humerus fractures: a triple-blind ınterobserver reliability comparison study. Adv Orthop. 2022;2022:5863813.
11. Huitema JM, van der Gaast N, Jaarsma RL, Doornberg JN, Edwards MJR, Hermans E. The effect of addition of 2DCT scans and 3DCT scans for the classification of tibial plateau fractures: a systematic review. Eur J Trauma Emerg Surg. 2024;50:71-9.
12. Vosoughi F, Menbari Oskouie I, Rahimdoost N, Pesantez R. Decoding tibial plateau fracture classifications: a century of individualized insights in a systematic review. EFORT Open Rev. 2025;10:316-26.
13. Boboc R, Băutu E, Gîrbacia F, Popovici N, Popovici DM. Augmented reality in cultural heritage: an overview of the last decade of applications. Appl. Sci. 2022;12:9859.
14. Assink N, Reininga IHF, Ten Duis K, Doornberg JN, Hoekstra H, Kraeima J, et al. Does 3D-assisted surgery of tibial plateau fractures improve surgical and patient outcome? A systematic review of 1074 patients. Eur J Trauma Emerg Surg. 2022;48:1737-49.
15. Alemayehu DG, Zhang Z, Tahir E, Gateau D, Zhang DF, Ma X. Preoperative planning using 3D printing technology in orthopedic surgery. Biomed Res Int. 2021;2021:7940242.
16. Colcuc C, Miersbach M, Cienfuegos M, Grüneweller N, Vordemvenne T, Wähnert D. Comparison of virtual reality and computed tomography in the preoperative planning of complex tibial plateau fractures. Arch Orthop Trauma Surg. 2024;144:2631-9.
17. Shen S, Wang P, Li X, Han X, Tan H. Pre-operative simulation using a three-dimensional printing model for surgical treatment of old and complex tibial plateau fractures. Sci Rep. 2020;10:6044.
18. Montemagno M, Testa G, Panvini FMC, Puglisi G, Papotto G, Marchese E, et al. The novel ımpact of augmented reality and 3D printing in the diagnosis of complex acetabular fractures: a comparative randomized study in orthopedic residents. J Clin Med. 2024;13:3059.
19. Montgomery SJ, Kooner SS, Ludwig TE, Schneider PS. Impact of 3D printed calcaneal models on fracture understanding and confidence in orthopedic surgery residents. J Surg Educ. 2020;77:472-8.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.24381
Nam Kem Med J 2025;13(3):293-301

HLA-B27 Frequency and Clinical Correlates in Turkish Patients with Various Subtypes of Spondyloarthritis: A Single-Center Cohort Study

Türk Hastalarda Spondiloartritin Farklı Alt Tiplerinde HLA-B27 Sıklığı ve Klinik Korelasyonları: Tek Merkezli Bir Kohort Çalışması

Dilara BULUT GÖKTEN[1], Hilmi TOZKIR[2], Ömer Atakan SOĞUR[1], Rıdvan MERCAN[1]

[1]Tekirdağ Namık Kemal University Faculty of Medicine, Department of Internal Medicine, Division, of Rheumatology, Tekirdağ, Türkiye
[2]Tekirdağ Namık Kemal University Faculty of Medicine, Department of Medical Genetics, Tekirdağ, Türkiye

ABSTRACT

Aim: This study aimed to evaluate the frequency of human leukocyte antigen (HLA)-B27 positivity and its clinical correlations in Turkish patients with ankylosing spondylitis (AS), psoriatic arthritis (PsA), inflammatory bowel disease (IBD)-associated arthritis, peripheral spondyloarthritis (pSpA), and non-radiographic axial SpA (nr-axSpA).

Materials and Methods: In this retrospective single-center cohort study, data from 524 AS, 380 PsA, 28 IBD-associated SpA, 45 pSpA, and 236 nr-axSpA patients were analyzed. Clinical features, HLA-B27 status, inflammatory markers, and imaging results were collected. Statistical analyses included the Chi-square tests, the Mann-Whitney U tests, Spearman's rank correlations, and logistic regression.

Results: HLA-B27 positivity was detected in 58.2% AS, 31.1% PsA, 25.0% IBD-SpA, 40.0% pSpA, and 35.2% nr-axSpA patients. In AS, HLA-B27 positivity correlated with higher C-reactive protein levels (p=0.001), elevated bath AS disease activity index (BASDAI) (p=0.0234), and bath ankylosing spondylitis functional index (BASFI) (p=0.0207) scores, male sex (p=0.0017), and younger age at onset (p=0.034). In PsA, higher PsA disease activity score were significantly associated with HLA-B27 positivity (p<0.0001). For IBD-SpA, HLA-B27 positivity correlated with younger age at disease onset (p=0.016), increased BASDAI (p=0.0003), and BASFI (p=0.0004). In nr-axSpA, male sex (p=0.00015), elevated BASDAI (p<0.0001), BASFI (p<0.0001), and increased biologic usage (p=0.0277) were significantly associated with HLA-B27 positivity, while non-steroidal anti-inflammatory drug responsiveness was higher in negatives (p=0.0058). Magnetic resonance imaging sacroiliitis negatively correlated with HLA-B27 positivity in AS (rho=-0.140, p=0.002).

Conclusion: HLA-B27 positivity varies across SpA subtypes, significantly correlating with male sex, disease activity, and functional impairment scores in axial SpA groups, with the highest prevalence in AS. While certain associations with disease activity and treatment patterns were observed, the overall clinical impact of HLA-B27 was limited. These findings highlight the complex and heterogeneous nature of HLA-B27's role in SpA and emphasize the need for further prospective studies to clarify its prognostic significance.

Keywords: HLA-B27, spondyloarthritis, seronegative spondyloarthritis

ÖZ

Amaç: Ankilozan spondilit (AS), psöriatik artrit (PsA), enflamatuvar bağırsak hastalığı (İBH)-ilişkili artrit, periferik spondiloartrit (pSpA) ve non-radiographic axial SpA (nr-axSpA) tanılı Türk hastalarda insan lökosit antijeni (HLA)-B27 pozitifliği sıklığını ve bunun klinik korelasyonlarını değerlendirmektir.

Gereç ve Yöntem: Bu retrospektif, tek merkezli kohort çalışmasında 524 AS, 380 PsA, 28 İBH-ilişkili SpA, 45 pSpA ve 236 nr-axSpA hastasına ait veriler analiz edildi. Klinik özellikler, HLA-B27 durumu, enflamatuvar belirteçler ve görüntüleme sonuçları toplandı. İstatistiksel analizlerde ki-kare testi, Mann-Whitney U testi, Spearman korelasyonları ve lojistik regresyon kullanıldı.

Address for Correspondence: Dilara Bulut GÖKTEN MD, Tekirdağ Namık Kemal University Faculty of Medicine, Department of Internal Medicine, Division, of Rheumatology, Tekirdağ, Türkiye
E-mail: dilarabulutgokten@gmail.com **ORCID ID:** orcid.org/0000-0002-9226-7532
Received: 07.05.2025 **Accepted:** 23.06.2025 **Publication Date:** 07.10.2025

Cite this article as: Bulut Gökten D, Tozkır H, Soğur ÖA, Mercan R. HLA-B27 frequency and clinical correlates in turkish patients with various subtypes of spondyloarthritis: a single-center cohort study. Nam Kem Med J. 2025;13(3):293-301

Bulgular: HLA-B27 pozitifliği AS hastalarının %58,2'sinde, PsA hastalarının %31,1'inde, İBH-ilişkili SpA hastalarının %25,0'ında, pSpA hastalarının %40,0'ında ve nr-axSpA hastalarının %35,2'sinde pozitif saptandı. AS grubunda HLA-B27 pozitifliği, yüksek C-reaktif protein düzeyleri (p=0,001), artmış bath AS hastalık aktivite indeksi (BASDAİ) (p=0,0234) ve bath ankilozan spondilit fonksiyonel indeksi (BASFİ) (p=0,0207) skorları, erkek cinsiyet (p=0,0017) ve daha genç başlangıç yaşı (p=0,034) ile ilişkiliydi. PsA grubunda, daha yüksek PsA hastalık skorları HLA-B27 pozitifliği ile anlamlı şekilde ilişkiliydi (p<0,0001). İBH-ilişkili SpA hastalarında, HLA-B27 pozitifliği daha genç başlangıç yaşı (p=0,016), artmış BASDAİ (p=0,0003) ve BASFİ (p=0,0004) ile koreleydi. nr-axSpA grubunda, erkek cinsiyet (p=0,00015), artmış BASDAİ (p<0,0001), BASFİ (p<0,0001) ve biyolojik tedavi kullanımı (p=0,0277) HLA-B27 pozitifliği ile anlamlı ilişkiliydi; buna karşılık steroid olmayan anti-enflamatuvar ilaç yanıtı negatiflerde daha yüksekti (p=0,0058). AS hastalarında manyetik rezonans görüntüleme ile saptanan sakroileit HLA-B27 pozitifliği ile negatif korelasyon gösterdi (rho=-0,140, p=0,002).

Sonuç: HLA-B27 pozitifliği SpA alt tipleri arasında farklılık göstermekte olup, aksiyal SpA gruplarında erkek cinsiyet, hastalık aktivitesi ve fonksiyonel bozulma skorları ile anlamlı ilişkiler göstermekte ve en yüksek sıklık AS grubunda gözlenmektedir. Hastalık aktivitesi ve tedavi paternleri ile bazı ilişkiler saptansa da HLA-B27'nin genel klinik etkisi sınırlıdır. Bulgular, HLA-B27'nin SpA'daki karmaşık ve heterojen rolünü vurgulamakta ve prognostik önemini netleştirmek için ileriye dönük çalışmalara ihtiyaç olduğunu göstermektedir.

Anahtar Kelimeler: HLA-B27, spondiloartrit, seronegatif spondiloartrit

INTRODUCTION

Spondyloarthropathies (SpA) comprise a set of conditions that predominantly involve the spine and peripheral joints. SpA is typically subclassified into conditions such as ankylosing spondylitis (AS), psoriatic arthritis (PsA), arthritis associated with inflammatory bowel disease (IBD), reactive arthritis, and undifferentiated variants[1]. In recent years, these diseases have been further classified into additional subgroups: peripheral SpA (pSpA) and non-radiographic axial SpA (nr-axSpA)[2]. One of the key genetic predispositions associated with SpA is the presence of human leukocyte antigen (HLA)-B27. It has been linked to the onset of axial SpA (axSpA), with or without peripheral arthritis, and is also linked to enthesitis, acute anterior uveitis, more pronounced radiological progression, and gastrointestinal inflammation[3].

AxSpA involves spinal column and sacroiliac joints (SIJ). It is further subdivided into two categories: radiographic axSpA (r-axSpA, formerly known as AS) and nr-axSpA, the latter defined by the absence of radiographic SIJ changes[4]. A well-established relationship exists between the frequency of HLA-B27 and occurrence of SpA within various samples. The strongest link has been identified with AS, with some countries reporting HLA-B27 positivity rates of up to 90% among individuals with AS, although the literature presents varying results[5]. While HLA-B27 is highly prevalent in AS, its frequency in nr-axSpA is comparatively lower, averaging about 50%[6]. In individuals with psoriasis, PsA can develop as a chronic inflammatory disease. Estimates of axial skeleton involvement in PsA cases differ substantially, varying between 25% and 70%[7]. While HLA-B27 appears in a notable portion of axial PsA patients (14-40%) and less frequently in peripheral cases (~10%), its link to PsA is not as substantial as in AS[8,9].

The occurrence of arthralgia among individuals with IBD has been documented to vary between 6% and 46%. Arthritis associated with IBD, affects approximately 5% to 20% of individuals with IBD[10]. The frequency of HLA-B27 in individuals with arthritis related to IBD-associated arthritis is relatively low, approximately 40% in those with axial SpA and around 10% in those with pSpA. However, it is significantly higher in those with spondylitis or sacroiliitis associated with IBD, reaching approximately 60%[11]. pSpA refers to various SpA subsets, with enthesitis, arthritis, and dactylitis being the predominant clinical manifestations. HLA-B27 positivity has also been linked to the axial involvement in pSpA as well[12,13].

Considering the differences in HLA-B27 frequency in SpA observed in the literature and the inter-racial variations, this work explores the proportion of HLA-B27 among a Turkish cohort diagnosed with axSpA, nr-axSpA, pSpA, enteropathic arthritis, and PsA, and to examine the association between its presence and clinical features.

MATERIALS AND METHODS

The retrospective cross-sectional study was carried out with approval from the local ethics committee, adhering to the principles of the Helsinki Declaration, between 2020-2025, at the department of rheumatology of a university hospital utilizing data from electronic databases and hospital medical records Approval was obtained from the Non-Interventional and Clinical Research Ethics Committee of Tekirdağ Namık Kemal University (decision no: 2024.324.12.08, date: 31.12.2024). The study population comprised 380 PsA cases, 45 pSpA, 236 nr-axSpA, 28 IBD-associated SpA, and 524 AS cases. Classification of PsA was based on the CASPAR criteria[14], AS was diagnosed based on the modified New York criteria, while nr-axSpA was classified using the ASAS classification criteria[15,16]. pSpA was defined using the ASAS classification criteria for pSpA[17], while IBD-associated SpA was defined using the Amor and European Spondyloarthropathy Study Group criteria[18,19]. Each participant met at least one of the aforementioned classification criteria. In cases where patients exhibited overlapping clinical features, classification was determined according to the predominant

clinical phenotype and the criteria most strongly met. Each patient was assigned to a single SpA subtype to ensure mutually exclusive groupings and to facilitate clearer subgroup analyses. Information on demographics (age and sex), disease duration, diagnostic year, time to diagnosis, family history, biologic treatment, and non-steroidal anti-inflammatory drug (NSAID) response was collected. The presence of clinical features such as (inflammatory low back pain), peripheral arthritis, enthesitis, dactylitis, uveitis, psoriasis and IBD were evaluated.

Initial inflammatory markers, including C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR), obtained from patients' outpatient visits, were recorded. Initial composite index scores, including the bath AS disease activity index (BASDAI) and the bath AS functional index (BASFI), PsA disease activity score (PASDAS)[20] were utilized for the assessment of disease severity and physical function.

The association between these recorded measures and HLA-B27 positivity was statistically analyzed. Individuals under aged 18 years; those with incomplete laboratory or medical data; lacking radiological imaging; with uncertain diagnoses; active infections; malignancies; advanced organ failure; or who had undergone major treatment changes, defined as the initiation or switching of biologic or conventional disease-modifying anti-rheumatic drugs therapy within the three months prior to the recorded evaluation, were excluded to minimize potential acute effects of therapy on disease activity and inflammatory markers.

HLA-B27 allele detection and genotyping were carried out using the Genvinset® HLA-B27v5 kit (Blackhills Diagnostic Resources, Spain) as per the manufacturer's instructions. The QIAamp DNA Blood Mini Kit was employed to isolate genomic DNA from whole blood and QIAcube automated system (Qiagen, Germany), providing consistent yields and reducing contamination risk[21]. DNA extraction, amplification, and fluorescence-based detection were performed as described by the kit instructions.

All patients underwent imaging of the SIJ using pelvic radiographs and magnetic resonance imaging (MRI). Sacroiliitis on X-rays was evaluated in accordance with modified New York criteria[16]. MRI evaluations focused on detecting bone marrow edema, following the ASAS guidelines[15].

Statistical Analysis

All statistical procedures were carried out using SPSS Statistics 27.0. Categorical variables were compared using the chi-square test or the Fisher's exact test, depending on data suitability. The distribution of continuous variables was assessed with the Shapiro-Wilk test. For non-normally distributed data, the Mann-Whitney U test was employed. Associations between continuous variables and HLA-B27 status were explored using Spearman's rank correlation. A p-value <0.05 was considered statistically significant.

RESULTS

A total of 524 individuals diagnosed with AS were enrolled, of whom 58.2% tested positive for HLA-B27. A significantly older age at disease onset (p=0.034) was observed among HLA-B27 negative individuals, whereas male predominance was greater in the HLA-B27 positive group (p=0.002). Dactylitis was significantly associated with HLA-B27 negativity (p=0.022). No meaningful relationship was observed between HLA-B27 and peripheral arthritis, psoriasis, enthesitis, uveitis, IBD, or family history. Mean CRP level was 14.6±21.0 mg/L [interquartile range (IQR): 14.23] and mean ESR was 22.4±16.0 mm/h (IQR: 19.0). CRP levels were markedly increased in individuals with HLA-B27 positivity (p=0.001), in contrast, ESR levels showed no statistically significant variation (p=0.057). The baseline BASDAI and BASFI scores were notably elevated in HLA-B27 positive individuals compared to negatives (6.10±3.20 vs. 4.85±3.60 and 5.80±3.10 vs. 4.35±3.70), with p-values of 0.0234 and 0.0207) No significant differences were found regarding NSAID responsiveness or biological therapy use. Radiographic sacroiliitis was detected in all AS patients irrespective of HLA-B27 status. MRI positivity was significantly more frequent among HLA-B27 negative individuals. [p=0.002; odds ratio (OR)=0.41; Spearman's rho=-0.140].

A total of 380 PsA patients were analyzed. Of the patients, 31.1% exhibited HLA-B27 positivity. No meaningful correlation was observed between HLA-B27 status and either age at disease onset (p=0.845) or sex (p=0.836). Additionally, HLA-B27 positivity showed no significant association with NSAID response, use of biologic agents, or clinical manifestations. Average CRP and ESR levels were 11.48±22.40 mg/L (IQR: 9.0) and 26.24±22.06 mm/h (IQR: 25.0), respectively, and did not significantly differ according to HLA-B27 status. A mean PASDAS score of 6.06±0.9 was recorded, with significantly higher values among HLA-B27 positive individuals (p<0.0001). No significant correlation was found between radiographic sacroiliitis and HLA-B27 positivity (Spearman's rho=-0.086, p=0.125; OR=0.89) or between MRI positivity and HLA-B27 status (Spearman's rho=-0.055, p=0.345; OR=0.77).

Twenty-eight patients with IBD-SpA were included. One-quarter (25.0%) of the cohort tested positive for HLA-B27. Disease onset time was significantly earlier in HLA-B27 positive patients (p=0.016), while no association was found between HLA-B27 status and sex (p=1.000). Clinical features did not significantly differ according to HLA-B27 status. Average CRP and ESR levels were 19.30± 3.41 mg/L (median: 8.65 mg/L, IQR: 6.6 mg/L) and 38.07±25.63 mm/h (median: 31.0 mm/h, IQR: 41.0

mm/h), with no significant differences according to HLA-B27 status. NSAID response and biologic therapy use were similarly unrelated to HLA-B27 positivity. Elevated BASDAI and BASFI scores were significantly associated with HLA-B27 positivity (BASDAI: 7.58±0.41 vs. 5.67±0.64, p=0.0003; BASFI: 6.99±0.44 vs. 5.49±0.73, p=0.0004). All patients exhibited radiographic sacroiliitis, regardless of HLA-B27 status (p=1.0), whereas MRI-detected sacroiliitis was negatively correlated with HLA-B27 positivity (rho=-0.707, p<0.001) (see Table 1 and 2).

Among the 45 individuals with pSpA, HLA-B27 positivity was identified in 40.0%. No meaningful associations were found between HLA-B27 status and sex or disease onset time. None of the clinical features, including dactylitis, enthesitis, or uveitis, were significantly associated with HLA-B27 positivity. Mean CRP and ESR levels were 11.24±16.01 mg/L and 26.32±19.13 mm/h. CRP, ESR, NSAID response and biologic usage showed no significant differences according to HLA-B27 status. Radiographic sacroiliitis was observed in 35% of pSpA patients. MRI-confirmed sacroiliitis was present in 77.8% of HLA-B27 positive patients. Sacroiliitis findings on both X-ray and MRI showed no significant correlation with HLA-B27 status.

Among the 236 nr-axSpA patients evaluated, HLA-B27 positivity was notably higher in males compared to females (p=0.00015) while no significant variation was found in the mean age at symptom onset. No significant relationship was found between HLA-B27 positivity and clinical features, CRP (18.27±34.40 vs. 11.57±20.06 mg/L; p=0.103), or ESR levels (29.25±24.71 mm/h; p=0.914). However, HLA-B27 positive patients had significantly higher disease activity and functional limitation scores, as reflected by BASDAI (6.73±0.67 vs. 5.66±0.70, p<0.0001) and BASFI (6.49±0.64 vs. 5.29±0.68, p<0.0001). Biologic treatment was more frequently used among HLA-B27 positive patients (19.3% vs. 8.5%, p=0.0277), whereas NSAID response was significantly more common in HLA-B27 negative patients (90.2% vs. 75.9%, p=0.0058). MRI positivity showed no significant correlation with HLA-B27 status (Spearman's rho=0.037, p=0.683) (see Table 3).

Regarding medications, biologic therapy was most commonly used in patients with AS (55.3%) and PsA (40.5%), followed by enteropathic arthritis (50.0%), pSpA (15.6%), and nr-axSpA (12.3%). Across all subtypes, tumor necrosis factor (TNF) inhibitors were the predominant class. The most frequently prescribed agents included adalimumab (n=142), etanercept (n=125), infliximab (n=73), golimumab (n=39), and certolizumab pegol (n=31). IL-17 inhibitors were used in 40 AS, 22 PsA, and 2 enteropathic arthritis patients, while JAK inhibitors were prescribed to 26 AS, 13 PsA, and 1 enteropathic arthritis patient. All biologic users in the nr-axSpA and pSpA groups were treated exclusively with TNF inhibitor.

Table 1. Demographic characteristics of all patients

Variables	Ax-SpA	PsA	EA	pSpA	Nr-AxSpA
Patient number (n)	524	380	28	45	236
Age, years (mean ± SD)	45.9±11.2	49.55±11.72	45.07±14.38	44.69±11.25	48.58±11.00
Symptom duration (onset) (years, mean ± SD)	12.02±7.55	9.23±7.89	6.48±5.67	5.13±4.22	8.14±608
Diagnostic delay (years, mean ± SD)	4.18±3.26	5.72±4.02	4.85±3.79	4.18±3.24	5.33±4.45
Sex (male, %)	320 (61.1)	134 (35.3)	12 (42.9)	12 (26.7)	49 (20.8)
HLA-B27 positivity (n, %)	305 (58.2)	118 (31.1)	7 (25)	18 (40)	83 (35.2)
Arthritis (n, %)	132 (25.2)	227 (59.7)	6 (21.4)	37 (82.2)	102 (43.2)
iLBP (n, %)	347 (66.2)	57 (15)	14 (50)	15 (33.3)	172 (72.9)
Dactylitis (n, %)	28 (5.3)	130 (34.2)	4 (14.3)	8 (17.8)	32 (13.6)
Psoriasis (n, %)	33 (6.3)	379 (100)	2 (7.1)	0 (0)	10 (4.2)
Enthesitis (n, %)	152 (29)	116 (30.5)	10 (35.7)	19 (42.2)	77 (32.6)
Uveitis (n, %)	78 (14.9)	13 (3.4)	2 (7.1)	3 (6.7)	12 (5.1)
Family history positivity (n, %)	117 (22.3)	55 (14.5)	6 (21.4)	10 (22.2)	40 (16.9)
IBD (n, %)	21 (4)	5 (1.3)	28 (100)	0 (0)	3 (1.3)
NSAID responsiveness (n, %)	237 (45.2)	226 (59.5)	10 (35.7)	44 (97.8)	201 (85.2)
Biologic usage (n, %)	290 (55.3)	154 (40.5)	14 (50)	7 (15.6)	29 (12.3)

Ax-Spa: Axial spondyloarthritis (ankylosing spondylitis), PsA: Psoriatic arthritis, EA: Enteropathic arthritis (inflammatory bowel disease associated spondyloarthritis), pSpA: Peripheral spondyloarthritis, Nr-AxSpA: Non radiographic axial spondyloarthritis, SD: Standard deviation, IBD: Inflammatory bowel disease, NSAID: Non-steroidal anti inflammatory drugs, iLBP: Inflammatory low back pain

Table 2. Comparison of clinical parameters according to HLA-B27 status in ankylosing spondylitis, psoriatic arthritis and inflammatory bowel disease associated spondyloarthritis cohorts

Variable	AS HLA-B27 Positive n (%)	AS HLA-B27 Negative n (%)	p-value	PsA HLA-B27 Positive n (%)	PsA HLA-B27 Negative n (%)	p-value	EA HLA-B27 Positive n (%)	EA HLA-B27 Negative n (%)	p-value
Male sex	204 (66.9)	101 (53.0)	**0.002**	43 (36.4)	91 (34.7)	0.836	3 (42.9)	9 (42.9)	1.000
iLBP	203 (66.5)	144 (65.7)	0.922	16 (13.5)	41 (15.6)	0.709	5 (71.4)	9 (42.9)	0.383
Sacroiliitis (X-ray)	305 (100)	219 (100)	1.000	31 (26.2)	86 (32.8)	0.125	7 (100)	21 (100)	1.000
Sacroiliitis (MRI positivity)	248 (81.3)	200 (91.3)	**0.002**	31 (26.2)	83 (31.6)	0.345	3 (42.9)	21 (100)	**0.0018**
Peripheral arthritis	74 (24.2)	58 (26.4)	0.634	66 (55.9)	161 (61.5)	0.367	3 (42.9)	3 (14.3)	0.287
Psoriasis	21 (6.8)	12 (5.4)	0.638	117 (99.2)	262 (100.0)	0.682	2 (28.6)	0 (0.0)	0.056
Enthesitis	88 (28.8)	64 (29.2)	1.000	40 (33.8)	76 (29)	0.402	2 (28.6)	8 (38.1)	1.000
Dactylitis	10 (3.2)	18 (8.2)	**0.022**	36 (30.5)	94 (35.8)	0.366	2 (28.6)	2 (9.5)	0.253
Family history	75 (24.5)	42 (19.1)	0.174	17 (14.4)	38 (14.5)	1.000	0 (0.0)	6 (28.6)	0.287
Uveitis	52 (17)	26 (11.8)	0.129	4 (3.3)	9 (3.4)	1.000	0 (0.0)	2 (9.5)	1.000
IBD	12 (3.9)	9 (4.1)	1.000	1 (0.8)	4 (1.5)	0.956	7 (100)	21 (100)	1.000

Statistically significant values are marked at p<0.05

AS: Ankylosing spondylitis, PsA:Psoriatic arthritis, EA: Enteropathic arthritis, iLBP: Inflammatory low back pain, HLA-B27: Human leukocyte antigen B27, MRI: Magnetic resonance ımaging, IBD: Inflammatory bowel disease

Table 3. Comparison of clinical parameters according to HLA-B27 status in peripheral spondyloarthritis and non-radiographic axial spondyloarthritis cohorts

Variable	pSpA HLA-B27 Positive n (%)	pSpA HLA-B27 Negative n (%)	p-value	Nr-axSpA HLA-B27 Positive n (%)	Nr-axSpA HLA-B27 Negative n (%)	p-value
Male sex	6 (33.3)	6 (22.2)	0.499	29 (34.9)	20 (13.1)	**0.00015**
iLBP	7 (38.9)	8 (29.6)	0.538	64 (77.1)	108 (70.6)	0.356
Sacroiliitis (X-ray)	8 (40)	8 (29.6)	0.281	0 (0.0)	0 (0.0)	1.000
Sacroiliitis (MRI positivity)	14 (77.8)	15 (55.6)	0.227	45 (54.2)	78 (51.0)	0.683
Peripheral arthritis	13 (72.2)	24 (88.9)	0.301	32 (38.6)	70 (45.8)	0.353
Psoriasis	0 (0.0)	0 (0.0)	1.000	3 (3.6)	7 (4.6)	0.991
Enthesitis	4 (22.2)	15 (55.6)	0.056	24 (28.9)	53 (34.6)	0.453
Dactylitis	3 (16.7)	5 (18.5)	1.000	13 (15.7)	19 (12.4)	0.620
Family history	4 (22.2)	6 (22.2)	1.000	15 (18.1)	25 (16.3)	0.875
Uveitis	2 (11.1)	1 (3.7)	0.714	9 (10.8)	3 (2)	**0.0079**
IBD	0 (0.0)	0 (0.0)	1.000	1 (1.2)	2 (1.3)	1.000

Statistically significant values are marked at p<0.05

PSpA: Peripheral spondyloarthritis, Nr-AxSpA: Non radiographic axial spondyloarthritis, iLBP: Inflammatory low back pain, HLA-B27: Human leukocyte antigen B27, MRI: Magnetic resonance imaging, IBD: Inflammatory bowel disease

DISCUSSION

This study assessed HLA-B27 positivity and its clinical, laboratory, and radiological correlations in a cohort comprising AS, PsA, nr-axSpA, pSpA, and enteropathic arthritis patients from Türkiye's Thrace region. Among 524 AS patients, the HLA-B27 frequency was 58.2%; among 380 PsA patients, 31.1%; among 28 IBD-associated SpA patients, 25%; among 45 pSpA patients, 40%; and among 236 nr-axSpA patients, 35.2%. These rates were lower for AS compared to Western populations but appeared comparable to other Turkish cohorts. For PsA, the frequency was similar to data reported from North America and Brazil[22,23]. The rate observed for enteropathic arthropathy was lower than that reported for enteropathic spondylitis in the literature but higher than that reported specifically for IBD-associated arthritis[22]. The frequency observed for pSpA was similar to international studies, such as the ASAS perSpA study, which reported a prevalence of 35.8%[13]. In contrast, the rate observed in nr-axSpA was somewhat lower compared to certain other Turkish cohorts. In the overall cohort, HLA-B27 positivity showed no significant association with most clinical manifestations. Despite early studies suggesting a link between HLA-B27 positivity and increased extra-articular involvement, such an association was not observed in this study. A significant relationship was observed between HLA-B27 positivity and initial BASDAI and BASFI scores in all groups; however, no such correlation was found with CRP or ESR. Among AS and nr-axSpA patients, HLA-B27 positivity was significantly associated with male sex although the relationship with age at onset was inconsistent across groups. Moreover, HLA-B27 positivity correlated with more severe radiographic sacroiliitis in AS and IBD-associated SpA while no such association was found in the remaining subtypes.

In this cohort, an HLA-B27 positivity rate of 58.2% was identified among AS patients, which is slightly lower than the 70–91% reported in Turkish cohorts[24,25], the 69% reported in Qatar by Abdelrahman et al.[26] (including 82% among Qataris, 72% among Jordanians, and 90% among Egyptians), and the 80.5% reported in a Greek cohort[27], but remains within the broader Middle Eastern range of 26.2% (Lebanon) to 91% (Türkiye)[28]. HLA-B27 carriage rates in the general population of Arab and Middle Eastern nations (0.3-6.8%) has been noted to be considerably lower than in Western populations (6-25%)[28], potentially affecting the diagnostic utility of HLA-B27-based referral strategies in these regions. Furthermore, the considerable methodological heterogeneity across regional studies, such as variations in sample sizes, classification criteria, and HLA-B27 testing techniques, should be carefully considered when comparing prevalence estimates across populations. The differences observed in HLA-B27 prevalence across studies, even within a single country, are likely influenced by genetic and environmental diversity among subpopulations. HLA-B27 positivity was significantly associated with male sex and earlier disease onset, consistent with findings from Arévalo et al.[29], as well as the DESIR and GESPIC cohorts[30,31]. Unlike Arévalo et al.[29], who found no CRP or ESR differences, this research observed significantly higher CRP levels in HLA-B27-positive patients, though ESR was similar. With respect to disease activity, our finding of elevated BASDAI and BASFI scores in HLA-B27 positive patients differs from Arévalo et al.[29], who observed higher values in those lacking HLA-B27. Radiographically, current study paradoxically showed an inverse relationship with MRI sacroiliitis, an area not assessed in Arévalo et al.[29].

In this cohort, HLA-B27 positivity showed no significant association with uveitis, enthesitis, peripheral arthritis, or IBD; however, dactylitis was significantly more common among HLA-B27 positive individuals. These findings contrast with the study of Zhang et al.,[32] where HLA-B27 positivity was linked to higher uveitis prevalence but lower rates of dactylitis and peripheral arthritis, with no clear association for enthesitis or IBD. Similarly, another Chinese cohort reported no significant differences in uveitis, enthesitis, or peripheral arthritis between HLA-B27-positive and negative AS patients, aligning with the present findings, although no relationship with dactylitis was identified in their cohort[33]. Collectively, these findings indicate that the association between HLA-B27 and extra-articular manifestations is multifaceted and potentially population-dependent.

The HLA-B27 positivity rate among PsA patients in this study was 31.1%, aligning with international reports such as the 27.3% observed by Ruiz et al.[23] and the 22.8% in a Sri Lankan SpA cohort[34], though slightly higher than the 17.6% reported by Öğretmen et al.[35]. While some studies, such as those by Ruiz et al.[23] and Kidnapillai et al.,[34] noted a male predominance among HLA-B27-positive patients, no significant association with male sex was observed here. Although earlier research has indicated a strong connection between HLA-B27 positivity and axial manifestations, especially given the reported 60% positivity rate in PsA patients with axial involvemen, the predominance of pSpA in this cohort may explain the absence of such associations. While Bonfiglioli et al.[37] identified significant associations between HLA-B27 positivity, male sex, and axial involvement in PsA, no analyses were conducted regarding composite disease activity measures such as PASDAS. Importantly, no prior research has directly explored the link between HLA-B27 status and PASDAS in PsA, with most studies instead concentrating on axial involvement measured by BASDAI or AS disease activity score (ASAS). The significant association identified here, with higher PASDAS values among HLA-B27-positive patients, suggests a potentially meaningful link between HLA-B27 status and overall disease activity in PsA.

Within this group of patients diagnosed with IBD-associated SpA, the frequency of HLA-B27 was identified as 25%, which is comparable to the 29% documented by Turkcapar et al.[38] and falls within the broader frequency range of 30-80% reported by Peluso et al.[39]. However, this rate is lower than the 46.7% noted in the Guinea cohort[40] but exceeds the 7.9% described by Huber et al.[41]. Similarly, the Brazilian research by Toledo et al.[42] found no link between HLA-B27 and enteropathic SpA, supporting the observation that HLA-B27 negativity is common in IBD-related SpA cases. Notably, Toledo et al.[42] highlighted that radiological sacroiliitis was significantly associated with HLA-B27, even beyond AS, whereas intestinal involvement tended to occur more frequently in HLA-B27-negative individuals. Regarding disease activity, HLA-B27 showed a significant connection with BASDAI and BASFI scores, reflecting its potential impact on functional status, although earlier studies offered limited insights into these measures. Notably, no significant link was identified between HLA-B27 positivity and radiographic sacroiliitis in this cohort; however, an inverse association with MRI findings was observed, highlighting a complex interplay between genetic markers and clinical expression. Across different studies, NSAID use was generally low, likely due to gastrointestinal safety concerns in IBD, while the proportion of patients on biologic therapies varied (for instance, 39.4% reported by Huber et al.[41]).

In this cohort of pSpA, the HLA-B27 positivity rate was 40%, notably higher than the ~27% reported in ASAS-based international cohorts[43]. Özsoy et al.[44] reported a lower HLA-B27 positivity specifically among Turkish pSpA patients (9.8%) although their overall SpA cohort showed a higher rate (74.7%) largely due to axial cases. Similarly, the ASAS perSpA study found a 35.8% HLA-B27 positivity in pSpA/PsA patients, reporting associations with earlier disease onset, male sex, axial involvement, tarsitis, and uveitis, but no link with dactylitis, psoriasis, or IBD[13]. While prior research has consistently shown that pSpA displays a weaker association with HLA-B27 compared to axial SpA[43], a similar pattern was observed here, with no significant associations identified between HLA-B27 status and extra-articular manifestations, inflammatory markers, NSAID response, or MRI-detected sacroiliitis. Interestingly, Özsoy et al.[44] also found no significant differences between HLA-B27 carriers and non-carriers in terms of dactylitis, enthesitis, or IBD; however, they reported a higher prevalence of uveitis among individuals expressing HLA-B27. Moreover, Arevalo Salaet et al.[13] emphasized that despite HLA-B27's links to axial features and uveitis, no association was observed with peripheral joint damage, reinforcing that structural damage pathways may be independent of HLA-B27 status.

In the current study, the HLA-B27 positivity rate among nr-axSpA patients was 35.2%, placing it within the lower-to-mid global range and notably below the 63.3% reported in the Turkish nr-axSpA cohort[45], as well as the 41.9% observed in Malaysia[46], 54.3% in Mexico[6] and 72.2% in a large cohort from China[47]. Male sex was significantly associated with HLA-B27 positivity, consistent with previous observation in Mexican cohort. No significant associations were found between HLA-B27 status and acute-phase reactants, MRI findings, or specific clinical features such as uveitis or dactylitis, aligning with earlier reports where these associations were not prominently emphasized. While Özdemirel et al.[45] reported no relationship between HLA-B27 and disease activity in Turkish nr-axSpA patients, the present study identified significant associations between HLA-B27 positivity and higher BASDAI and BASFI scores, indicating greater subjective disease burden. This may partly explain the more frequent use of biologic therapies and the lower responsiveness to NSAIDs among HLA-B27-positive individuals.

This appears to be the first study to systematically evaluate HLA-B27 prevalence across multiple SpA subtypes, including AS, nr-axSpA, IBD-associated SpA, PsA, and pSpA, in the Thrace region of Türkiye. While the key clinical characteristics of Turkish SpA patients were consistent with previous studies, several distinct differences were observed in this cohort. One of the most striking findings of the present study was the relatively low frequency of HLA-B27 positivity across all SpA subtypes, including AS (58.2%) and nr-axSpA (38.1%). These rates are notably lower than those reported in several European cohorts, where HLA-B27 positivity exceeds 80% in AS and approximately 60-70% in nr-axSpA. In contrast, our findings are more consistent with regional data from the Middle East and Mediterranean countries, where the prevalence of HLA-B27 in SpA patients tends to be lower, possibly due to ethnic and genetic diversity. Additionally, population-based studies have shown that the general prevalence of HLA-B27 in Türkiye is considerably lower than in Northern European countries, which may contribute to the relatively low detection rates observed in our cohort. Environmental exposures, differing referral patterns, and variability in disease phenotypes could also play a role in this discrepancy. These findings underscore the importance of considering regional genetic backgrounds when interpreting HLA-B27 positivity in clinical practice and research.

Study Limitations

This study has several limitations inherent to its retrospective and single-center design. Although standardized classification criteria were used, diagnostic accuracy may have been influenced by variability in physician assessments over time. The absence of longitudinal follow-up precluded the evaluation of disease progression and long-term outcomes, particularly relevant for nr-axSpA patients who may later transition to

radiographic disease. Small sample sizes in certain subgroups may have limited the statistical power. All radiographic and MRI evaluations were performed by a single radiologist, which, while ensuring consistency, may have introduced observer bias due to the lack of inter-reader validation. Additionally, data on smoking status and comorbidities were not consistently documented and were therefore excluded, potentially limiting the depth of clinical characterization. In patients with pSpA, joint activity could not be assessed using standardized joint counts or composite activity scores due to inconsistent recording in the medical files. Lastly, genetic and environmental factors specific to the study population may limit the generalizability of these findings. Future large-scale, prospective, multicenter studies are needed to validate these results and further elucidate the clinical significance of HLA-B27 across diverse SpA subtypes.

CONCLUSION

This study provides a comprehensive overview of HLA-B27 frequency and its clinical associations across multiple SpA subtypes in a large Turkish cohort. While HLA-B27 positivity rates varied between disease groups, its overall clinical impact appeared limited, with no consistent associations identified for extra-articular manifestations or inflammatory markers. Importantly, elevated disease activity and functional limitation scores were observed in HLA-B27–positive individuals, most prominently in those with axial SpA and nr-axSpA. These findings underscore the complex and subtype-specific nature of HLA-B27's clinical relevance. Future multicenter, prospective research are needed to confirm these observations and further clarify prognostic role of HLA-B27 in diverse populations.

Ethics

Ethics Committee Approval: Approval was obtained from the Non-Interventional and Clinical Research Ethics Committee of Tekirdağ Namık Kemal University (decision no: 2024.324.12.08, date: 31.12.2024).

Informed Consent: This is retrospective cross-sectional study.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Pan H, Ye Y, Wang D, Zheng J, Zhao Z, Zhao K. Ankylosing spondylitis and the genetic correlation of extra-articular manifestations: evidence from bidirectional Mendelian randomization and Bayesian colocalization. Medicine (Baltimore). 2025;104:e41388.
2. Robinson PC, Sengupta R, Siebert S. Non-radiographic axial spondyloarthritis (nr-axSpA): advances in classification, imaging and therapy. Rheumatol Ther. 2019;6:165-77.
3. Navid F, Chen L, Bowness P, Colbert RA. HLA-B27 and spondyloarthritis: at the crossroads of innate and adaptive immunity. Nat Rev Rheumatol. 2025;21:77-87.
4. Boonen A, Sieper J, van der Heijde D, Dougados M, Bukowski JF, Valluri S, et al. The burden of non-radiographic axial spondyloarthritis. Semin Arthritis Rheum. 2015;44:556-62.
5. Akassou A, Bakri Y. Does HLA-B27 Status influence ankylosing spondylitis phenotype? Clin Med Insights Arthritis Musculoskelet Disord. 2018;11:1179544117751627.
6. Londono J, Pacheco-Tena C, Santos AM, Cardiel MH, Rodríguez-Salas G, Rueda I, et al. Differences between radiographic and non-radiographic axial spondyloarthritis patients in a Mexican cohort. Sci Rep. 2024;14:10342.
7. Lopez-Medina C, Ziade N. Axial disease in psoriatic arthritis: how can we define it, and does it have an impact on treatment? Mediterr J Rheumatol. 2022;33:142-9.
8. Chandran V, Tolusso DC, Cook RJ, Gladman DD. Risk factors for axial inflammatory arthritis in patients with psoriatic arthritis. J Rheumatol. 2010;37:809-15.
9. Feld J, Ye JY, Chandran V, Inman RD, Haroon N, Cook R, et al. Is axial psoriatic arthritis distinct from ankylosing spondylitis with and without concomitant psoriasis? Rheumatology. 2019;59:1340-6.
10. Shahid Z, Brent L, Lucke M. Enteropathic Arthritis. StatPearls. 2024.
11. Rios Rodriguez V, Duran TI, Torgutalp M, López-Medina C, Dougados M, Kishimoto M, et al. Comparing clinical profiles in spondyloarthritis with Crohn's disease or ulcerative colitis: insights from the ASAS-PerSpA study. Rheumatol Adv Pract. 2024;8:rkae064.
12. Carron P, De Craemer AS, Van den Bosch F. Peripheral spondyloarthritis: a neglected entity-state of the art. RMD Open. 2020;6:e001136.
13. Arevalo Salaet M, López-Medina C, Moreno M, Navarro-Compan V, Calvet Fontova J, Llop M, et al. Association between HLA-B27 and peripheral spondyloarthritis phenotype: results from the ASAS perSpA study. RMD Open. 2022;8:e002696.
14. Taylor W, Gladman D, Helliwell P, Marchesoni A, Mease P, Mielants H, et al. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum. 2006;54:2665-73.
15. Rudwaleit M, van der Heijde D, Landewé R, Listing J, Akkoc N, Brandt J, et al. The development of assessment of spondyloarthritis international society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis. 2009;68:777-83.
16. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum. 1984;27:361-8.
17. Rudwaleit M, van der Heijde D, Landewé R, Akkoc N, Brandt J, Chou CT, et al. The assessment of spondyloarthritis international society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis. 2011;70:25-31.
18. Dougados M, van der Linden S, Juhlin R, Huitfeldt B, Amor B, Calin A, et al. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. Arthritis Rheum. 1991;34:1218-27.
19. Amor B, Dougados M, Mijiyawa M. Critères de classification des spondylarthropathies [Criteria of the classification of spondylarthropathies]. Rev Rhum Mal Osteoartic. 1990;57:85-9.

20. Helliwell PS, Deodhar A, Gottlieb AB, Boehncke WH, Xu XL, Xu S, et al. Composite measures of disease activity in psoriatic arthritis: comparative instrument performance based on the efficacy of guselkumab in an interventional phase II trial. Arthritis Care Res (Hoboken). 2020;72:1579-88.

21. Geiger K, Zach C, Leiherer A, Fraunberger P, Drexel H, Muendlein A. Real-time PCR based HLA-B*27 screening directly in whole blood. HLA. 2020;95:189-95.

22. Hochberg MC, Gravallese EM, Smolen JS, van der Heijde D, Weinblatt ME, Weisman MH. Rheumatology E-Book: Elsevier Health Sciences; 2022.

23. Ruiz DG, Azevedo MN, Lupi O. HLA-B27 frequency in a group of patients with psoriatic arthritis. An Bras Dermatol. 2012;87:847-50.

24. Gunal EK, Sarvan FO, Kamali S, Gul A, Inanc M, Carin M, et al. Low frequency of HLA-B27 in ankylosing spondylitis patients from Turkey. Joint Bone Spine. 2008;75:299-302.

25. Akkoç N, Yarkan H, Kenar G, Khan MA. Ankylosing spondylitis: HLA-B*27-positive versus HLA-B*27-negative disease. Curr Rheumatol Rep. 2017;19:26.

26. Abdelrahman MH, Mahdy S, Khanjar IA, Siam AM, Malallah HA, Al-Emadi SA, et al. Prevalence of HLA-B27 in patients with ankylosing spondylitis in Qatar. Int J Rheumatol. 2012;2012:860213.

27. Alamanos Y, Papadopoulos NG, Voulgari PV, Karakatsanis A, Siozos C, Drosos AA. Epidemiology of ankylosing spondylitis in Northwest Greece, 1983-2002. Rheumatology (Oxford). 2004;43:615-8.

28. Ziade NR. HLA B27 antigen in Middle Eastern and Arab countries: systematic review of the strength of association with axial spondyloarthritis and methodological gaps. BMC Musculoskelet Disord. 2017;18:280.

29. Arévalo M, Gratacós Masmitjà J, Moreno M, Calvet J, Orellana C, Ruiz D, et al. Influence of HLA-B27 on the ankylosing spondylitis phenotype: results from the REGISPONSER database. Arthritis Res Ther. 2018;20:221.

30. Chung HY, Machado P, van der Heijde D, D'Agostino MA, Dougados M. HLA-B27 positive patients differ from HLA-B27 negative patients in clinical presentation and imaging: results from the DESIR cohort of patients with recent onset axial spondyloarthritis. Ann Rheum Dis. 2011;70:1930-6.

31. Rudwaleit M, Haibel H, Baraliakos X, Listing J, Märker-Hermann E, Zeidler H, et al. The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis Rheum. 2009;60:717-27.

32. Zhang S, Wang Y, Peng L, Su J, Zeng X, Li M, et al. Comparison of clinical features in HLA-B27 positive and negative patients with axial spondyloarthritis: results from a cohort of 4,131 patients. Front Med (Lausanne). 2020;7:609562.

33. Yang M, Xu M, Pan X, Hu Z, Li Q, Wei Y, et al. Epidemiological comparison of clinical manifestations according to HLA-B*27 carrier status of Chinese ankylosing spondylitis patients. Tissue Antigens. 2013;82:338-43.

34. Kidnapillai S, Sirisena ND, Dissanayake VH. HLA-B27 allele frequency in Sri Lankan patients with spondyloarthritides. Ceylon Med J. 2016;61:71-3.

35. Öğretmen Z, Hız MM, Sılan F, Koşar Ş, Özdemir Ö. Is the HLA B27 genotype a risc faktor for psoriatic arthritis and psoriasis vulgaris? TURKDERM - Turkish Archives of Dermatology and Venereology. 2014;48:131-4.

36. Sheehan NJ. The ramifications of HLA-B27. J R Soc Med. 2004;97:10-4.

37. Bonfiglioli R, Conde RA, Sampaio-Barros PD, Louzada-Junior P, Donadi EA, Bertolo MB. Frequency of HLA-B27 alleles in Brazilian patients with psoriatic arthritis. Clin Rheumatol. 2008;27:709-12.

38. Turkcapar N, Toruner M, Soykan I, Aydintug OT, Cetinkaya H, Duzgun N, et al. The prevalence of extraintestinal manifestations and HLA association in patients with inflammatory bowel disease. Rheumatol Int. 2006;26:663-8.

39. Peluso R, Di Minno MN, Iervolino S, Manguso F, Tramontano G, Ambrosino P, et al. Enteropathic spondyloarthritis: from diagnosis to treatment. Clin Dev Immunol. 2013;2013:631408.

40. Kaba C, Mohamed A, Guelngar CO, Emmanuel A, Aminou SY, Diallo MH, et al. Spondyloarthritis associated with inflammatory bowel disease: study of 15 observations in guinea. Open Journal of Rheumatology and Autoimmune Diseases. 2021;11:89-95.

41. Huber MK, Valim V, Serrano ÉV, Mendonça JA, Lourenço RB, Espírito Santo TMR, et al. Prevalence of spondyloarthritis in inflammatory bowel disease according ASAS and ultrassonography and its correlation with plasma calprotectin. Adv Rheumatol. 2024;64:27.

42. Toledo RAd, Toledo RAd, Camargo U, Camargo AVdS, Xavier DH, Batista MF, et al. HLA-B*27 – Frequency of clinical signs in Brazilian patients with spondyloarthritis. Biomarkers and Genomic Medicine. 2015;7:72-7.

43. Paramarta J, Van de Schoot M, Turina M, Ambarus C, Bijlsma J, De Rycke L, et al. FRI0459 Hla-b27 positive versus hla-b27 negative spondyloarthritis: different phenotypes but similar disease burden. Annals of the Rheumatic Diseases. 2013;72:A530.

44. Özsoy Z, Öztürk Y, Öztürk Z, Beyan E. HLA-B27 positivity and associated factors in spondyloartritis patients from Turkey: a single-center study. Anatolian Curr Med J / ACMJ / acmj. 2025;7:223-9.

45. Özdemirel AE, Tecer D, Torgutalp ŞŞ, Küçükşahin O. Prevalence of HLA-B27 in Turkish patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis. Journal of Physical Medicine & Rehabilitation Sciences. 2021;24:17-22.

46. Ong PS, Nur-Aida-Sabrina MR, Haziqah-Itqan A, Lau S, Mollyza MZ, Azmillah R, et al. Comparison of non-radiographic axial spondyloarthritis and ankylosing spondylitis patients in Malaysia. The Egyptian Rheumatologist. 2022;44:357-60.

47. Huang SX, Li HG, Liang HJ, Wang DM, Peng JH, Shen FC, et al. Comparison of clinical characteristics between adult-onset and juvenile-onset non-radiographic axial spondyloarthritis in Chinese patients: results from the COCAS cohort. Eur J Med Res. 2023;28:388.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.71677
Nam Kem Med J 2025;13(3):302-310

Empagliflozin Inhibits OSCC Proliferation and Migration by Suppressing SLC2A3 and NLRP3

Empagliflozin Oral Squamöz Hücreli Karsinomada SLC2A3 ve NLRP3 İfadesini Baskılayarak Proliferasyonu ve Migrasyonu Engeller

Gülçin TEZCAN

Bursa Uludağ University Faculty of Dentistry, Department of Fundamental Sciences, Bursa, Türkiye

ABSTRACT

Aim: High glucose and inflammation facilitate oral squamous cell carcinoma (OSCC) progression. The glucose transporter solute carrier family 2 member 3 (SLC2A3) and the nucleotide-binding oligomerization domain, leucine-rich repeat-containing family pyrin domain containing 3 (NLRP3), contribute to chronic inflammation and poor prognosis. This study investigates the regulatory effects of the antidiabetic drug empagliflozin (emp) on SLC2A3 and NLRP3 in OSCC and their impact on cancer cell growth.

Materials and Methods: Human OSCC cell lines, SCCL-MT1 and UPSI-SCC-131, were cultured in Dulbecco's Modified Eagle Medium with supplements. Emp, lipopolysaccharide, and adenosine triphosphate were used for treatments. Real-time polymerase chain reaction quantified SLC2A3 and NLRP3 expressions. Enzyme-linked immunosorbent assay measured interleukin-1beta (IL-1β) levels, and cell proliferation was assessed using the xCELLigence system. Migration was evaluated via a scratch wound assay. Protein-protein interactions were predicted using the STRING database. Data analysis was conducted using GraphPad Prism.

Results: Emp, an SGLT2 inhibitor, significantly reduced SLC2A3 ($p<0.0001$, $p=0.0002$; respectively) and NLRP3 expression ($p=0.0008$; $p=0.0006$; respectively), leading to decreased IL-1β release ($p=0.0190$, $p<0.0001$; respectively), proliferation ($p=0.024$; $p<0.0001$; respectively), and migration ($p=0.0021$; $p=0.0004$, respectively) in SCCL-MT1 and UPSI-SCC131 OSCC cell lines. These findings suggest emp's potential as a therapeutic agent for OSCC by targeting glucose metabolism and inflammation.

Conclusion: These findings suggest that emp effectively modulates glucose metabolism and inflammation in OSCC by inhibiting SLC2A3 and NLRP3 expression and IL-1β release, thereby reducing cancer cell proliferation and migration, highlighting its possible role as a treatment option for managing OSCC progression.

Keywords: Oral squamous cell carcinoma, SLC2A3, NLRP3, empagliflozin

ÖZ

Amaç: Oral squamöz hücreli karsinom (OSCC), ilerlemesi için yüksek glikoz metabolizması ve enflamasyona bağımlıdır. Glukoz taşıyıcı çözücü taşıyıcı ailesi 2 üyesi 3 (SLC2A3), kötü prognoz ile ilişkilidir, nükleotid bağlayıcı oligomerleşme bölgesi, lösin-zengin tekrar içeren aile pirin domain içeren 3 (NLRP3) ise kronik iltihaplanmaya katkı sağlar. Bu çalışma, antidiabetik ilaç empagliflozinin (emp) OSCC'de SLC2A3 ve NLRP3 üzerindeki düzenleyici etkilerini ve kanser hücresi büyümesi üzerindeki etkilerini araştırmaktadır.

Gereç ve Yöntem: İnsan OSCC hücre hatları SCCL-MT1 ve UPSI-SCC-131, suplementlerle zenginleştirilmiş Dulbecco's Modified Eagle Medium kültüre edilmiştir. Lipopolisakkarit maruziyeti ile enflamasyon modellenmiş ve adenozin trifosfat ile enflamazom aktivasyonu sağlanmıştır. Emp tedavi için kullanılmıştır. SLC2A3 ve NLRP3 ifadeleri, gerçek zamanlı polimeraz zincir reaksiyonu ile belirlenmiştir. Enzim bağlantılı immünosorbent testi, interlökin-1 beta (IL-1β) düzeylerini ölçerken, hücre proliferasyonu xCELLigence sistemi ile değerlendirilmiştir. Migrasyon sıyrık yara testi ile değerlendirilmiştir. Protein-protein etkileşimleri STRING veritabanı kullanılarak tahmin edilmiştir. İstatistiksel analiz GraphPad Prism ile yapılmıştır.

Address for Correspondence: Gülçin TEZCAN MD, Bursa Uludağ University Faculty of Dentistry, Department of Fundamental Sciences, Bursa, Türkiye
E-mail: gulcintezcan@uludag.edu.tr **ORCID ID:** orcid.org/0000-0002-5956-8755
Received: 04.03.2025 **Accepted:** 25.06.2025 **Publication Date:** 07.10.2025

Cite this article as: Tezcan G. Empagliflozin inhibits OSCC proliferation and migration by suppressing SLC2A3 and NLRP3. Nam Kem Med J. 2025;13(3):302-310

Bulgular: Emp, bir SGLT2 inhibitörü olarak, SCCL-MT1 ve UPSI-SCC131 OSCC hücre hatlarında SLC2A3 (sırasıyla p<0,0001, p=0,0002) ve NLRP3 ekspresyonunu (sırasıyla p=0,0008; p=0,0006) anlamlı düzeyde azaltmış; bu da IL-1β salınımında (sırasıyla p=0,0190, p<0,0001), hücre proliferasyonunda (sırasıyla p=0,024; p<0,0001) ve migrasyonunda (sırasıyla p=0,0021; p=0,0004) azalmaya yol açmıştır. Bu bulgular, glukoz metabolizması ve enflamasyonu hedef alarak OSCC tedavisinde Emp'in potansiyel bir terapötik ajan olabileceğini düşündürmektedir.

Sonuç: Bulgularımız, emp'nin SLC2A3 ve NLRP3 ifadesini ve IL-1β salınımını inhibe ederek OSCC'de glikoz metabolizması ve enflamasyonu etkin bir şekilde modüle ettiğini, böylece kanser hücre proliferasyonu ve migrasyonunu azalttığını ve OSCC ilerlemesinin kontrolü için potansiyel bir terapötik ajan olabileceğini göstermektedir.

Anahtar Kelimeler: Oral squamöz hücreli karsinom, SLC2A3, NLRP3, empagliflozin

INTRODUCTION

Elevated glucose metabolism and inflammatory processes are two key factors that facilitate tumor growth, metastasis, and resistance to therapy in oral squamous cell carcinoma (OSCC)[1]. Cancer cells must adapt to rapid microenvironmental changes, such as hypoxia, nutrient scarcity, and acidic pH, particularly when oxidative phosphorylation alone cannot sustain energy production. To compensate, they reprogram glucose metabolism to ensure sufficient adenosine triphosphate (ATP) production and meet biosynthetic demands essential for cell proliferation[2,3]. This metabolic shift enables cancer cells to generate ATP and biosynthetic precursors via aerobic glycolysis, leading to significantly increased glucose uptake in highly proliferative tumor cells[3,4]. This adaptation supplies the substrates necessary for cancer cell proliferation and contributes to tumor progression, metastasis, and long-term survival[4,5].

Additionally, elevated glucose levels in cancer cells can induce reactive oxygen species, which consequently drive the formation of cytoplasmic multiprotein complexes such as the nucleotide-binding oligomerization domain, leucine-rich repeat-containing family pyrin domain containing 3 (NLRP3)[6]. This protein complex, known as the inflammasome, activates inflammatory responses by promoting cytokine secretion, such as pro-inflammatory interleukin-1beta (IL-1β), and releasing it into the extracellular space, thereby intensifying cellular inflammation[7].

Metabolic reprogramming of cancer cells is primarily regulated by hypoxia-inducible factor-1α , which controls the expression of metabolic enzymes, including glucose transporters (GLUTs)[8,9]. Given this metabolic reliance, targeting glucose metabolism offers a promising approach to enhance therapeutic efficacy[10,11]. Among the GLUT family, GLUT3 plays a crucial role in maintaining basal glucose transport, particularly in cells requiring high glucose affinity for survival[12]. In OSCC tumors, GLUT3 expression, encoded by the solute carrier family 2 member 3 *(SLC2A3)* gene, has been linked to poor prognosis and an aggressive cancer phenotype due to enhanced glucose metabolism[13,14]. Moreover, increased NLRP3 expression in OSCC has been associated with tumor progression and metastasis[15].

Given these findings, antidiabetic drugs have been investigated for their potential to modulate cancer cell metabolism and inflammatory responses. While studies have highlighted the anticancer effects of anti-diabetic drugs, including metformin, the role of sodium-glucose cotransporter-2 inhibitors, such as empagliflozin (emp), in cancer treatment remains poorly understood[16-19]. Although emp has been reported to suppress NLRP3 in obesity models[20], its impact on NLRP3 expression remains largely unknown in cancer. Thus, this study examined the influence of emp on SLC2A3, which encodes GLUT3, and NLRP3 in OSCC cell lines. Additionally, we explored the role of Emp-mediated SLC2A3-NLRP3 modulation in OSCC cell proliferation.

MATERIALS AND METHODS

Cells and Reagents

The human recurrent oral cavity squamous cell carcinoma cell line, SCCL-MT1, was generously provided by Bursa University, while the UPSI-SCC-131 OSCC cell line was a gift from (Bursa University). Cells were cultured in Dulbecco's Modified Eagle Medium-high glucose medium supplemented with fetal bovine serum (10%), antibiotic solution (penicillin & streptomycin, 50 U/mL), L-glutamine (2 mM), and sodium pyruvate (1 mM) and maintained under standard culture conditions (37 °C, 5% CO_2, humidified atmosphere). All experiments were conducted using passage 15 cells.

The cell culture media was from HyClone (Logan, UT, USA), while the supplements were provided by BIOCHROME (Berlin, Germany). Lipopolysaccharide (LPS) from Escherichia coli O111:B4 was purchased from Sigma (St. Louis, MO, USA), ATP was purchased from InvivoGen (San Diego, CA, USA) emp was supplied by Boehringer Ingelheim (Biberach, Germany).

Real-Time Polymerase Chain Reaction

Total RNA was isolated by the Quick-RNA Miniprep Kit (Zymo Research, Orange, CA, USA). The concentration and purity of RNA were evaluated by analyzing the A260/280 and A260/230 ratios with a Beckman Coulter DU 730 spectrophotometer (Brea, CA, USA). The ProtoScript M-MuLV First Strand cDNA Synthesis Kit (New England Biolabs, Ipswich, MA, USA) was

used for reverse transcription. The expression level of SLC2A3 was analyzed using the TaqMan™ Gene Expression Assay (Thermo Fisher Scientific, USA), while NLRP3 expression was analyzed with primers reported by Tezcan et al.[21]. Actin-beta (ACTB) served as the reference gene for normalization. For SLC2A3, the quantitative polymerase chain reaction (qPCR) reaction was performed according to the manufacturer's instructions using TaqMan® Gene Expression Assays (Applied Biosystems, Foster City, CA, USA). For NLRP3 and ACTB, qPCR was carried out in a 10 µL reaction volume containing 200 ng of cDNA, 10 µM of each primer, and Luna® Universal qPCR Master Mix (New England Biolabs, Ipswich, MA, USA). qPCR was performed on a StepOnePlus™ real time-PCR system (Thermo Fisher Scientific, Waltham, MA, USA) with the following cycling conditions: initial denaturation at 95 °C for 2 minutes, followed by 40 cycles of denaturation at 95 °C for 10 seconds and annealing/extension at 56 °C for 60 seconds. Data analysis, including Ct value determination and fold change calculations, was conducted using the StepOnePlus™ software and the $2^{-\Delta}$Ct method. For each comparison group, experiments were performed with three independent biological replicates, each analyzed in technical duplicate.

Enzyme-Linked Immunosorbent Assay (ELISA)

A sandwich ELISA approach (SunRed Bio, Shanghai, China) was employed, utilizing a biotin-labeled detection antibody specific to the target IL-1β. Samples and standards were incubated in wells pre-coated with a capture antibody for 1 hour at 37 °C, followed by three washes with phosphate-buffered saline (PBS). The biotin-labeled detection antibody was added and incubated for 30 minutes, followed by the streptavidin-horseradish peroxidase conjugate. After a 30-minute incubation with the substrate solution in the dark, color development was measured at 450 nm (reference: 650 nm) using a microplate reader. Each comparison group was analyzed in three biological repeats, each analyzed in technical duplicates. The results were calculated based on the standard curve generated from IL-1β reference standards.

Dynamic Cell Proliferation Analysis

SCCL-MT1 and UPSI-SCC-131 cells (5×10^4 per well) were plated into an E-Plate 16 (ACEA Biosciences, San Diego, CA, USA). Cell growth was perpetually monitored every 30 minutes over a 96-hour period using the xCELLigence real-time cell analysis (RTCA) dual purpose system (ACEA Biosciences, San Diego, CA, USA), which measures electrical impedance as an indicator of cell viability and proliferation. Background impedance was recorded before cell seeding to ensure accurate normalization. For each of the two different cell lines, the experiment was conducted with three independent biological replicates per comparison group. The resulting proliferation curves were generated and analyzed using the RTCA Software (ACEA Biosciences).

In-vitro Wound Healing Scratch Assay

A scratched wound-healing assay was conducted to visualize cell migration. Once the monolayer achieved confluence, a scratch was introduced using a 1000 µL pipette tip, creating a standardized wound area. Detached cells and debris were carefully removed by washing with PBS. Images of the wounded region were recorded after scratching and again 24 hours post-treatment using an inverted microscope (Nikon, Tokyo, Japan). Wound closure was quantified by measuring the wound area at both time points using ImageJ v1.54f software (National Institutes of Health, Bethesda, MD, USA). The experiment was performed in three biological replicates per comparison group, with each replicate analyzed in technical duplicate to ensure reproducibility.

Statistical Analysis

Protein-protein interactions were evaluated through the STRING v12.0, (https://string-db.org/) (accessed on February 24, 2025). Statistical analysis was achieved by GraphPad Prism (v14.4.0) (GraphPad Software, Boston, MA, USA). Data were confirmed to follow a normal distribution based on Q-Q plot analysis and Shapiro-Wilk tests, with $p>0.05$ indicating no significant deviation from normality. Homogeneity of variances was evaluated using Levene's test. Group differences were analyzed using one-way ANOVA, followed by Tukey's post-hoc test for pairwise comparisons to control for Type I error inflation due to multiple testing. Results are reported as mean ± standard error, with significance defined as $p<0.05$. All statistical tests were conducted with at least three independent biological replicates to ensure data robustness and reproducibility.

RESULTS

Empagliflozin Suppressed SLC2A3 Expression in OSCC Cells

To establish an inflammation-induced OSCC cell model, SCCL-MT1 and UPSI-SCC131 cells were exposed to 1 µg/mL LPS for 3 hours[22]. Additionally, cells were treated with 500 nM Emp, an SGLT2 inhibitor, for 24 hours to suppress glucose transport as described in our previous study[19]. Our results demonstrated that emp significantly downregulated SLC2A3 RNA expression in both SCCL-MT1 ($p<0.0001$), and UPSI-SCC131 cells ($p=0.0002$). Moreover, SLC2A3 expression was further reduced in LPS-induced cells when subsequently exposed to Emp, compared to cells treated with LPS alone ($p<0.0001$) (Figure 1 A, B).

Empagliflozin Reduced NLRP3 Expression and Suppresses IL-1β Release in LPS-Induced Inflammation

Genetic interaction analysis predicted that both SLC2A3 and NLRP3 interact with IL-1β (Figure 2A). Supporting these findings, emp reduced NLRP3 expression in SCCL-MT1 (p=0.0008) and UPSI-SCC131 cells (p=0.0006) (Figure 2B, C). Similarly, in LPS-treated SCCL-MT1 and UPSI-SCC131 cells, the emp treatment led to a decrease in NLRP3 expression than the cells triggered by LPS-only (p<0.0001). To evaluate the effect of emp on IL-1β release upon inflammasome activation, cells were treated with 1 µg/mL LPS for 3 hours, followed by 24-hour emp treatment and 25-minute exposure to 5 mM ATP (Figure 2D). IL-1β release was significantly increased in inflammasome-activated cells compared to those cells under LPS-induced inflammation conditions (p<0.0001). However, emp decreased IL-1β release both in LPS-treated cells (SCCL-MT1: p=0.0190; UPSI-SCC131) (p<0.0001) and in inflammasome-activated cells with LPS-ATP treatment (p<0.0001) (Figure 2E, F), suggesting that Emp, through the inhibition of SLC2A3, may also reduce NLRP3-mediated inflammatory cytokine production.

Empagliflozin Slowed Cell Proliferation and Migration in OSCC Cells

Emp treatment decreased the proliferation of both SCCL-MT1 and UPSI-SCC131 cell lines (Figure 3A, B). In contrast, LPS treatment increased cell proliferation compared to untreated controls. However, emp treatment also inhibited proliferation in LPS-pretreated cells relative to those treated with LPS alone (Figure 3A, B).

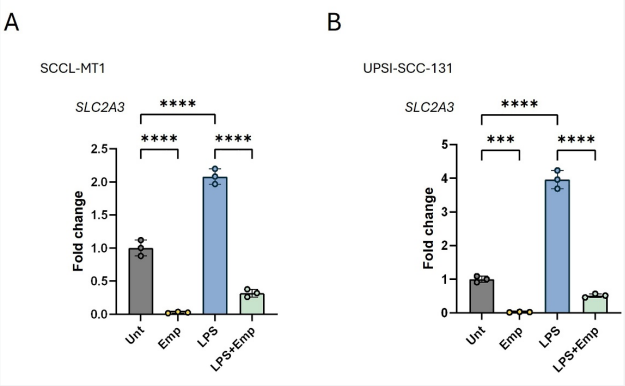


Figure 1. Effect of Emp on SLC2A3 expression in OSCC cells. (A) SCCL-MT1 cells, (B) UPSI-SCC131 cells. Comparison was performed between the unt & emp and LPS & emp+LPS. Only statistically significant differences are shown ***p<0.001, ****p<0.0001

Unt: Untreated, LPS: Lipopolysaccharide, emp: Empagliflozin, OSCC: Oral squamous cell carcinoma

Wound healing assays further supported these findings, showing that emp decreased cell migration in both SCCL-MT1 (p=0.024) and UPSI-SCC131 cells (p<0.0001) (Figure 3C, D). In contrast; LPS enhanced migration, accelerating cell movement (Figure 3C, D). Notably, emp also reduced wound healing in

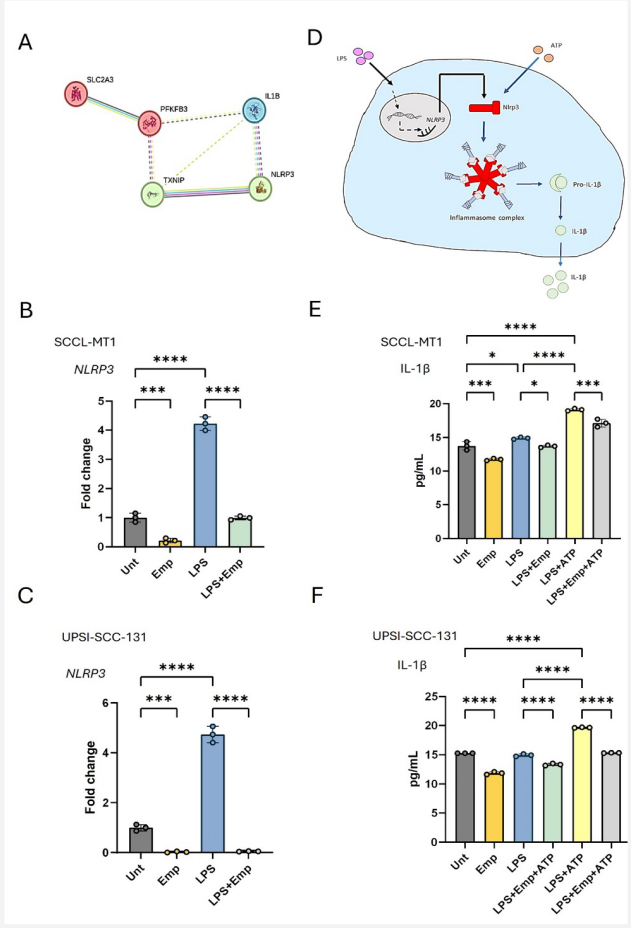


Figure 2. Effect of emp on the NLRP3 inflammasome in OSCC cells. (A) Interaction between SLC2A3, NLRP3, and IL-1β. (B) NLRP3 expression in SCCL-MT1 cells and (C) in UPSI-SCC131 cells. (D) Schematic representation of how LPS and ATP exposure influence NLRP3 transcription and inflammasome complex formation in vitro. (E) IL-1β release from SCCL-MT1 cells and (F) from UPSI-SCC131 cells. Emp, LPS, and LPS+ATP groups were compared to the Unt (control) group. To evaluate the effect of emp under inflammatory conditions, emp+LPS and emp+LPS+ATP groups were compared to LPS. For RT-qPCR analysis, only the comparison between emp+LPS and LPS was performed. Only statistically significant differences are shown.*p<0.05, ***p<0.001, ****p<0.0001

Unt: Untreated, LPS: Lipopolysaccharide, ATP: Adenosine triphosphate, emp: Empagliflozin, OSCC: Oral squamous cell carcinoma, IL-1β: Interleukin-1beta, RT-qPCR: Reverse transcription quantitative polymerase chain reaction

LPS-pretreated cells (SCCL-MT1: p=0.0021; UPSI-SCC131: p=0.0004) (Figure 3C, D). These results suggest that Emp-mediated suppression of SLC2A3 and NLRP3 reduces both the growth rate and invasive potential of OSCC cells, leading to a decrease in aggressive cancer cell behavior.

DISCUSSION

This study investigates the role of emp on NLRP3 inflammasome activation in SCCL-MT1 and UPSI-SCC131, OSCC cell lines. SCCL-MT1, derived from buccal mucosa, is an invasive and recurrent OSCC model with strong immunomodulatory activity[23]. Previous studies have demonstrated that SCCL-MT1 promotes immunosuppression within the tumor microenvironment by stimulating myeloid-derived suppressor cells via IL-1β secretion, leading to effector T-cell suppression and an increase in regulatory T-cells[23,24]. Similarly, UPSI-SCC131, originating from a patient with a history of tobacco and alcohol use, is resistant to radiotherapy[25,26]. This resistance has been associated with elevated IL-1β secretion[27]. Due to their clinically relevant characteristics such as invasiveness, immunomodulatory activity, and therapy resistance, the effects of emp on OSCC were investigated using the SCCL-MT1 and UPSI-SCC131 cell lines in this study.

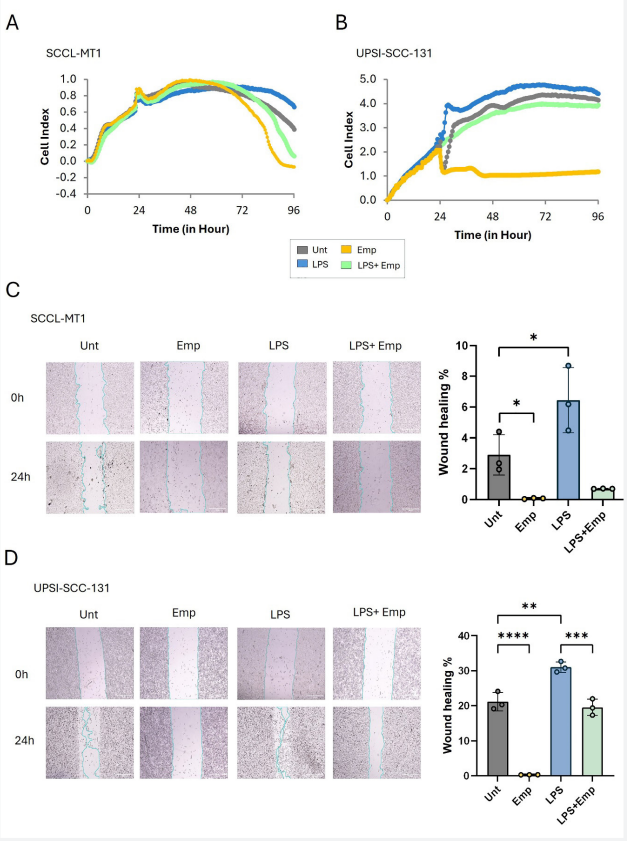


Figure 3. Effect of emp on OSCC cell motility. (A) Real-time proliferation of SCCL-MT1 cells and (B) UPSI-SCC131 cells. (C) Healing of scratched wound in SCCL-MT1 cells and (D) UPSI-SCC131 cells. Comparisons were performed between the Unt & emp and the Unt & LPS groups. In addition, to assess emp's effect under inflammatory conditions, emp+LPS was compared to LPS. Only statistically significant differences are shown. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

Unt: Untreated, LPS: Lipopolysaccharide, Emp: Empagliflozin, OSCC: Oral squamous cell carcinoma

Empagliflozin-Mediated Mechanistic Link Between SLC2A3 and NLRP3

IL-1β production and secretion *in vitro* require a two-step activation of the NLRP3 inflammasome. In the priming phase, pathogen-associated stimuli such as LPS activate Toll-like receptors, leading to NF-κB signaling and enhanced NLRP3 transcription. In the activation phase, ATP stimulates the P2X7 receptor, inducing potassium efflux and promoting NLRP3 inflammasome assembly. This process culminates in pro-caspase-1 activation and IL-1β secretion[28-30]. Therefore, we evaluated the effect of emp on NLRP3 activation by stimulating SCCL-MT1 and UPSI-SCC131 cells with LPS, followed by ATP exposure, and analyzing IL-1β secretion.

Building on these mechanistic insights, our findings demonstrate that treatment with Emp, apharmacological inhibitor of glucose transport[31], downregulates SLC2A3 (GLUT3) mRNA expression, indicating reduced glucose uptake in OSCC cells. Emp also suppresses NLRP3 expression and inhibits inflammasome activation, leading to decreased IL-1β secretion. These results suggest a functional link between SLC2A3 and NLRP3 in the regulation of glucose metabolism and inflammation in cancer cells.

A recent study has associated elevated SLC2A3 expression with the activation of epithelial-mesenchymal transition and NF-κB signaling pathways, contributing to poor prognosis in patients with head and neck squamous cell carcinoma[32]. NF-κB plays a pivotal role in initiating inflammasome activation by upregulating the transcription of NLRP3 and pro-IL-1β during the priming phase[33]. In our study, LPS was used to model pathogen-associated molecular patterns. Importantly, even in the absence of LPS stimulation, emp suppressed SLC2A3 and NLRP3 expression, along with IL-1β secretion. These findings suggest that SLC2A3-mediated glucose uptake alone may contribute to inflammasome activation. Indeed, previous studies have shown that elevated glucose transport triggers inflammatory signaling in macrophages[34], the cell type where NLRP3 activation is most pronounced[35].

To further explore this connection, we conducted a gene–gene interaction analysis, which suggested a potential mechanistic link between SLC2A3 and NLRP3 through the glycolysis-regulating enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3 (PFKFB3). Cancer cells overexpress GLUTs such as SLC2A3 to increase glucose influx, which elevates intracellular fructose-2,6-bisphosphate (F2,6BP) levels. This metabolite allosterically activates phosphofructokinase-1, the rate-limiting enzyme in glycolysis. PFKFB3 amplifies this effect by promoting F2,6BP synthesis, accelerating glycolytic flux and enhancing lactate production. Lactate is exported from the cell, bypassing mitochondrial oxidative metabolism and promoting reliance on aerobic glycolysis. This shift supports rapid proliferation and fosters a hypoxic tumor microenvironment[36]. Emerging evidence indicates that PFKFB3 not only promotes glycolytic flux but also contributes to NLRP3 inflammasome activation. Suppression of PFKFB3 has been shown to reduce glycolysis[36], and in macrophages, NLRP3 can regulate glycolysis through PFKFB3[37]. Moreover, PFKFB3-induced hypoxia has been reported to trigger NLRP3 activation, while downregulation of PFKFB3 reduced NLRP3 downstream targets, including caspase-1 and IL-1β, in an atherosclerosis model[38]. Notably, LPS has been shown to enhance ROS generation induced by high glucose–associated hypoxia, thereby exacerbating NLRP3 activation[39].

Taken together, these findings support the hypothesis that the SLC2A3-PFKFB3-NLRP3 axis may represent a potential mechanistic link between altered glucose metabolism and inflammation in OSCC (Figure 4). Although based on in silico predictions, this pathway suggests that by suppressing SLC2A3, emp may attenuate NLRP3 inflammasome activation, potentially through modulation of PFKFB3, thereby exerting both metabolic and anti-inflammatory effects in oral cancer.

Therapeutic Implications of Emp on OSCC

The role of the NLRP3 inflammasome in cancer remains controversial. While some studies associate NLRP3 overexpression with tumor aggressiveness and poor prognosis, others report minimal effects upon its inhibition[40]. Our previous work using Glibenclamide and VX765, two NLRP3 inhibitors, showed only modest reductions in OSCC cell proliferation[22,41]. In contrast, emp exhibited a more pronounced anti-cancer effect, significantly decreasing OSCC cell proliferation and impairing wound healing, suggesting greater therapeutic potential through dual metabolic and inflammatory modulation.

Notably, SLC2A3 is known to promote OSCC progression through the TGF-β signaling[42]. In addition, TGF-β signaling has been implicated in chronic inflammation through an IL-1β-driven autocrine mechanism[43]. Although not tested in this

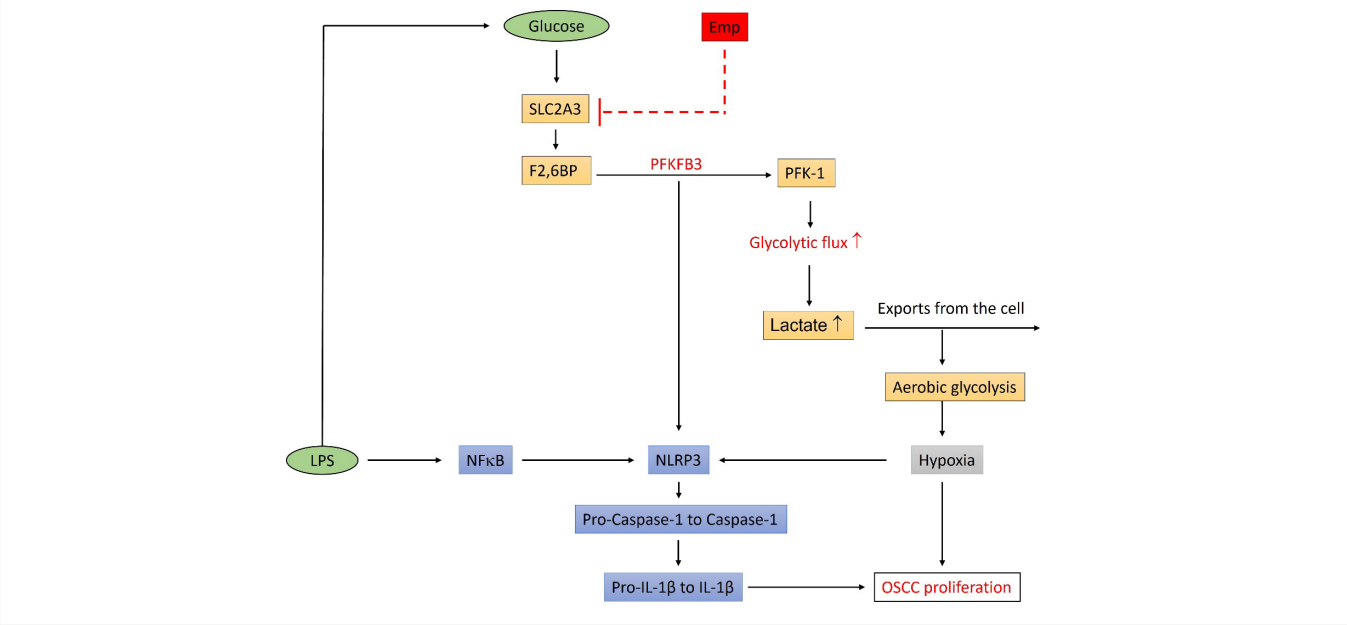


Figure 4. Schematic representation of the hypothetical SLC2A3-PFKFB3-NLRP3 Axis and emp's action in OSCC. Emp suppresses glucose uptake by downregulating SLC2A3, which may lead to decreased PFKFB3 activity and subsequently reduce NLRP3 inflammasome activation. This proposed pathway links altered glucose metabolism to inflammation in OSCC

LPS: Lipopolysaccharide, emp: Empagliflozin, OSCC: Oral squamous cell carcinoma, LPS: Lipopolysaccharide, SLC2A3: Solute carrier family 2 member 3

study, emp's reported inhibition of TGF-β signaling[44] suggests a potential additional mechanism for its effects in OSCC, which warrants further investigation. The convergence of SLC2A3-mediated glucose transport and TGF-β–IL-1β signaling pathways may represent a novel therapeutic target axis in OSCC.

Collectively, these data support the hypothesis that glucose transport contributes to OSCC aggressiveness, and its inhibition by emp may help suppress this phenotype. Under LPS-induced inflammatory conditions, proliferation and migration rates increased but were effectively reduced by Emp. These findings highlight emp's potential not only as an antidiabetic agent but also as a repurposed therapeutic candidate targeting metabolic-inflammation crosstalk in OSCC.

Comparison of Emp and Conventional NLRP3 Inhibitors

Mechanistically, Glibenclamide prevents inflammasome formation by blocking NLRP3 monomer assembly, while VX765 inhibits caspase-1 activation, which is required for IL-1β maturation[22,41,45]. Although emp does not directly target the inflammasome complex, our results suggest it reduces NLRP3 expression by limiting glucose availability, a key upstream stimulus. Thus, emp may act earlier in the NLRP3 pathway and exert additional effects via SLC2A3-associated signaling.

Overall, this study reveals that metabolic regulation of glucose uptake plays a critical role in the progression of inflammation associated with OSCC cells. Notably, the indirect interaction between SLC2A3 and NLRP3 may contribute to shaping this inflammatory response. In this context, the suppression of proliferation and migration observed after treatment with glucose transport inhibitors such as emp likely results from the downregulation of NLRP3 expression. This reduction may limit NLRP3-associated signaling pathways involved in inflammation, thereby decreasing the secretion of pro-tumorigenic cytokines.

Indeed, the effects of emp under both basal and LPS-induced inflammatory conditions highlight the potential of glucose uptake inhibition to reduce pro-inflammatory signaling and disrupt autocrine cytokine loops in cancer cells, thereby mitigating the aggressive phenotype of OSCC.

Study Limitations

This study is an *in vitro* investigation conducted using two commercial cell lines. In OSCC inflammation, the inflammatory activities of non-cancerous cells within the tumor microenvironment and their interactions with cancer cells play a role. However, the experimental setup in this study consists of a homogeneous system containing only cancer cells, thereby excluding the influence of the tumor microenvironment. This represents a limitation of our study.

Furthermore, the study employed a single concentration of Emp, selected based on preliminary cytotoxicity and efficacy data. Although effective, the use of a single dose restricts dose–response interpretation and does not account for possible concentration-dependent effects or therapeutic windows. Additional experiments using multiple concentrations and time points would enhance the pharmacodynamic understanding of emp in OSCC.

Another important limitation is the lack of *in vivo* validation. Without animal model data, it is difficult to predict the systemic effects of Emp, its pharmacokinetics, or its potential interactions within a living organism.

Moreover, as this is not a patient-based study, the focus is solely on the effects of emp on cancer cell biology, disregarding potential variations arising from patient heterogeneity, etiological differences, and epigenetic diversity. These factors may limit the clinical generalizability and translational applicability of our findings.

CONCLUSION

In conclusion, our findings demonstrate that Emp, an SGLT2 inhibitor, suppresses SLC2A3 and NLRP3 expression, thereby reducing OSCC cell proliferation and migration. Its dual action, disrupting glucose metabolism and mitigating chronic inflammation via NLRP3 inhibition, may attenuate the aggressive phenotype of OSCC. Taken together, these insights highlight emp's potential as a therapeutic agent for OSCC, though further studies are needed to validate its clinical efficacy.

Ethics

Ethics Committee Approval: The study does not involve human or animal subjects, ethical approval is not required. The relevant statement has also been included in the manuscript.

Informed Consent: This study is an in vitro study conducted using two commercially available cell lines. No biological material derived from patients or experimental animals was used, and no clinical or personal data from any patient were included.

Footnotes

Authorship Contributions

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

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

REFERENCES

1. Chen Y, Han Z, Zhang L, Gao C, Wei J, Yang X, et al. TIMELESS promotes reprogramming of glucose metabolism in oral squamous cell carcinoma. J Transl Med. 2024;22:21.
2. Oizel K, Chauvin C, Oliver L, Gratas C, Geraldo F, Jarry U, et al. Efficient mitochondrial glutamine targeting prevails over glioblastoma metabolic plasticity. Clin Cancer Res. 2017;23:6292-304.
3. Porporato PE, Dhup S, Dadhich RK, Copetti T, Sonveaux P. Anticancer targets in the glycolytic metabolism of tumors: a comprehensive review. Front Pharmacol. 2011;2:49.
4. Liberti MV, Locasale JW. The Warburg Effect: How does it benefit cancer cells? Trends Biochem Sci. 2016;41:211-8.
5. Levine AJ, Puzio-Kuter AM. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science. 2010;330:1340-4.
6. Ortega MA, De Leon-Oliva D, García-Montero C, Fraile-Martinez O, Boaru DL, de Castro AV, et al. Reframing the link between metabolism and NLRP3 inflammasome: therapeutic opportunities. Front Immunol. 2023;14:1232629.
7. Guo H, Callaway JB, Ting JP. Inflammasomes: mechanism of action, role in disease, and therapeutics. Nat Med. 2015;21:677-87.
8. Gupta S, Roy A, Dwarakanath BS. Metabolic cooperation and competition in the tumor microenvironment: implications for therapy. Front Oncol. 2017;7:68.
9. Semenza GL. Targeting HIF-1 for cancer therapy. Nat Rev Cancer. 2003;3:721-32.
10. Vander Heiden MG, DeBerardinis RJ. Understanding the Intersections between metabolism and Cancer Biology. Cell. 2017;168:657-69.
11. Martinez-Outschoorn UE, Peiris-Pagés M, Pestell RG, Sotgia F, Lisanti MP. Cancer metabolism: a therapeutic perspective. Nat Rev Clin Oncol. 2017;14:11-31.
12. Yager JY. Glucose and perinatal brain injury questions and controversies. Elsevier. 2019:141-61.
13. Ayala FRR, Rocha RM, Carvalho KC, Carvalho AL, Da Cunha IW, Lourenço SV, et al. Glut1 and Glut3 as Potential Prognostic Markers for Oral Squamous Cell Carcinoma. Molecules. 2010;15:2374-87.
14. Kumari A, Jha A, Tiwari A, Nath N, Kumar A, Nagini S, et al. Role and regulation of GLUT1/3 during oral cancer progression and therapy resistance. Arch Oral Biol. 2023;150:105688.
15. Wang H, Luo Q, Feng X, Zhang R, Li J, Chen F. NLRP3 promotes tumor growth and metastasis in human oral squamous cell carcinoma. BMC Cancer. 2018;18:500.
16. Dąbrowski M. Diabetes, antidiabetic medications and cancer risk in type 2 diabetes: focus on SGLT-2 inhibitors. Int J Mol Sci. 2021;22:1680.
17. Dutka M, Bobiński R, Francuz T, Garczorz W, Zimmer K, Ilczak T, et al. SGLT-2 Inhibitors in cancer treatment-mechanisms of action and emerging new perspectives. Cancers (Basel). 2022;14:5811.
18. Wu W, Wang Y, Xie J, Fan S. Empagliflozin: a potential anticancer drug. Discov Oncol. 2023;14:127.
19. Kocaeli AA, Tekin C, Ercelik M, Tezcan G, Aksoy SA, Kocaeli H, et al. Concurrent presence of diabetes affects the GLUT3 programming of glucose metabolism in glioblastoma. Eur Rev Med Pharmacol Sci. 2023;27:8110-8.
20. Benetti E, Mastrocola R, Vitarelli G, Cutrin JC, Nigro D, Chiazza F, et al. Empagliflozin protects against diet-induced NLRP-3 inflammasome activation and lipid accumulation. J Pharmacol Exp Ther. 2016;359:45-53.
21. Tezcan G, Alsaadi M, Hamza S, Garanina EE, Martynova EV, Ziganshina GR, et al. Azithromycin and ceftriaxone differentially activate NLRP3 in LPS primed cancer cells. Int J Mol Sci. 2022;23:9484.
22. Hamza S, Garanina EE, Shkair L, Alsaadi M, Khaiboullina SF, Tezcan G. Implications of NLRP3 Suppression using glibenclamide and miR-223 against colorectal cancer. Pharmaceuticals (Basel). 2024;17:299.
23. Russell SM, Lechner MG, Gong L, Megiel C, Liebertz DJ, Masood R, et al. USC-HN2, a new model cell line for recurrent oral cavity squamous cell carcinoma with immunosuppressive characteristics. Oral Oncol. 2011;47:810-7.
24. Lechner MG, Liebertz DJ, Epstein AL. Characterization of cytokine-induced myeloid-derived suppressor cells from normal human peripheral blood mononuclear cells. J Immunol. 2010;185:2273-84.
25. White JS, Weissfeld JL, Ragin CC, Rossie KM, Martin CL, Shuster M, et al. The influence of clinical and demographic risk factors on the establishment of head and neck squamous cell carcinoma cell lines. Oral Oncol. 2007;43:701-12.
26. Sürer Şİ, Elçitepe TB, Akçay D, Daşkın E, Çalıbaşı Kocal G, Arıcan Alıcıkuş Z, et al. A promising, novel radiosensitizer nanodrug complex for oral cavity cancer: cetuximab and cisplatin-conjugated gold nanoparticles. Balkan Med J. 2021;38:278-86.
27. Tiwari DK, Hannen R, Unger K, Kohl S, Heß J, Lauber K, et al. IL1 pathway in HPV-negative HNSCC cells is an indicator of radioresistance after photon and carbon ion irradiation without functional involvement. Front Oncol. 2022;12:878675.
28. Tengesdal IW, Dinarello CA, Marchetti C. NLRP3 and cancer: pathogenesis and therapeutic opportunities. Pharmacol Ther. 2023;251:108545.
29. Gurung P, Li B, Subbarao Malireddi RK, Lamkanfi M, Geiger TL, Kanneganti TD. Chronic TLR stimulation controls NLRP3 inflammasome activation through IL-10 mediated regulation of NLRP3 expression and caspase-8 activation. Sci Rep. 2015;5:14488.
30. Rotondo JC, Mazziotta C, Lanzillotti C, Stefani C, Badiale G, Campione G, et al. The role of purinergic P2X7 receptor in inflammation and cancer: novel molecular insights and clinical applications. Cancers (Basel). 2022;14:1116.
31. Zügner E, Yang HC, Kotzbeck P, Boulgaropoulos B, Sourij H, Hagvall S, et al. Differential In vitro effects of SGLT2 inhibitors on mitochondrial oxidative phosphorylation, glucose uptake and cell metabolism. Int J Mol Sci. 2022;23:7966.
32. Chai F, Zhang J, Fu T, Jiang P, Huang Y, Wang L, et al. Identification of SLC2A3 as a prognostic indicator correlated with the NF-κB/EMT axis and immune response in head and neck squamous cell carcinoma. Channels (Austin). 2023;17:2208928.
33. Liu T, Zhang L, Joo D, Sun SC. NF-κB signaling in inflammation. Signal Transduct Target Ther. 2017;2:17023.
34. Johnson AR, Milner JJ, Makowski L. The inflammation highway: metabolism accelerates inflammatory traffic in obesity. Immunol Rev. 2012;249:218-38.
35. Lebreton F, Berishvili E, Parnaud G, Rouget C, Bosco D, Berney T, et al. NLRP3 inflammasome is expressed and regulated in human islets. Cell Death Dis. 2018;9:726.
36. Lu L, Chen Y, Zhu Y. The molecular basis of targeting PFKFB3 as a therapeutic strategy against cancer. Oncotarget. 2017;8:62793-802.
37. Finucane OM, Sugrue J, Rubio-Araiz A, Guillot-Sestier MV, Lynch MA. The NLRP3 inflammasome modulates glycolysis by increasing PFKFB3 in an IL-1β-dependent manner in macrophages. Sci Rep. 2019;9:4034.
38. Wang X, Liu X, Wu W, Liao L, Zhou M, Wang X, et al. Hypoxia activates macrophage-NLRP3 inflammasome promoting atherosclerosis via PFKFB3-driven glycolysis. FASEB J. 2024;38:e23854.
39. Qiu Z, He Y, Ming H, Lei S, Leng Y, Xia ZY. Lipopolysaccharide (LPS) aggravates high glucose- and hypoxia/reoxygenation-induced injury through activating ROS-Dependent NLRP3 inflammasome-mediated pyroptosis in H9C2 cardiomyocytes. J Diabetes Res. 2019;17:8151836.
40. Moossavi M, Parsamanesh N, Bahrami A, Atkin SL, Sahebkar A. Role of the NLRP3 inflammasome in cancer. Mol Cancer. 2018;17:158.
41. Tezcan G, Garanina EE, Alsaadi M, Gilazieva ZE, Martinova EV, Markelova MI, et al. Therapeutic potential of pharmacological targeting NLRP3 inflammasome complex in cancer. Front Immunol. 2021;11:607881.

42. Jiang W, Xu S, Li P. SLC2A3 promotes tumor progression through lactic acid-promoted TGF-β signaling pathway in oral squamous cell carcinoma. PLoS One. 2024;19:e0301724.

43. Kang H, Seo E, Oh YS, Jun HS. TGF-β activates NLRP3 inflammasome by an autocrine production of TGF-β in LX-2 human hepatic stellate cells. Mol Cell Biochem. 2022;477:1329-38.

44. Shentu Y, Li Y, Xie S, Jiang H, Sun S, Lin R, et al. Empagliflozin, a sodium glucose cotransporter-2 inhibitor, ameliorates peritoneal fibrosis via suppressing TGF-β/Smad signaling. Int Immunopharmacol. 2021;93:107374.

45. Zahid A, Li B, Kombe AJK, Jin T, Tao J. Pharmacological Inhibitors of the NLRP3 inflammasome. Front Immunol. 2019;10:2538.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.37640
Nam Kem Med J 2025;13(3):311-321

Cosmetic Use and Attitudes Toward Female Genital Cosmetic Procedures Among Women: A Single Center Cross-Sectional Study

Kadınlarda Kozmetik Kullanımı ve Kadın Genital Estetik İşlemlerine Yönelik Tutumlar: Tek Merkezli Kesitsel Bir Çalışma

Özge ZORLU[1], İlke ÖZER ASLAN[2], Mustafa Törehan ASLAN[3]

[1]*Tekirdağ Namık Kemal University Faculty of Medicine, Department of Dermatology and Venereology, Tekirdağ, Türkiye*
[2]*Tekirdağ Namık Kemal University Faculty of Medicine, Department of Obstetrics and Gynecology, Tekirdağ, Türkiye*
[3]*Tekirdağ Namık Kemal University Faculty of Medicine, Department of Pediatrics, Division of Neonatology, Tekirdağ, Türkiye*

ABSTRACT

Aim: Female genital cosmetic procedures (GCP) are becoming a trending topic as the demand for GCP and cosmetic usage has increased. This study aimed to assess women's changes/intended changes in cosmetic use parameters and their attitudes toward GCP following/in the event of pregnancy or delivery.

Materials and Methods: This cross-sectional and questionnaire-based study included 98 pregnant women (PW), 82 non-PW (NPW) of reproductive age, and 96 puerperal women (PuW). Knowledge about GCP, willingness to have any GCP, and cosmetic use parameters were established, and the Female Genital Self-image scale (FGSIS) was applied.

Results: Women who wanted GCP had lower FGSIS scores (p=0.015). The proportion of women with a university or higher education level was significantly higher among women who considered having GCP in the NPW (76.5%, p=0.021) and PuW (42.5%, p=0.015) groups. More NPW stated they would change their criteria and information sources related to cosmetics in the event of pregnancy/delivery compared to other groups (p<0.001). A significant proportion of NPW thought to stop using blush, powder, and concealer in the case of pregnancy, and lipstick, mascara, makeup remover, and foundation in the case of pregnancy/delivery (p<0.05).

Conclusion: Healthcare professionals should be aware of and address the demand for GCP and cosmetic usage among women, including pregnant and PuW.

Keywords: Cosmetics, genital cosmetic, female genital self-image, pregnancy, puerperium

ÖZ

Amaç: Kadın genital kozmetik uygulamalar (GKU), bu tür işlemlere ve kozmetik ürün kullanımına olan talebin artmasıyla birlikte giderek daha popüler bir konu haline gelmektedir. Bu çalışma, kadınların gebelik veya doğum sonrasında/halinde kozmetik kullanım alışkanlıklarındaki değişiklikleri ya da planladıkları değişiklikleri ve GKU'ya yönelik tutumlarını değerlendirmeyi amaçlamıştır.

Gereç ve Yöntem: Bu kesitsel ve anket temelli çalışmaya 98 gebe kadın (GK), üreme çağında olan 82 gebe olmayan kadın (GOK) ve 96 lohusa kadın (LK) dahil edilmiştir. Katılımcıların GKU hakkındaki bilgi düzeyleri, herhangi bir GKU yaptırma isteklilikleri ve kozmetik kullanımına ilişkin parametreleri belirlenmiş; ayrıca Kadın Genital Kendilik Algısı ölçeği (KGKAÖ) uygulanmıştır.

Bulgular: GKU yaptırmak isteyen kadınların KGKAÖ puanları daha düşüktü (p=0,015). GKU yaptırmayı düşünen kadınlar arasında üniversite veya üzeri eğitim düzeyine sahip olanların oranı, GOK (%76,5, p=0,021) ve LK (%42,5, p=0,015) gruplarında anlamlı düzeyde yüksekti. GOK grubundaki daha fazla kadın, gebelik/doğum durumunda kozmetik ürünlere ilişkin kriterlerini ve bilgi kaynaklarını değiştireceğini belirtti (p<0,001). Anlamlı

Address for Correspondence: Özge ZORLU MD, Tekirdağ Namık Kemal University Faculty of Medicine, Department of Dermatology and Venereology, Tekirdağ, Türkiye
E-mail: zorluzg@gmail.com **ORCID ID:** orcid.org/0000-0001-5555-130X
Received: 30.04.2025 **Accepted:** 30.06.2025 **Publication Date:** 07.10.2025

Cite this article as: Zorlu Ö, Aslan İÖ, Aslan MT. Cosmetic use and attitudes toward female genital cosmetic procedures among women: a single center cross-sectional study. Nam Kem Med J. 2025;13(3):311-321

bir oranda GOK, gebelik halinde allık, pudra ve kapatıcı kullanımını bırakmayı; gebelik/doğum durumunda ise ruj, maskara, makyaj temizleyici ve fondöten kullanımını bırakmayı düşündüğünü ifade etti (p<0,05).

Sonuç: Sağlık çalışanları, gebeler ve LK da dahil olmak üzere kadınlar arasında GKU'ya ve kozmetik ürün kullanımına yönelik talebin farkında olmalı ve bu taleplere yönelik gerekli değerlendirme ve yönlendirmeleri yapmalıdır.

Anahtar Kelimeler: Kozmetik, genital kozmetik, kadın genital kendilik algısı, gebelik, lohusalık

INTRODUCTION

Female genital cosmetic procedures (GCP) are defined as legal procedures performed to improve the appearance of the genital area in the absence of medical indications, such as labiaplasty, vaginoplasty, and hymenoplasty[1]. An increasing number of women are pursuing surgical modification of the genitalia for cosmetic reasons. According to the Aesthetic Plastic Surgery National Databank statistics, vaginoplasty, labiaplasty, and clitoral hood reduction surgeries have increased by 374%, 36%, and 128%, respectively, from 2020 to 2021[2].

Cosmetics, including personal care products (PCP) or make-up products (MUP), that contain various chemical substances are widely used in daily life[3,4]. The association between cosmetic use and increased personal exposure to substances such as benzophenone, paraben, and phthalate has been shown in various studies[5,6]. In addition, individual exposure to some of these substances, especially phthalates and phenols, is reportedly related to adverse pregnancy outcomes[7,8].

Pregnant women (PW) may experience skin changes that are not attractive, such as melasma, striae, or acne. Likewise, some problems in the genital region related to delivery methods may exist in the puerperium. General cosmetic use habits can change, and a need for modification of the genitalia may arise for cosmetic reasons. As such, women are vulnerable to the potential risks of chemicals in cosmetics.

Understanding women's attitudes towards GCP and their demand for these procedures would contribute to the awareness and knowledge of healthcare professionals. This study aimed to assess the changes in cosmetics use parameters (the criteria of choice of cosmetic products, information sources considered while choosing cosmetics, regularly used cosmetic products), GCP knowledge, and the attitudes toward GCP following pregnancy/delivery, or intended changes in the event of pregnancy/delivery. Secondly, we assessed the relationship between the Female Genital Self-image scale (FGSIS) and attitudes toward GCP.

MATERIALS AND METHODS

Study Population

This cross-sectional questionnaire-based study was conducted between April 2023 and October 2023 in the obstetrics and gynecology (OB/GYN) outpatient clinic of our tertiary referral hospital.

The study population consisted of three groups: PW, non-pregnant women of reproductive age (NPW), and puerperal women (PuW). Women 18-45 years of age from the same geographical region who agreed to participate in the study and gave informed consent were included.

Group exclusion criteria were:

- Depression in pregnancy, ectopic or molar pregnancy, having a fetus with an intrauterine anomaly for PW;
- Giving birth ≥6 weeks ago, postpartum depression, having a baby with a congenital anomaly for PuW;
- History of previous pregnancies, current pregnancy, being in the puerperal or perimenopausal period for NPW.

The study was approved by the Non-Interventional and Clinical Research Ethics Committee of Tekirdağ Namık Kemal University (decision no: 2023.28.02.06, date: 28.02.2023) and was carried out following the principles of the 1975 Declaration of Helsinki as revised in 2000. All participants signed informed consent before participating in the study.

Data Collection

All participants were asked to complete a self-administered questionnaire comprising socio-demographic data (age, education status, socioeconomic status, residence, occupation) and obstetrical data (gestational week, delivery week, birth weight, previous pregnancy number, first maternal age, preterm birth, multiple pregnancies, delivery method, abortus) for PW and PuW groups.

The criteria for choice of cosmetics (ingredients, price, brand, odor, advice of doctors or friends, appearance, habit, user comments, net contents, package, satisfaction), information sources considered while choosing cosmetics (commercials, social media phenomena, dermatologists' or friends' advice, internet, estheticians, TV programs, user comments on blogs), purchasing frequency (once a month, quarterly, twice a year, once a year, other), choices of make-up ingredients (fat-free, alcohol-free, paraben-free, unpolished, non-comedogenic, natural), and regularly used cosmetic products were established by an OB/GYN resident directly asking each woman.

Use habits were established for 36 cosmetics: 14 MUP (blush, lipliner, lipstick, mascara, eye shadow, eyeliner, eye pencil, eyebrow shadow, powder, concealer, foundation, make-up remover, nail polish, and nail polish remover) and 22 PCP [eight for general care (sunscreen, moisturizing cream, soap, shower gel, body peeling, perfume, deodorant, and collagen pills), seven for face (daily face cream, face night cream, facial cleanser, anti-aging serum, facial mask, facial tonic, and facial peeling), four for the genital region (cleanser, care lotions, deodorant, and bleaching cream), and three for hair (shampoo, dye, and mask)]. The use of cosmetics was evaluated based on general and regular use, irrespective of the frequency.

The thought of sufficient knowledge about GCP [labiaplasty, vaginoplasty (tightening, rejuvenation)], perineoplasty, vulvar or perianal bleaching, laser, cesarean section (C/S) scar revision, hymenoplasty, clitoral hood reduction, G-spot amplification) and willingness to have any GCP were established by an OB/GYN resident directly asking.

The changes/intended changes in cosmetic use parameters and willingness to have GCP following pregnancy or delivery were assessed by asking the following question: "Since the beginning of pregnancy/delivery (or in the event of pregnancy/delivery), have you changed (or will you change) the use of the criteria/information sources related to GCP?" (Yes/No).

Finally, the FGSIS, a Likert-type four-point (1: Strongly disagree to 4: Strongly agree) self-reported questionnaire consisting of seven items, was applied[9]. The total possible score ranged between 7 and 28, with a higher score indicating better genital self-image.

Statistical Analysis

The statistical power for our sample size was 0.872, with an effect size of 0.30 and an alpha-type error of 0.05, which were calculated with the help of the G*Power 3.1.9.4 program using the χ^2 -test family.

The Shapiro-Wilk test was used to assess the normality of variables. According to the normality results, continuous variables were presented as medians [interquartile range (IQR)], the Mann-Whitney U test was used for comparisons between two groups, and the Kruskal-Wallis test was used for comparisons between the three groups (e.g., FGSIS). Categorical variables were reported as n (%). The Pearson χ^2 or Fisher's exact test was used to compare categorical variables (e.g., socio-demographic data, attitudes towards GCP/cosmetics). Pairwise comparisons were performed using the post-hoc test with the Bonferroni adjustment if the result of the chi-square test was significant. A cross-over analysis using McNemar's χ^2 test was performed to compare the proportions of criteria, information sources, cosmetic use (MUP or PCP), make-up ingredients, and GCP before and after pregnancy or before and after delivery.

SPSS v.25 (IBM Corp, Armonk, NY, US) software was used for statistical analysis, and a p-value<0.05 was considered statistically significant.

RESULTS

Socio-demographic and Obstetric Data

The socio-demographic and obstetric data of the participants are presented in Table 1. Median (IQR) gestational week for PW was 35 (33-37). Of PW, 6 (7.1%) were in the second trimester, and 79 (92.9%) were in the third trimester. Multiple pregnancies occurred in 8 (8.2%) of 98 PW.

Median (IQR) delivery week for PuW was 38 (37-39). Preterm birth occurred in 14 (22.6%) women. The median (IQR) birth weight was 3200g (2760-3500). Low birth weight was seen in 8 (14%) deliveries, and no babies had high birth weights. Delivery methods were vaginal in 22 (22.9%) and C/S in 74 (77.1%) PuW, and 4 (4.2%) of 96 had multiple pregnancies.

GCP and FGSIS Scores

The GCP and FGSIS scores of the study groups are presented in Table 2. There were no significant differences in considering GCP between the study groups, except for labiaplasty, of which the percentage was highest in the NPW group and lowest in the PW group (p=0.015) (Figure 1).

FGSIS scores were significantly lower in the participants who would like to have at least one GCP (median, IQR=21, 18-25) compared to those who did not want any (median, IQR=23, 20-27) (p=0.015). There was no statistically significant relationship between considering having GCP and socio-demographic or obstetric features, except education level. In PW, NPW, and PuW groups, 44.8%, 76.5% (Z-score=3.9, p=0.004) and 35% of women, respectively, who thought about having GCP had a university or higher education level (p=0.008).

Changing Attitudes Concerning GCP

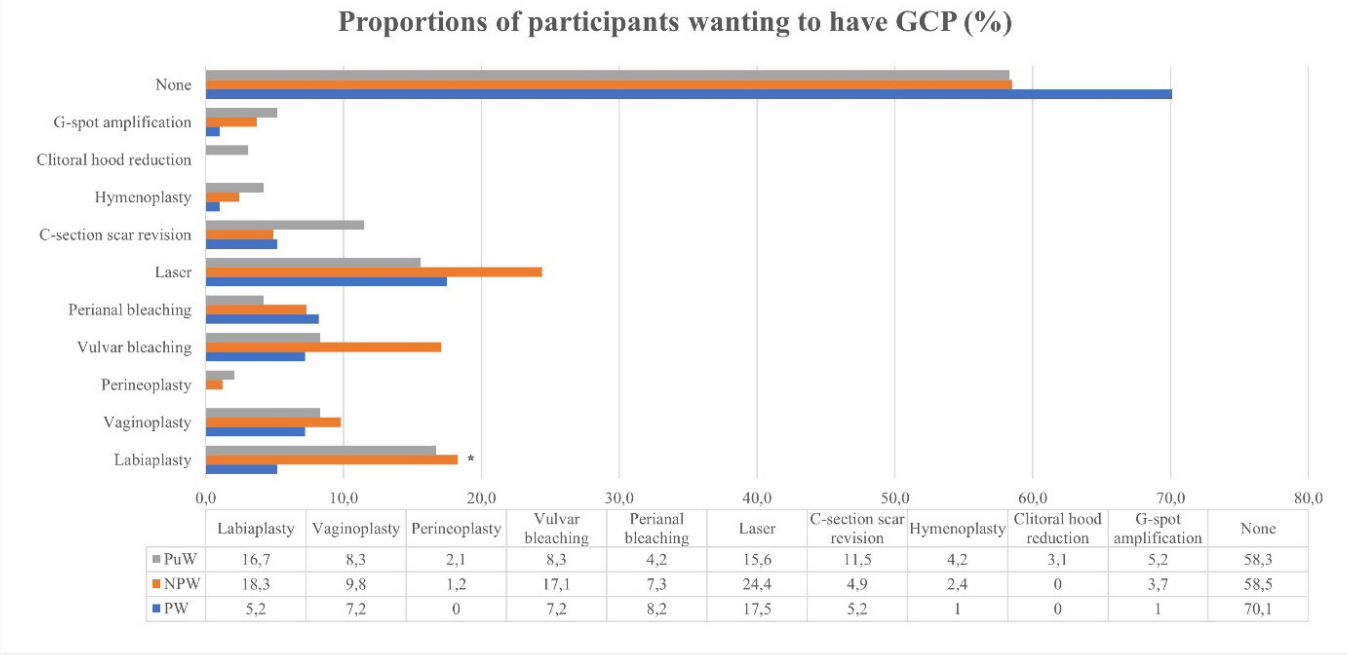
More NPW stated that their attitudes toward GCP would change in the event of pregnancy or childbirth compared to the other groups (p<0.001 and p<0.001, respectively) (Table 3). The attitudes of PW toward GCP were similar to those before pregnancy or in the event of delivery (p>0.05). Likewise, PuW had similar attitudes compared with pregnancy or before pregnancy (p>0.05).

Table 1. Socio-demographic and obstetric data of the participants (n=276)				
Characteristics (n, %)	**PW (n=98)**	**NPW (n=82)**	**PuW (n=96)**	**p**
Age (years)				
≤20	11 (11.7%)	6 (7.3%)	15 (15.6%)	
21-25	34 (36.2%)	27 (32.9%)	18 (18.8%)	
26-30	21 (22.3%)	11 (13.4%)	25 (26%)	0.052
31-35	15 (16%)	17 (20.7%)	19 (19.8%)	
>35	13 (13.8%)	21 (25.6%)	19 (19.8%)	
Education				
Elementary school	29 (29.6%)	17 (20.7%)	38 (39.6%)	
High school	35 (35.7%)	17 (20.7%)	32 (33.3%)	<0.001*
University and higher	34 (34.7%)[a]	48 (58.5%)[b]	26 (27.1%)[a]	
Marital status				
Married	94 (95.9%)[c]	53 (%64.6)[d]	83 (87.4%)[e]	<0.001**
Other	4 (4.1%)[f]	29 (%35.4)[g]	12 (12.6%)[h]	
Employed				
Yes	15 (15.3%)[i]	45 (54.9%)[j]	32 (33.3%)[k]	<0.001***
No	83 (84.7%)[l]	37 (45.1%)[m]	64 (66.7%)[n]	
Socioeconomics				
Low income	32 (32.7%)	19 (23.2%)	42 (43.8%)	
Middle income	57 (58.2%)	52 (63.4%)	46 (47.9%)	0.064
High income	9 (9.2%)	11 (13.4%)	8 (8.3%)	
Residence				
Urban	90 (95.7%)	77 (97.5%)	86 (89.6%)	0.062
Non-urban	4 (4.3%)	2 (2.5%)	10 (10.4%)	
Previous pregnancy				
0	20 (26.3%)		23 (30.3%)	0.605
1	32 (42.1%)		26 (34.2%)	
≥2	24 (31.6%)		27 (35.5%)	
First maternal age (years)				
≤20	10 (30.3%)		16 (40%)	0.230
21-25	9 (27.3%)		13 (32.5%)	
26-30	12 (36.4%)		7 (17.5%)	
31-35	1 (3%)		4 (10%)	
> 35	1 (3%)		0	
Preterm birth				
Yes	12 (21.8%)		12 (22.6%)	0.918
No	43 (78.2%)		41 (77.4%)	
Multiple pregnancy				
Yes	2 (3.6%)		3 (5.9%)	0.670
No	53 (96.4%)		48 (94.1%)	
Delivery method				
Vaginal	18 (35.3%)		15 (29.4%)	0.717
C/S	32 (62.7%)		34 (66.7%)	
Both	1 (2%)		2 (3.9%)	
Abortus				
Yes	17 (30.9%)		11 (20.8%)	0.229
No	38 (69.1%)		42 (79.2%)	

*: Pairwise comparisons; [b] (Z-score = 4.3, p < 0.001) was significantly different from [a]. **: Pairwise comparisons; the differences between [c] and [d], [d] and [e], [f] and [g], and [g] and [h] were significant (Z-scores were 4.1 for [c], -5.6 for [d], -4.1 for [f], and 5.6 for [g]. p < 0.001). ***: Pairwise comparisons; the differences between [i] and [j], [j] and [k], [i] and [k], [l] and [m], [m] and [n], and l and n were significant (Z-scores were -4.7 for i, 4.9 for j, 4.7 for l, and -4.9 for m. p < 0.001). PW: Pregnant women, NPW: Non-pregnant women of reproductive age, PuW: Puerperal women, yrs: years, C/S: Cesarean section, Z-score: Adjusted residual

Table 2. Knowledge and attitudes toward GCP and the FGSIS group scores				
	PW	NPW	PuW	p
Sufficient GCP knowledge				
Yes	27 (28.7%)	26 (31.7%)	27 (28.4%)	0.873
No	67 (71.3%)	56 (68.3%)	68 (71.6%)	
Interested in at least one GCP				
Yes	29 (29.9%)	34 (41.5%)	40 (41.7%)	0.161
No	68 (70.1%)	48 (58.5%)	56 (58.3%)	
Among participants considering having GCP				
Elementary school	5 (17.2%)	4 (11.8%)	9 (22.5%)	0.008*
High school	11 (37.9%)	4 (11.8%)	17 (42.5%)	
University and higher	13 (44.8%)[a]	26 (76.5%)[b]	14 (35%)[a]	
Considering having labiaplasty				
Yes	5 (5.2%)[c]	15 (18.3%)[d]	16 (16.7%)[d]	0.015**
No	92 (94.8%)[e]	67 (81.7%)[f]	80 (83.3%)[f]	
FGSIS [median (IQR)]	21 (19-25)	23 (19-27)	22 (19-27%)	0.518

*: Pairwise comparisons; [b] (Z-score =3.9, p=0.004) was significantly different from [a]. **: Pairwise comparisons; the differences between [c] and [d], and e and [f] were significant (Z-scores were -3.2 for [c] and 3.2 for [e]. p=0.006). PW: Pregnant women, NPW: Non-pregnant women of reproductive age, PuW: Puerperal women, GCP: Genital cosmetic procedures, FGSIS: Female Genital Self-image scale, IQR: Interquartile range, Z-score: Adjusted residual



	Labiaplasty	Vaginoplasty	Perineoplasty	Vulvar bleaching	Perianal bleaching	Laser	C-section scar revision	Hymenoplasty	Clitoral hood reduction	G-spot amplification	None
PuW	16,7	8,3	2,1	8,3	4,2	15,6	11,5	4,2	3,1	5,2	58,3
NPW	18,3	9,8	1,2	17,1	7,3	24,4	4,9	2,4	0	3,7	58,5
PW	5,2	7,2	0	7,2	8,2	17,5	5,2	1	0	1	70,1

Figure 1. Proportions of participants wanting to have GCP according to study group

**: p=0.015. Pairwise comparisons: The proportion of PW was significantly different from the others. (Z-score: -3.2, p=0.006). GCP: Genital cosmetic procedures, PW: Pregnant women, NPW: Non-pregnant women of reproductive age, PuW: Puerperal women, Z-score: Adjusted residuals*

According to the cross-over analysis, most NPW who wanted a laser procedure stated that they would not have it during pregnancy (p=0.006). In addition, 10 and 16 women stated they would have C/S scar revision done in the event of pregnancy and childbirth, respectively (p=0.002) (Table 4). The attitudes toward the other GCP were similar to the intended attitudes in the event of pregnancy or childbirth (p>0.05).

Cosmetic Product Use

The criteria for choice of cosmetics, information sources, and purchasing frequencies of cosmetics are summarized in Table 5. The criteria of choice of cosmetics changed in 34 (36.2%) PW and 14 (15.2%) PuW, compared to before pregnancy, and 49 (69.5%) NPW in the event of pregnancy (p<0.001).

Table 3. Changes in attitudes toward GCP				
Changes/intended changes during/in the event of pregnancy	**PW**	**NPW**	**PuW**	**p**
No	70 (74.5%)[a]	34 (41.5%)[b]	60 (63.2%)[a]	<0.001*
Yes	8 (8.5%)[c]	23 (28%)[d]	5 (5.3%)[c]	
Unknown	16 (17%)	25 (30.5%)	30 (31.6%)	
Changes/intended changes after/in the event of childbirth	**PW**	**NPW**	**PuW**	**p**
No	63 (67%)[e]	34 (41.5%)[f]	64 (68.1%)[e]	<0.001**
Yes	5 (5.3%)[g]	19 (23.2%)[h]	8 (8.5%)[g]	
Unknown	26 (27.7%)	29 (35.4%)	22 (23.4%)	

*: Pairwise comparisons; the differences between [b] and [a], and [d] and [c] were significant (Z-scores were -4.2 for [b] and 4.7 for [d]. p<0.001). **: Pairwise comparisons; the differences between [f] and [e], and [h] and [g] were significant (Z-score of [f]: -4, p=0.003; Z-score of [h]: 3.9, p=0.004). PW: Pregnant women, NPW: Non-pregnant women of reproductive age, PuW: puerperal women, Z-score: Adjusted residual

Table 4. The cross-over analysis of changes in the attitudes toward GCP in the NPW group				
GCP	**Current attitudes**	**Intended attitudes in the event of pregnancy**		**p**
Laser	No (n=42)	No (n=41)	Yes (n=1)	0.006
	Yes (n=15)	No (n=11)	Yes (n=4)	
C-section scar revision	No (n=56)	No (n=46)	Yes (n=10)	0.002
	Yes (n=1)	No (n=0)	Yes (n=1)	
GCP	**Current attitudes**	**Intended attitudes in the event of labor**		**p**
Laser	No (n=38)	No (n=35)	Yes (n=3)	0.227
	Yes (n=16)	No (n=8)	Yes (n=8)	
C-section scar revision	No (n=52)	No (n=36)	Yes (n=16)	<0.001
	Yes (n= 2)	No (n=0)	Yes (n=2)	

GCP: Genital cosmetic procedures, NPW: Non-pregnant women of reproductive age

Additionally, the criteria of choice of cosmetics changed in 13 (14%) PuW after delivery, 16 (18.2%) PW, and 38 (46.3%) NPW in the event of delivery (p<0.001). The criteria of choice were ingredients and doctors' advice for NPW in the case of pregnancy or delivery, and ingredients for PW in the event of delivery.

The information sources considered while choosing cosmetics changed in 13 (15.1%) PW and 6 (6.5%) PuW compared to before pregnancy and 25 (31.3%) NPW in the event of pregnancy (p<0.001). On the other hand, a change was reported by 4 (4.3%) PuW after delivery, 9 (10.2%) PW, and 24 (30%) NPW in the case of delivery (p<0.001). In the case of pregnancy or delivery, dermatologists' advice was the leading information source for NPW. Similarly, dermatologists' advice was the most frequent information source among PW in the event of delivery.

The proportions of participants using MUP regularly are presented in Figure 2. In the PW group, six women stated that they gave up using foundation in pregnancy (p=0.031). Other changes in PW or PuW groups were not significant. The significant changes in the NPW group are presented in Table 6.

The preferred make-up ingredients according to the study groups are summarized in Table 7. In the PW group, nine (10.2%) women stated a change in make-up ingredients compared to before pregnancy, whereas four (4.6%) noted an intended change in the event of delivery. In the NPW group, 16 (20.5%) and 12 (15.4%) women stated an intended change in the make-up ingredients in the event of pregnancy and delivery, respectively; however, these changes were not statistically significant (p>0.05).

The proportions of participants using PCP regularly are presented in Figure 3. The use of face cream (p<0.001), shower gel (p=0.035), anti-aging serum (p=0.003), face night cream (p=0.002), face cleaner (p<0.001), face peeling (p=0.013), sunscreen (p=0.001), moisturizing cream (p=0.002), hair dye (p=0.001), and collagen pills (p=0.001) were significantly more frequent in the NPW group. All participants used more than one PCP. There were no significant changes in any group in the case of pregnancy or delivery in the use of PCP.

Table 5. Criteria for choice of cosmetics, information sources, and purchasing frequencies of cosmetics				
Criteria	**PW**	**NPW**	**PuW**	**p**
Ingredients	54 (56.3%)[a]	44 (53.7%)[a]	31 (32.3%)[b]	0.001*
Price	26 (27.1%)	29 (35.4%)	35 (36.5%)	0.325
Brand	43 (44.8%)[c]	42 (51.2%)[c]	18 (18.8%)[d]	< 0.001**
Odor	28 (29.2%)	24 (29.3%)	20 (20.8%)	0.323
Doctors' advice	39 (40.6%)	39 (47.6%)	34 (35.4%)	0.259
Appearance	4 (4.2%)	5 (6.1%)	7 (7.3%)	0.648
Friends' advice	9 (9.4%)	19 (23.2%)	21 (21.9%)	0.026***
Habit	16 (16.7%)	20 (24.4%)	13 (13.5%)	0.158
User comments	29 (30.2%)	34 (41.5%)	25 (26%)	0.079
Net contents	2 (2.1%)	3 (3.7%)	1 (1%)	0.491
Package	0	6 (7.3%)	2 (2.1%)	0.013***
Satisfaction	47 (49%)	50 (61%)	39 (40.6%)	0.025***
Other	1 (1%)	0	4 (4.2%)	0.091
Information sources	**PW**	**NPW**	**PuW**	**p**
Commercials	10 (10.3%)	16 (19.5%)	16 (16.7%)	0.209
Social media influencers	12 (12.4%)	16 (19.5%)	11 (11.5%)	0.251
Salesperson	9 (9.3%)	8 (9.8%)	18 (18.8%)	0.090
Dermatologists	64 (66%)[e]	60 (73.2%)[e]	45 (46.9%)[f]	0.001****
Internet	11 (11.3%)	15 (18.3%)	12 (12.5%)	0.364
Friends' advice	30 (30.9%)	36 (43.9%)	45 (46.9%)	0.058
Estheticians	19 (19.6%)	18 (22%)	20 (20.8%)	0.927
TV programs	0	4 (4.9%)	7 (7.3%)	0.032***
User comments on blogs	26 (26.8%)	27 (32.9%)	17 (17.7%)	0.063
Purchasing frequency	**PW**	**NPW**	**PuW**	**p**
Once a month	15 (16%)	16 (19.8%)	11 (12.5%)	0.518
Quarterly	15 (16%)	22 (27.2%)	17 (19.3%)	
Twice a year	20 (21.3%)	17 (21%)	20 (22.7%)	
Once a year	23 (24.5%)	12 (14.8%)	18 (20.5%)	
Other	21 (22.3%)	14 (17.3%)	22 (25%)	

*: Pairwise comparisons; [b] (Z-score: -3.6, p < 0.001) was significantly different from [a]. **: Pairwise comparisons: [d] (Z-score: -4.7, p<0.001) was significantly different from [c]. ***: Pairwise comparisons were insignificant according to post-hoc analysis with the Bonferroni adjustment. ****: Pairwise comparisons; [f] (Z-score: -3.6, p < 0.001) was significantly different from [e]. PW: pregnant women, NPW: Non-pregnant women of reproductive age, PuW: Puerperal women, Z-score: Adjusted residual

DISCUSSION

This study provides information on the attitudes toward GCP and the general use patterns for 36 widely used cosmetic products in PW, NPW, and PuW. More than half of the participants in all groups had negative attitudes toward GCP. Although FGSIS scores did not significantly differ between the groups, the scores were significantly lower among women who would like to have at least one GCP. In the event of pregnancy or delivery, the leading information source for cosmetics was dermatologists' advice, and the criteria of choice were ingredients and doctors' advice in the NPW group. Ingredients was the criterion of choice, and dermatologists' advice was the leading information source for PW in the event of delivery.

Genital Cosmetic Procedures and FGSIS Scores

Many women experience undesirable changes in their genitalia, affecting sexual life, self-consciousness, and quality of life, whether due to childbirth, physical factors, or as a result of menopause. Nowadays, those unwanted changes can be improved with GCP.

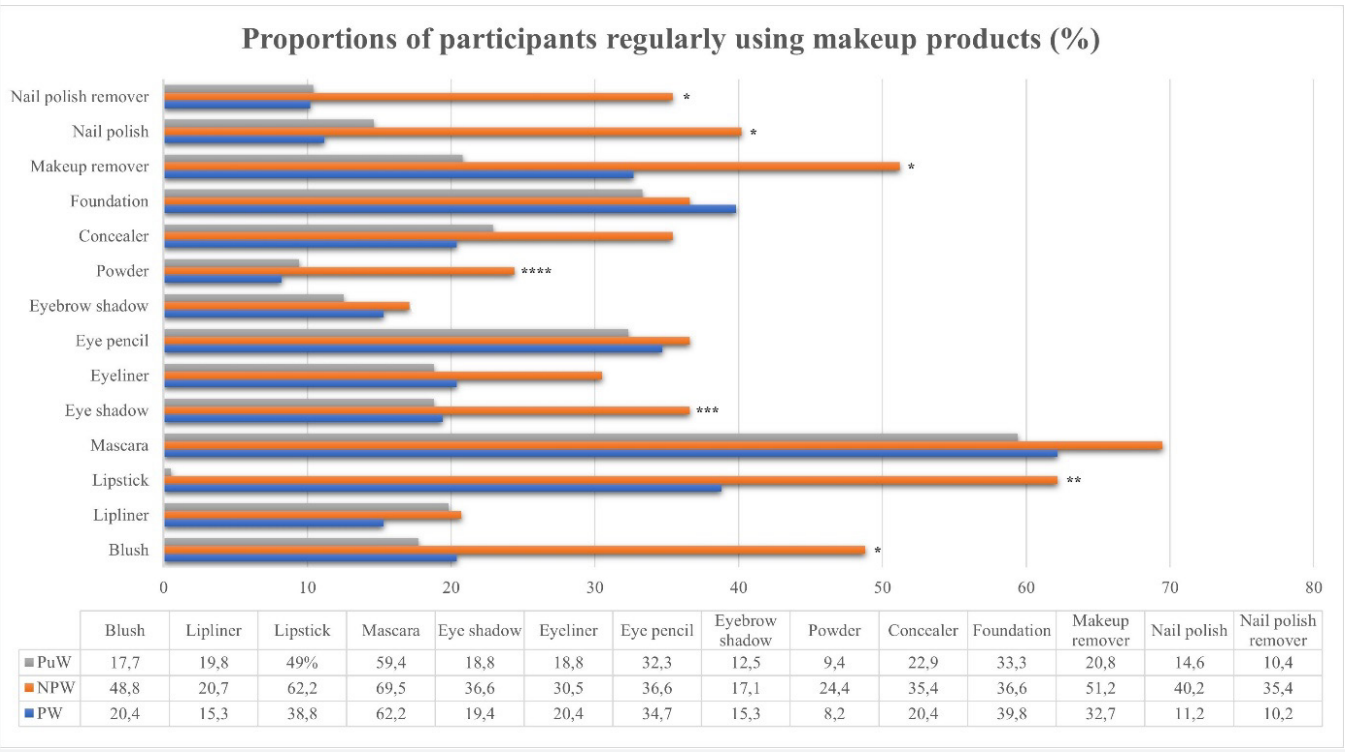
Multiple factors play a role in the increasing demand for GCP among women, including information on social media, the Internet, and TV or the absence of accurate information about normal genital anatomy. In addition, women's beauty perceptions may be changed by exposure to images of modified vulvas[10]. Therefore, dissatisfaction with genital self-image

Table 6. The cross-over analysis of changes in the use of MUP in the NPW group in the events of pregnancy or labor				
Product	**Current use**	**Intended use in the event of pregnancy**		**p**
Blush	No (n=42)	No (n=42)	Yes (n=0)	0.031
	Yes (n=40)	No (n=6)	Yes (n=34)	
Lipstick	No (n=31)	No (n=31)	Yes (n=0)	0.001
	Yes (n=51)	No (n=11)	Yes (n=40)	
Mascara	No (n=25)	No (n=25)	Yes (n=0)	0.008
	Yes (n=57)	No (n=8)	Yes (n=49)	
Make-up cleaner	No (n=40)	No (n=40)	Yes (n=0)	0.016
	Yes (n=42)	No (n=7)	Yes (n=35)	
owder	No (n=62)	No (n=62)	Yes (n=0)	0.031
	Yes (n=20)	No (n=6)	Yes (n=14)	
Concealer	No (n=53)	No (n=53)	Yes (n=0)	0.031
	Yes (n=29)	No (n=6)	Yes (n=23)	
Foundation	No (n=52)	No (n=52)	Yes (n=0)	0.016
	Yes (n=30)	No (n=7)	Yes (n=23)	
Product	**Current use**	**Intended use in the event of labor**		**p**
Lipstick	No (n=31)	No (n=31)	Yes (n=0)	0.004
	Yes (n=51)	No (n=9)	Yes (n=42)	
Mascara	No (n=25)	No (n=25)	Yes (n=0)	0.016
	Yes (n=57)	No (n=7)	Yes (n=50)	
Make-up cleaner	No (n=40)	No (n=40)	Yes (n=0)	0.008
	Yes (n=42)	No (n=8)	Yes (n=34)	
Foundation	No (n=52)	No (n=52)	Yes (n=0)	0.031
	Yes (n=30)	No (n=6)	Yes (n=24)	
MUP: Makeup products. NPW: Non-pregnant women of reproductive age				

Table 7. Makeup ingredient choices				
Makeup ingredients	**PW**	**NPW**	**PuW**	**p**
Fat-free	17 (17.5%)	21 (25.9%)	14 (14.6%)	0.144
Alcohol-free	36 (37.1%)	39 (48.1%)	32 (33.3%)	0.117
Paraben-free	37 (38.1%)	40 (49.4%)[a]	24 (25%)b	0.003*
Unpolished	23 (23.7%)	34 (42%)	25 (26%)	0.018**
Non-comedogenic	35 (36.1%)[c]	38 (46.9%)[c]	19 (19.8%)[d]	0.001***
Natural products	62 (63.9%)	41 (50.6%)	55 (57.3%)	0.201
*: Pairwise comparisons; [b] (Z-score: -3.1, p=0.0082) was significantly different from [a]. **: Pairwise comparisons were insignificant according to post-hoc analysis with the Bonferroni adjustment. ***: Pairwise comparisons, [d] (Z-score: -3.5, p < 0.001) was significantly different from [c]. PW: pregnant women, NPW: Non-pregnant women of reproductive age, PuW: Puerperal women, Z-score: Adjusted residual				

may also be the reason for the increasing demand for GCP. On the other hand, cultural or religious reasons, such as the great importance of virginity, may cause women to seek GCP, like hymenoplasty. We do not know the exact statistics of GCP among the female population or how much information the women or healthcare professionals have about GCP in Türkiye. Thus, further studies are needed on this topic.

Most participants did not think they had sufficient knowledge about GCP. In addition, more than half of the women in all groups stated that they did not want GCP (Table 2). This attitude may be attributed to the paucity of knowledge about GCP or to the shyness and private approaches to genital problems in the study population. The other attributable reasons are financial resources to pay for such procedures and education level. Most participants were low/middle income in all groups. The proportion of women with an education level of university or higher was significantly higher among those who considered having GCP in the NPW and PuW groups. Therefore, we speculated that the lower the education level, the lower the knowledge or positive attitudes toward GCP.



	Blush	Lipliner	Lipstick	Mascara	Eye shadow	Eyeliner	Eye pencil	Eyebrow shadow	Powder	Concealer	Foundation	Makeup remover	Nail polish	Nail polish remover
PuW	17,7	19,8	49%	59,4	18,8	18,8	32,3	12,5	9,4	22,9	33,3	20,8	14,6	10,4
NPW	48,8	20,7	62,2	69,5	36,6	30,5	36,6	17,1	24,4	35,4	36,6	51,2	40,2	35,4
PW	20,4	15,3	38,8	62,2	19,4	20,4	34,7	15,3	8,2	20,4	39,8	32,7	11,2	10,2

Figure 2. Proportions of participants regularly using makeup products

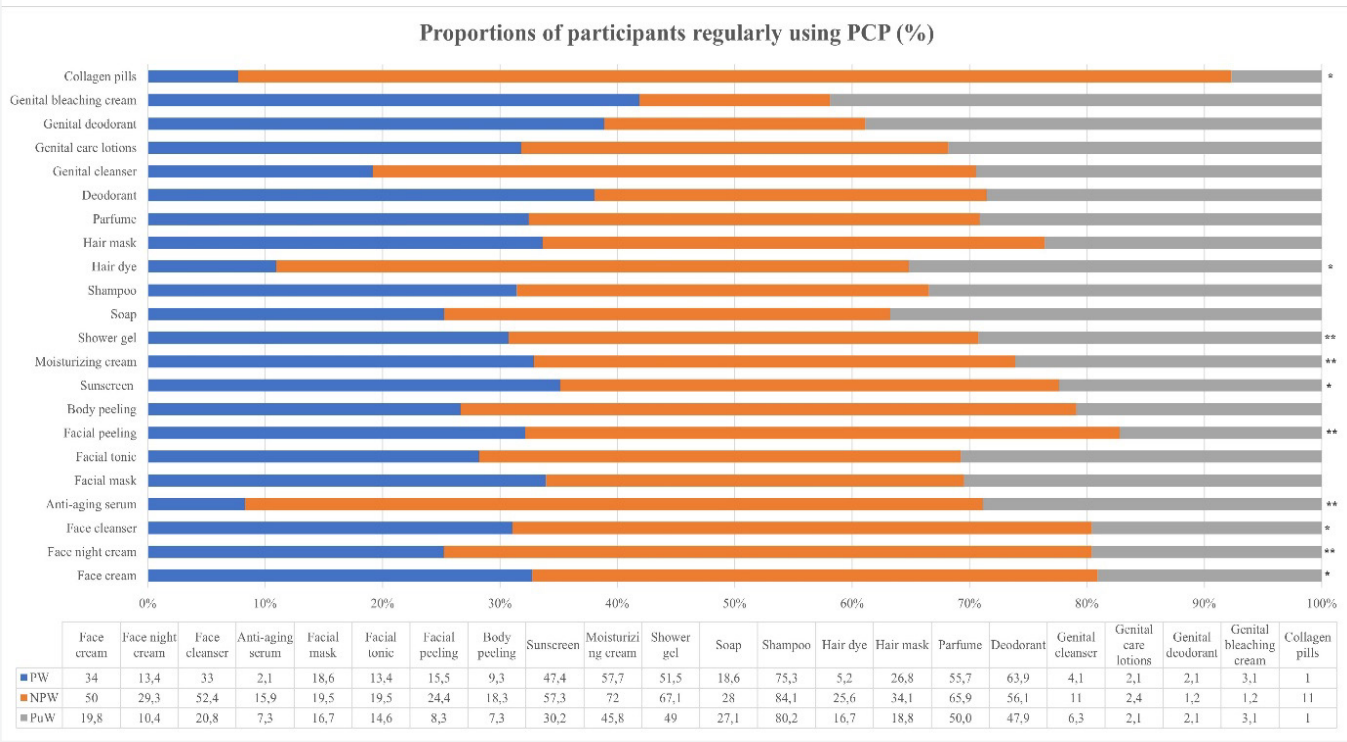
**: p<0.001. Post-hoc tests: Nail polish remover; Z-score of NPW: 5, p<0.001. Nail polish; Z-score of NPW: 5.1, p<0.001. Make-up remover; Z-score of NPW: 3.9, p<0.001. Blush; Z-score of NPW: 5, p < 0.001. **: p=0.007. Pairwise comparisons were insignificant according to post-hoc analysis with the Bonferroni adjustment. ***: p=0.008. Post-hoc tests: Eye shadow; Z-score of NPW: 3.1, p=0.0082. ****: p=0.002. Post-hoc tests: Powder; Z-score of NPW: 3.5, p<0.001. PW: pregnant women, NPW: Non-pregnant women of reproductive age, PuW: Puerperal women, Z-score: Adjusted residuals*

The most desirable GCP were laser, perianal/vulvar bleaching, and vaginoplasty for PW, laser, labioplasty, vulvar bleaching, and vaginoplasty for NPW, and labioplasty, laser, C/S scar revision, vaginoplasty, and vulvar bleaching for PuW. For the NPW group, the proportion of women who thought to change their attitudes toward GCP in the event of pregnancy or delivery was significantly higher compared to the other groups. The pregnant and PuW had similar attitudes compared to before or after pregnancy or in the event of delivery. However, for NPW, only the attitudes toward laser and C/S scar revision significantly changed. The majority of women had negative attitudes toward laser in the event of pregnancy. C/S scar revision was the leading GCP that women would have in the case of pregnancy or delivery (Table 4).

Seeking GCP can be related to the negative effect of genital self-image on quality of life, dissatisfaction with genital appearance, or sexual performance[1,10]. Compatible with the literature, FGSIS scores were significantly lower in women who wanted GCP. However, participants were not evaluated regarding psychological background, such as depression, anxiety, or body dysmorphic disorder.

Cosmetic Product Use

Most PCP, especially genital hygiene and skincare products such as shampoo, shower gel, moisturizing cream, perfume, and deodorant, were widely used by all study groups as expected, compatible with the literature[11,12]. However, NPW generally used less PCP and MUP outside of pregnancy compared to the other studies[13-16]. Some PCP usages were significantly lower in PW and PuW compared to NPW, including face creams (day and night), facial cleansers, anti-aging sera, facial peeling, hair dyes, sunscreens, moisturizing creams, shower gels, and collagen pills. However, it was uncertain whether this trend was due to awareness of the potentially harmful effects of PCP or reduced self-care of women during those periods. Likewise, using some MUP, including blush, lipstick, eyeshadow, powder, make-up remover, nail polish, and nail polish remover,



	Face cream	Face night cream	Face cleanser	Anti-aging serum	Facial mask	Facial tonic	Facial peeling	Body peeling	Sunscreen	Moisturizing cream	Shower gel	Soap	Shampoo	Hair dye	Hair mask	Parfume	Deodorant	Genital cleanser	Genital care lotions	Genital deodorant	Genital bleaching cream	Collagen pills
PW	34	13,4	33	2,1	18,6	13,4	15,5	9,3	47,4	57,7	51,5	18,6	75,3	5,2	26,8	55,7	63,9	4,1	2,1	2,1	3,1	1
NPW	50	29,3	52,4	15,9	19,5	19,5	24,4	18,3	57,3	72	67,1	28	84,1	25,6	34,1	65,9	56,1	11	2,4	1,2	1,2	11
PuW	19,8	10,4	20,8	7,3	16,7	14,6	8,3	7,3	30,2	45,8	49	27,1	80,2	16,7	18,8	50,0	47,9	6,3	2,1	2,1	3,1	1

Figure 3. Proportions of participants regularly using PCP

**: p≤0.001. Post-hoc tests: Collagen pills; Z-score of NPW: 3.8, p<0.001. Hair dye; Z-score of PW: -3.4, p<0.001. Sunscreen; Z-score of PuW: -3.5, p<0.001. Face cleanser; Z-score of NPW: 4.1, Z-score of PuW: -3.5, p<0.001). Face cream; Z-score of NPW: 3.7, Z-score of PuW: -3.6, p<0.001. **: p<0.05. Post-hoc tests: Moisturizing cream; Z-score of NPW: 3.1, p=0.0082. Anti-aging serum; Z-score of NPW: 3.1, p=0.0082. Face night cream; Z-score of NPW: 3.5, p<0.001. Pairwise comparisons of "shower gel" and "facial peeling" were insignificant according to post-hoc analysis with the Bonferroni adjustment. PCP: Personal care products, PW: Pregnant women, NPW: Non-pregnant women of reproductive age, PuW: Puerperal women, Z-score: Adjusted residuals*

was significantly lower in PW and PuW. On the other hand, a significant proportion of NPW thought to stop using blush, powder, and concealer in the case of pregnancy, as well as lipstick, mascara, make-up remover, and foundation in the case of pregnancy or delivery. However, changes or intended changes were insignificant in the other groups, possibly related to the unequal educational levels between the study groups.

Most studies on the cosmetic use patterns in PW gathered data using 24 to 48-hour recall questionnaires for different PCP or MUP[11,12,17,18]. Marie et al.[16] performed a questionnaire study about the routine use habits of cosmetics irrespective of usage frequency in pregnant and NPW. In that study, the products most commonly given up by PW were nail polish, nail polish remover, and hair dye. In addition, safe ingredients and odor were the new choice criteria among PW who changed their cosmetic use. In the event of pregnancy, NPW stated ingredients and professional advice as the criteria of choice[16]. In the present study, a few PW indicated that they had given up using foundations (p=0.031). The other indicated changes were not significant. In the case of pregnancy or delivery, the ingredients and doctors' advice were the criteria of choice.

Study Limitations

This study has some limitations, including a small sample size and potential information bias due to the cross-sectional and self-report questionnaire design. The possible confounders (education level, marital status, and employment status) that could influence the attitudes toward GCP and the prevalence of cosmetic use were not homogeneous between the study groups. Because 93.3% of PW were in the third trimester, there may be a potential recall bias about cosmetic use habits or attitudes toward GCP. This study could not assess the use habits according to the periods of pregnancy. Moreover, it did not analyze women's perceptions of risk related to cosmetics.

Although there was no significant relationship between the attitudes toward GCP and obstetric features, 77% of PuW gave birth by C/S, which could lead to an underestimation of the effects of vaginal birth on GCP demands. Lastly, the study could not evaluate women's exact level of knowledge about GCP one by one.

CONCLUSION

This study compared the attitudes toward GCP and cosmetic use in PW, NPW, and PuW. The results of this study would help raise healthcare professionals' awareness to inform women about GCP and cosmetics use during pregnancy or puerperium. Further studies with more substantial participant sizes representative of the general population are needed to verify and further this study's findings.

Ethics

Ethics Committee Approval: The study was approved by the Non-Interventional and Clinical Research Ethics Committee of Tekirdağ Namık Kemal University (decision no: 2023.28.02.06, date: 28.02.2023) and was carried out following the principles of the 1975 Declaration of Helsinki as revised in 2000.

Informed Consent: All participants signed informed consent before participating in the study.

Acknowledgment: The authors declared that they did not receive financial support

Footnotes

Authorship Contributions

Concept: Ö.Z., İ.Ö.A., M.T.A., Design: Ö.Z., İ.Ö.A., M.T.A., Data Collection or Processing: İ.Ö.A., M.T.A., Analysis or Interpretation: Ö.Z., Literature Search: Ö.Z., İ.Ö.A., M.T.A., Writing: Ö.Z.

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

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

REFERENCES

1. Shaw D, Allen L, Chan C, Kives S, Popadiuk C, Robertson D, et al. Guideline No. 423: female genital cosmetic surgery and procedures. J Obstet Gynaecol Can. 2022;44:204-14.
2. Cosmetic Surgery National Data Bank Statistics. Aesthet Surg J. 2018;38(suppl_3):1-24.
3. Bocca B, Pino A, Alimonti A, Forte G. Toxic metals contained in cosmetics: a status report. Regul Toxicol Pharmacol. 2014;68:447-67.
4. Zota AR, Shamasunder B. The environmental injustice of beauty: framing chemical exposures from beauty products as a health disparities concern. Am J Obstet Gynecol. 2017;217:418.e1-e6.
5. Larsson K, Ljung Björklund K, Palm B, Wennberg M, Kaj L, Lindh CH, et al. Exposure determinants of phthalates, parabens, bisphenol a and triclosan in Swedish mothers and their children. Environ Int. 2014;73:323-33.
6. Ko A, Kang HS, Park JH, Kwon JE, Moon GI, Hwang MS, et al. The association between urinary benzophenone concentrations and personal care product use in Korean adults. Arch Environ Contam Toxicol. 2016;70:640-6.
7. Philippat C, Mortamais M, Chevrier C, Petit C, Calafat AM, Ye X, et al. Exposure to phthalates and phenols during pregnancy and offspring size at birth. Environ Health Perspect. 2012;120:464-70.
8. Li H, Zheng J, Wang H, Huang G, Huang Q, Feng N, et al. Maternal cosmetics use during pregnancy and risks of adverse outcomes: a prospective cohort study. Sci Rep. 2019;9:8030.
9. Ellibes Kaya A, Yassa M, Dogan O, Basbug A, Pulatoglu C, Caliskan E. The female genital self-image scale (FGSIS): cross-cultural adaptation and validation of psychometric properties within a Turkish population. Int Urogynecol J. 2019;30:89-99.
10. Veale D, Eshkevari E, Ellison N, Costa A, Robinson D, Kavouni A, et al. Psychological characteristics and motivation of women seeking labiaplasty. Psychol Med. 2014;44:555-66.
11. Lang C, Fisher M, Neisa A, MacKinnon L, Kuchta S, MacPherson S, et al. Personal care product use in pregnancy and the postpartum period: implications for exposure assessment. Int J Environ Res Public Health. 2016;13:105.
12. Meeker JD, Cantonwine DE, Rivera-González LO, Ferguson KK, Mukherjee B, Calafat AM, et al. Distribution, variability, and predictors of urinary concentrations of phenols and parabens among pregnant women in Puerto Rico. Environ Sci Technol. 2013;47:3439-47.
13. Wu XM, Bennett DH, Ritz B, Cassady DL, Lee K, Hertz-Picciotto I. Usage pattern of personal care products in California households. Food Chem Toxicol. 2010;48:3109-19.
14. Shaaban H, Alhajri W. Usage patterns of cosmetic and personal care products among female population in Saudi Arabia: important factors for exposure and risk assessment. J Environ Public Health. 2020;2020:8434508.
15. Manová E, von Goetz N, Keller C, Siegrist M, Hungerbühler K. Use patterns of leave-on personal care products among Swiss-German children, adolescents, and adults. Int J Environ Res Public Health. 2013;10:2778-98.
16. Marie C, Cabut S, Vendittelli F, Sauvant-Rochat MP. Changes in cosmetics use during pregnancy and risk perception by women. Int J Environ Res Public Health. 2016;13:383.
17. Just AC, Adibi JJ, Rundle AG, Calafat AM, Camann DE, Hauser R, et al. Urinary and air phthalate concentrations and self-reported use of personal care products among minority pregnant women in New York city. J Expo Sci Environ Epidemiol. 2010;20:625-33.
18. Buckley JP, Palmieri RT, Matuszewski JM, Herring AH, Baird DD, Hartmann KE, et al. Consumer product exposures associated with urinary phthalate levels in pregnant women. J Expo Sci Environ Epidemiol. 2012;22:468-75.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.68725
Nam Kem Med J 2025;13(3):322-332

Breast Cancer Surgery Quality of Life Scale: Turkish Validity and Reliability Study

Meme Kanseri Cerrahisi Yaşam Kalitesi Ölçeği: Türkçe Geçerlik ve Güvenirlik Çalışması

Serap KORKMAZ[1], Ebru ESEN[2], Münevver SÖNMEZ[3]

[1]*Cyprus International University Faculty of Health Sciences, Department of Nursing, Nicosia, Northern Cyprus*
[2]*İstinye University, Liv Hospital Bahçeşehir, Department of General Surgery, İstanbul, Türkiye*
[3]*University of Health Sciences Türkiye Gülhane Faculty of Nursing, Ankara, Türkiye*

ABSTRACT

Aim: This study aimed to evaluate the validity and reliability of the Turkish version of the breast cancer surgery quality of life scale (BCSQOL) among patients who underwent surgical treatment for breast cancer.

Materials and Methods: The study sample consisted of 250 female patients who had undergone breast cancer surgery. Following the translation of the scale into Turkish, content and face validity analyses were conducted. Construct validity was assessed using confirmatory factor analysis. For reliability testing, the test-retest method was employed. Additionally, omega and Cronbach's alpha coefficients were calculated for the subdimensions of the scale.

Results: The chi-square/DF ratio was found to be 1.064, and the root mean square error of approximation value was 0.017. The comparative fit index, adjusted goodness-of-fit index, normed fit index, Tucker-Lewis index, and incremental fit index values indicated a "good fit", while the standardized root mean square residual and goodness of fit index values indicated an "acceptable fit" according to confirmatory factor analysis results. Cronbach's alpha coefficients for the subdimensions ranged between 0.79 and 0.95, and omega coefficients ranged between 0.97 and 1.00. The total Cronbach's alpha coefficient of the scale was 0.79. The scale comprises 49 items across 8 subdimensions: physical activity, pain, feelings, body image, physical health, sexual function, general health, and relationships with others.

Conclusion: Based on the findings, the Turkish version of the BCSQOL scale is a valid and reliable instrument for assessing quality of life in breast cancer patients following surgical treatment.

Keywords: Breast cancer, breast surgery, validity and reliability, quality of life

ÖZ

Amaç: Bu araştırmanın amacı meme kanseri nedeniyle cerrahi tedavi uygulanan hastalarda meme kanseri cerrahisi yaşam kalitesi ölçeğinin (MKCYKÖ), Türkçe versiyonunun, geçerlik ve güvenirliğini test etmektir.

Gereç ve Yöntem: Meme kanseri nedeniyle ameliyat olan 250 kadın hasta çalışmanın örneklemini oluşturdu. Ölçeğin Türkçe tercümesi yapıldıktan sonra kapsam ve yüzey geçerliği hesaplandı. Ölçeğin yapı geçerliği için doğrulayıcı faktör analizi uygulandı. Ölçeğin güvenirlik analizinde test-tekrar test yöntemi kullanıldı. Ayrıca ölçeğin ve alt boyutlarının omega ve Cronbach alfa katsayısı hesaplandı.

Bulgular: Çalışmada ki-kare istatistik değeri 1,064, yaklaşık hataların kök ortalama kare değeri ise 0,017 olarak elde edildi. Doğrulayıcı faktör analizi sonucunda elde edilen diğer uyum indekslerinden karşılaştırmalı uyum indeksi, ayarlanmış iyilik hızı indeksi, normlaştırılmış uyum indeksi, Tucker-Lewis indeksi ve artışlı uyum indeksi değerlerinin "mükemmel uyum", standartlaştırılmış kök ortalama kare artık ve iyilik uyum testi değerlerinin

Address for Correspondence: Münevver SÖNMEZ MD, University of Health Sciences Türkiye Gülhane Faculty of Nursing, Ankara, Türkiye
E-mail: munevver.sonmez@sbu.edu.tr **ORCID ID:** orcid.org/0000-0001-7646-1864
Received: 24.03.2025 **Accepted:** 14.07.2025 **Publication Date:** 07.10.2025

Cite this article as: Korkmaz S, Esen E, Sönmez M. Breast cancer surgery quality of life scale: Turkish validity and reliability study. Nam Kem Med J. 2025;13(3):322-332

ise "kabul edilebilir düzeyde uyum" sağladığı tespit edildi. Alt boyutlara ait Cronbach alfa katsayısının 0,79-0,95 arasında ve omega katsayısının ise 0,97-1,00 arasında değiştiği saptandı. Ölçeğin toplam Cronbach alfa değeri 0,79'dur. Ölçek toplam 49 madde ve fiziksel aktivite, ağrı, duygular, beden imajı, fiziksel sağlık, cinsel işlev, genel sağlık ve diğerleriyle ilişkiler olmak üzere 8 alt boyuttan oluşmaktadır.

Sonuç: Çalışma verileri değerlendirildiğinde; MKCYKÖ, Türkçe versiyonunun, meme kanseri hastalarda cerrahi tedavi sonrası yaşam kalitesinin değerlendirilmesinde geçerli ve güvenilir bir ölçüm aracı olduğu sonucuna varılmıştır.

Anahtar Kelimeler: Meme kanseri, meme cerrahisi, geçerlik ve güvenirlik, yaşam kalitesi

INTRODUCTION

Breast cancer is the most common type of cancer in women in Türkiye, as elsewhere around the world[1-3]. According to Global Cancer Statistics 2020 data for 36 cancer types in 185 countries, breast cancer takes place on the top in all age groups with a rate of 24.5% in women and 23.9% in Türkiye[4,5]. According to the report of the International Agency for Research on Cancer (2022), on the other hand, approximately 2.3 million new cases were detected worldwide in 2020, and the rate of new cases is increasing day by day[1,6]. However, the survival rate of breast cancer patients has increased in recent years due to developments in diagnosis and treatment[1]. One of the most important treatment methods for breast cancer is surgical treatment[3]. However, although surgical treatment increases the survival rate in patients, this may cause some problems[2] such as physical activity limitations, fatigue, pain, sleep disturbances and psychosocial issues (e.g., anxiety, depression). In particular, it leads to a decrease in self-esteem, aesthetic losses, sexual dysfunction, deterioration of body image[1,4]. These problems experienced after breast cancer surgeries may significantly reduce the postoperative quality of life of patients[3,7,8].

It is recommended to use a valid and reliable measurement tool to assess the quality of life of patients. In the literature, the Turkish adaptation, validity, and reliability study of the "functional assessment of cancer therapy-breast" scale was conducted by Yalçın and Gürkan[9] for this purpose. However, this scale primarily focuses on the general impact of cancer treatment rather than specifically targeting the postoperative experiences of patients who have undergone breast cancer surgery. Therefore, there is a need for a specific scale to assess the quality of life of patients who have undergone breast cancer surgery. Breast cancer surgery quality of life scale (BCSQOL), developed by El Farrah[10], is specifically designed to address the physical and psychosocial problems experienced by individuals undergoing breast cancer surgery. This scale allows for a more comprehensive assessment of postoperative changes affecting quality of life, including physical activity, pain, aesthetic concerns, and emotional difficulties. Therefore, the use of the BCSQOL in this study provides a more accurate and relevant assessment of quality of life for this patient population.

The Turkish adaptation of the BCSQOL scale has not been conducted to date. In this regard, it is both clinically and academically significant to translate and validate the quality of life scale, specifically developed for individuals who have undergone breast cancer surgery, into Turkish. The aim of this research was to evaluate BCSQOL the validity and reliability of its Turkish version.

Research Questions

Is the Turkish version of the BCSQOL scale valid?

Is the Turkish version of the BCSQOL scale reliable?

MATERIALS AND METHODS

Research Type

This study is methodological research aimed at the adaptation of a measurement scale. The data were collected between November 2022 and April 2023 at the breast clinic of a training and research hospital in Ankara.

Study Design and Eligibility Criteria

In scale validity and reliability studies, it is recommended to include a sample size that is 5 to 10 times the number of items in the scale[11]. Accordingly, the minimum required sample size was calculated as 245 participants, based on the 49 items of the BCSQOL (49 items × 5). To account for a potential 10% data loss, the target sample size was increased to 270 participants. Ultimately, the study was completed with 250 participants who met the inclusion criteria and were selected through a non-probability sampling method.

The inclusion criteria were as follows: female individuals aged 18 years or older; those who had undergone surgical treatment for a breast cancer diagnosis; were on the third postoperative day; had no central nervous system metastasis; had no psychiatric or neurological diagnoses (such as psychosis, depression, or delirium); were able to communicate in Turkish; voluntarily agreed to participate; and provided both verbal and written informed consent.

Exclusion criteria included: individuals who were receiving treatment for another type of cancer or had a previous breast cancer diagnosis and treatment; those who had only

undergone a breast biopsy; individuals with impaired cognitive functions; and those who declined to participate in the study.

Data were collected from eligible female patients on the third day after breast cancer surgery through face-to-face interviews conducted in the surgical clinic. Each data collection session lasted approximately 10 to 15 minutes.

To assess the test-retest reliability of the Turkish version of the scale, it was re-administered to 75 patients from the same sample group during their outpatient clinic follow-up visits 15 days later. Data were collected using the "Patient Information Form" and the Turkish version of the BCSQOL scale.

The form includes two items designed to collect demographic and clinical information, specifically the patient's age and the type of surgical procedure performed.

Breast Cancer Surgery Quality of Life Scale

The scale was developed by El Farrah in 2003 to evaluate the quality of life of individuals who had undergone breast cancer surgery. The scale includes 49 items and 8 subdimensions. Each subdimension is measured with a different Likert-type frequency. The subdimensions of the scale, the number of items, and the minimum and maximum scores that may be obtained from each subdimension are given in Table 1. Physical activity consists of 8 questions, each scored between 0 and 2, resulting in a total score range of 0-16. Pain includes 7 questions, each scored between 0 and 3, with a total score range of 0-21. The feelings subdimension consists of 6 questions, each scored between 0 and 3, with a total score range of 0-18. Body image includes 3 questions, with each scored between 0 and 3, resulting in a total score range of 0-9. Physical health consists of 7 questions, each scored between 0 and 1, with a total score range of 0-7. Sexual functioning includes 4 questions, each scored between 0 and 1, with a total score range of 0-4. General health consists of 7 questions, each scored between 0 and 2, resulting in a total score range of 0-14. Finally, the relationships with others subdimension includes 7 questions, each scored between 0 and 3, with a total score range of 0-21. The subdimension scores are obtained by dividing the score from each subdimension by the maximum score for that subdimension. This process converts the scores obtained for each subdimensions into a value between 0 and 1 (for instance, suppose the total score derived from the physical activity subdimensions is 10. This score is divided by 16, which is the maximum score that can be obtained from the subdimensions, and a value between 0 and 1 is found for that subdimensions. So the score for the physical activity subdimensions is 10/16=0.625). This process is repeated for each subdimensions. The result is a total score between 0 and 8, obtained by summing the scores from the eight subdimensions. A higher score indicates a better quality of life for the individual who has undergone breast cancer surgery. The items in the scale [post-surgical symptoms and physical functioning (5.2), sexuality (6, 6.4), General health (7.B.1, 7.B.2, 7.B.3, 7.B.4, 7.B.5] are reverse coded[10].

Statistical Analysis

SPSS version 26.0 (Armonk, NY: IBMCorp) package program and R-Project program were used to analyze the data[12]. Descriptive statistics for continuous variables were given with mean and standard deviation. Whether the data were normally distributed or not was evaluated using the Kolmogorov-Smirnov test as well as Skewness-Kurtosis values. In the literature, if the skewness and kurtosis values are between +2 and -2, it is accepted that the data show a normal distribution[13]. The content validity index (CVI) was calculated using the Davis technique to assess content validity. The conformity of the scale to construct validity was evaluated with Kaiser-Meyer-Olkin sampling adequacy (KMO) and Barlett sphericity tests. KMO and Bartlett's sphericity test are two essential tests used in multivariate statistical analyses, such as factor analysis. These tests are employed to assess whether the data are suitable for analysis prior to conducting factor analysis. Confirmatory factor analysis (CFA) was conducted using the lavaanPlot package in the R-Project software to evaluate construct validity[14]. The diagonal weighted least squares (DWLS) technique was used because scale items were defined categorically in CFA. DWLS is a statistical method used for estimating parameters in models. Fit indices values were calculated. No improvements were made to the CFA model, and error covariances were not freed. The graphical summary of the CFA was made in R-Project software with the lavaanPlot package. In the reliability analysis of the scale, item-total correlation coefficient, standardized Cronbach alpha (α), and omega (w) coefficients were calculated. Correlation analyses were conducted to assess the test-retest reliability. A significance level of $p<0.05$ was considered.

Permission to translate the scale into Turkish and to confirm its validity and reliability was obtained via e-mail from El Farrah, the researcher who developed the original scale. Ethical approval for the study University of Health Sciences Türkiye Gülhane Faculty of Medicine Clinical Research Ethics Committee (decision no: 2022/36, date: 25.10.2022). Prior to data collection, all participants were informed about the purpose and procedures of the study, and both written and verbal informed consent were obtained. The study was conducted in accordance with the principles outlined in the Declaration of Helsinki.

RESULTS

The mean age of the participants was 50.06±12.33 years. 17.2% (n=43) of individuals had surgery due to modified radical mastectomy+axillary lymph node dissection, 31.2% (n=78) breast conserving surgery+axillary lymph node dissection, 36.8% (n=92) breast conserving surgery+sentinel lymph node dissection, 14.8% (n=37) modified radical mastectomy+sentinel lymph node dissection.

Language Validity Results

The translation-back translation technique was used for the language validity of the BCSQOL. The original English statements were independently translated into Turkish by two native Turkish-speaking linguists proficient in English. The authors reviewed both translations and, with the assistance of a linguist, consolidated the statements that best reflected the meaning and scope of each item into a single version. An expert in Turkish language and literature further evaluated the suitability of the Turkish version. The finalized Turkish form was then back-translated into English by two different linguists proficient in English, who had not seen the original questionnaire. The researchers controlled both languages and made the necessary corrections and finalized the scale.

Content and Face Validity of Scale

Nine faculty members, who are experts in their respective fields, along with a nurse working in the breast care clinic, were consulted to evaluate the Turkish version of the scale. Their feedback focused on assessing whether the items in the scale adequately cover all relevant aspects of the subject, including the clarity and appropriateness of the expressions, as well as the content validity. Content validity was evaluated using the Davis technique. A minimum CVI of 0.80 is accepted for items[15]. In the study, the CVI of each item in the scale was calculated to be between 0.85 and 1.00. The scale items were modified based on the recommendations provided by the experts. After the content validity of the scale, a pilot study was conducted. To assess face validity, a pilot application was conducted with 10 individuals who had undergone breast cancer surgery and met the same criteria as the sample intended for the scale application. In the pilot study, no negative feedback was received regarding the expressions in the scale. Individuals participating in the pilot study were not included in the research sample.

It is important to determine the sample adequacy of the data set in scale adaptation studies. KMO and Bartlett's test of Sphericity are statistical tests used to evaluate whether your data set is suitable for factor analysis. In the study, the KMO value was calculated as 0.77, and it was determined that the sample size was sufficient for factor analysis and the data were homogeneously distributed. Bartlett's test value was calculated as χ^2=5055.721; p<0.001 and it was determined that there was sufficient correlation between the items for factor analysis.

Confirmatory Factor Analysis

CFA was performed to assess how well the prespecified factors were consistent with the observed data. According to CFA, it was determined that the structural equation model result of the scale was statistically significant at p<0.001 level, the standardized factor loads were positive in eight subdimensions of the scale, and 49 items and eight subdimensions constituting the scale were correlated to the scale structure (Table 2).

Table 1. Scoring breast cancer surgery quality of life scale

BCSQOL subdimensions	Number of items	Minumum score	Maximum score	Total score per items	Total score per subdimensions
Physical activity	8	0	2	0-16	16/16=1
Pain	7	0	3	0-21	21/21=1
Feelings	6	0	3	0-18	18/18=1
Body image	3	0	3	0-9	9/9=1
Physical health	7	0	1	0-7	7/7=1
Sexual functioning	4	0	1	0-4	4/4=1
General health	7	0	2	0-14	14/14=1
Relationship with others	7	0	3	0-21	21/21=1
Overall score	49				0-8

The answers were scored using a positively valued unipolar scale, with the value of 0 assigned to the worst Health related quality of life status, and 3 to the maximum on the 4-point scale. On the 3-point scale, the score ranged from 0 for the minumum to 2 for the maximum. On the 2-point scale, the score ranged from 0 for the minimum to 1 for the maximum

BCSQOL: Breast cancer surgery quality of life

Table 2. Results of confirmatory factor analysis

	Estimate	SE	Z-value	p-value	β_0
A1<---F1	1.000				0.372
A2<---F1	1.971	0.336	5.863	<0.001	0.804
A3<---F1	1.634	0.279	5.853	<0.001	0.708
A4<---F1	1.134	0.219	5.172	<0.001	0.472
A5<---F1	1.488	0.262	5.689	<0.001	0.638
A6<---F1	0.693	0.152	4.554	<0.001	0.296
A7<---F1	0.858	0.166	5.171	<0.001	0.442
A8<---F1	0.563	0.124	4.528	<0.001	0.265
P9<---F2	1.000				0.691
P10<---F2	1.278	0.065	19.616	<0.001	0.760
P11<---F2	1.844	0.089	20.797	<0.001	0.880
P12<---F2	1.459	0.072	20.269	<0.001	0.868
P13<---F2	1.547	0.079	19.652	<0.001	0.777
P14<---F2	1.478	0.074	20.034	<0.001	0.775
P15<---F2	1.705	0.082	20.724	<0.001	0.886
F16<---F3	1.000				0.693
F17<---F3	0.960	0.064	14.900	<0.001	0.621
F18<---F3	1.361	0.080	16.992	<0.001	0.849
F19<---F3	1.396	0.083	16.883	<0.001	0.841
F20<---F3	1.082	0.066	16.332	<0.001	0.758
F21<---F3	0.990	0.067	14.883	<0.001	0.626
B22<---F4	1.000				0.645
B23<---F4	1.101	0.191	5.774	<0.001	0.765
B24<---F4	0.915	0.159	5.762	<0.001	0.660
H25<---F5	1.000				0.469
H26<---F5	1.110	0.276	4.022	<0.001	0.611
H27<---F5	1.135	0.286	3.966	<0.001	0.572
H28<---F5	1.189	0.320	3.715	<0.001	0.540
H29<---F5	1.045	0.280	3.738	<0.001	0.527
H30<---F5	1.273	0.342	3.718	<0.001	0.519
H31<---F5	0.876	0.265	3.308	<0.001	0.386
S32<---F6	1.000				0.602
S33<---F6	1.508	0.431	3.500	<0.001	0.785
S34<---F6	1.621	0.453	3.574	<0.001	0.902
S35<---F6	0.579	0.214	2.712	<0.001	0.260
G36<---F7	1.000				0.485
G37<---F7	1.278	0.166	7.701	<0.001	0.852
G38<---F7	1.210	0.159	7.609	<0.001	0.765
G39<---F7	1.160	0.153	7.567	<0.001	0.715
G40<---F7	1.145	0.152	7.541	<0.001	0.739
G41<---F7	1.056	0.143	7.359	<0.001	0.590
G42<---F7	0.277	0.053	5.212	<0.001	0.252
R43<---F8	1.000				0.522
R44<---F8	1.069	0.147	7.250	<0.001	0.612
R45<---F8	1.128	0.153	7.388	<0.001	0.651
R46<---F8	1.083	0.149	7.274	<0.001	0.552
R47<---F8	0.981	0.141	6.944	<0.001	0.492
R48<---F8	1.187	0.158	7.510	<0.001	0.663
R49<---F8	1.149	0.153	7.531	<0.001	0.640

p<0.05 SE: Standard error, β_0: Standardized coefficient, F1: Physical activity, F2: Pain, F3P: Feelings, F4: Body image, F5: Physical health, F6: Sexual functioning, F7: General health, F8: Relationship with others

Figure 1 presents the CFA model of the scale. The diagram illustrates eight latent factors (F1-F8) and their associated items, with regression coefficients displayed between each factor and its corresponding items. The loadings for physical activity (F1), pain (F2), feelings (F3), body image (F4), physical health (F5), sexual functioning (F6), general health (F7), and relationship with others (F8) are all statistically significant ($p<0.001$). Correlations between latent factors are indicated by double-headed arrows, and the figure provides a visual representation of the multidimensional structure of the scale (Figure 1).

When the goodness-of-fit indices of the scale were examined, it was found that it showed a good fit with the chi-square test of fit (χ^2/df)=1.064, root mean square errors of approximation (RMSEA)=0.017, comparative fit index (CFI)=0.99, adjusted goodness-of-fit index (AGFI)=0.92, standardized root mean square residual (SRMR)=0.065, normed fit index (NFI)=0.99, trucker-lewis index (TLI)=0.99 and incremental fit index (IFI)=0.99; as well as acceptable fit with goodness of fit test (GFI)=0.93 (Table 3).

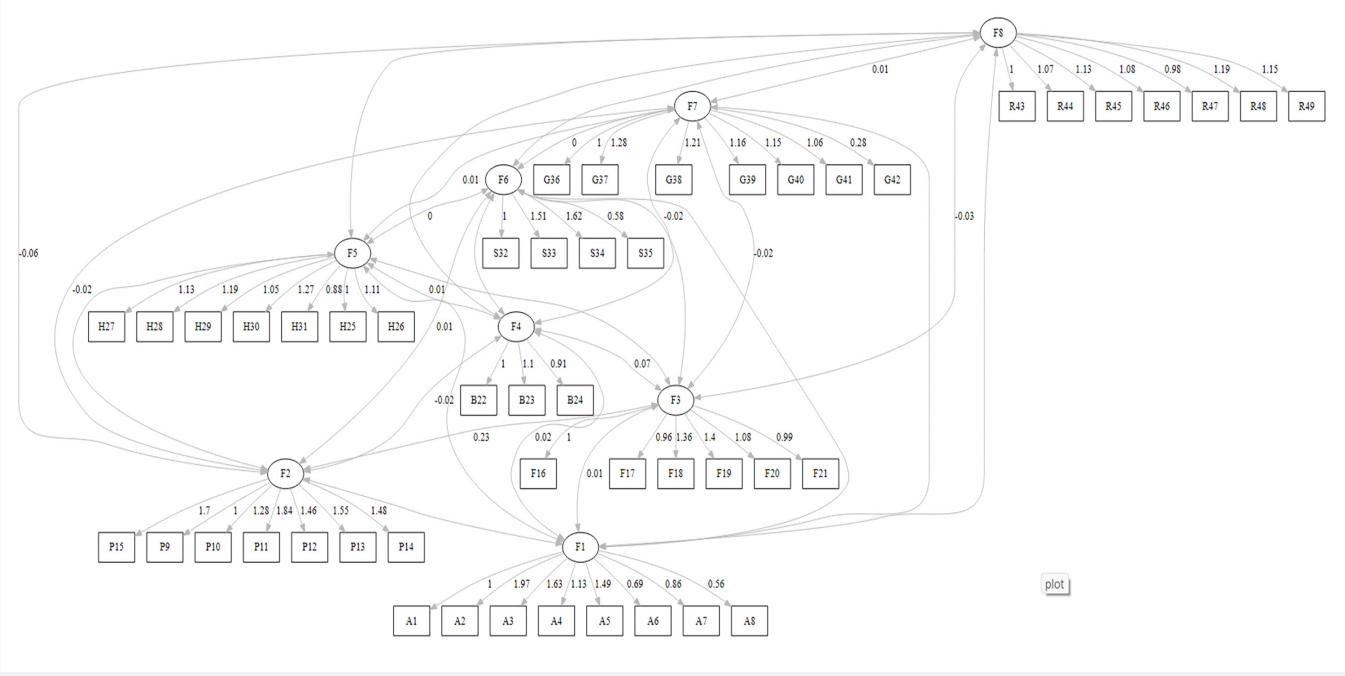


Figure 1. Confirmatory factor analysis results of the BCSQOL scale

BCSQOL: Breast cancer surgery quality of life

Table 3. Fit indices of CFA findings of the BCSQOL scale

Goodness of fit indices	Perfect fit indices	Acceptable fit indices	Fit index values	Obtained in the model
χ^2/df	$0\leq\chi^2/df\leq 3$	$3<\chi^2/df\leq 5$	1.064	"Perfect fit"
GFI	$0.95\leq GFI\leq 1$	$0.90\leq GFI\leq 0.95$	0.930	"Acceptable fit"
AGFI	$0.90\leq GFI\leq 1$	$0.85\leq AGFI\leq 0.90$	0.922	"Perfect fit"
CFI	$0.97\leq CFI\leq 1$	$0.95\leq CFI\leq 0.97$	0.991	"Perfect fit"
IFI	$0.95\leq IFI\leq 1.00$	$0.90\leq IFI\leq 0.95$	0.991	"Perfect fit"
TLI	$0.95\leq TLI\leq 1.00$	$0.90\leq TLI\leq 0.95$	0.990	"Perfect fit"
NFI	$0.95\leq NFI\leq 1.00$	$0.90\leq NFI<0.95$	0.990	"Perfect fit"
RMSEA	$0<RMSEA<0.05$	$0.05<RMSEA<0.08$	0.017	"Perfect fit"
SRMR	$0\leq SRMR\leq 0.05$	$0.05\leq SRMR\leq 0.10$	0.065	"Acceptable fit"

χ^2: Chi-square, df: Degrees of freedom, GFI: Goodness of fit index, AGFI: Adjusted goodness of fit index, IFI: Incremental fit index, TLI: Turker-Lewis index, CFI: Comparative fit index, NFI: Normed fit index, SRMR: Standardized root mean square residual, RMSEA: Root mean square error of approximation, CFA: Confirmatory factor analysis, BCSQOL: Breast cancer surgery quality of life

Reliability Analysis

In the study, coefficient values for an item-total correlation of the scale were 0.35-0.67 for the "physical activity" dimension, 0.70-0.91 for the "pain" dimension, 0.69-0.84 for "feelings" dimension, 0.56-0.69 for "body image" dimension, and 0.59-0.78 for "physical health" dimension, 0.54-0.93 for "sexual functioning" dimension, 0.32-0.85 for "general health" dimension, and 0.51-0.59 for "relationship with other" dimension.

Cronbach's alpha and omega coefficients of subdimensions were examined to measure the internal consistency of the BCSQOL scale. Cronbach's alpha for physical activity subdimensions was 0.81, omega coefficient was 0.87; Cronbach's alpha for pain subdimensions was 0.95, omega coefficient was 0.97; Cronbach's alpha for Feelings subdimensions was 0.90, omega coefficient was 0.95; Cronbach's alpha for body image subdimensions was 0.79, omega coefficient was 1.00; Cronbach's alpha for physical health subdimensions was 0.90, omega coefficient was 1.00; Cronbach's alpha for sexual functioning subdimensions was 0.90, omega coefficient was 1.00; Cronbach's alpha for general health subdimensions was 0.86, omega coefficient was 1.00; Cronbach's alpha for relationship with other subdimensions was 0.81, omega coefficient was 1.00. The total Cronbach's alpha coefficient of the scale is 0.79 (Table 4).

Test-retest Reliability

In the study, test-retest reliability was evaluated with the test-retest method. In this context, the scale was reapplied to 75 people from the same sample group, with an interval of 2 weeks. According to test-retest analysis results; it was found that there was no statistically significant difference between the mean scores of scale factors and total scale ($p>0.05$), the correlation coefficient ranged between 0.27-0.99, and there was a strong correlation between the two measurements ($p<0.001$) (Table 5).

In summary, CFA indicated that the scale structure demonstrated a statistically significant model fit, with all 49 items and eight subdimensions showing positive and meaningful factor loadings. The goodness-of-fit indices supported the validity of the model, with indicators reflecting both good and acceptable fit. The scale also exhibited strong internal consistency across subdimensions, as reflected by high Cronbach alpha and omega coefficients. Furthermore, the test-retest results showed that the scale was stable over time, as there were strong correlations between the two measurements and no significant changes in the average scores.

DISCUSSION

In this study, the psychometric properties of the Turkish version of the BCSQOL were examined for cross-cultural adaptation and validation. The language validity of the Turkish and English versions of the scale was assessed using the translation-back translation method in this study. Reverse-translated English scale was decided to be compatible with the original English scale. Thus, the language adaptation of the Turkish scale was provided.

Content Validity

Scale validity is defined as the ability of a measurement tool to accurately measure the concept or feature it wants to measure[16]. In this study, the content and construct validity of the scale were examined per validity analysis of the scale.

Although different methods are used in the evaluation of content validity, the most preferred method is to ask for an expert's opinion[17]. In the study, a total of 10 experts were asked for their opinions to evaluate the content validity. In the study, the CVI value of the scale items was found to be between 0.85-1.00, and the total CVI value was found to be 0.90. The values obtained from the study showed that the scale adapted to Turkish was sufficient in terms of quantity and quality to evaluate the quality of life of individuals who had undergone surgery for breast cancer and that the scale met the desired criteria for content validity.

Construct Validity

To assess the construct validity the number of samples should be sufficient, and data should be collected from participants 5 or 10 times the number of items in the scale to ensure construct validity[11]. KMO test is performed to evaluate sample fitness. Bartlett's test is performed to assess the normality of the data, and its chi-square result should be statistically significant. In the study, the KMO value was calculated as 0.77, and it was determined that the sample size was "good" for factor analysis and the data were homogeneously distributed. Bartlett's value was calculated as $\chi^2=5055.721$; $p<0.001$ and it was determined that there was sufficient correlation between the items for construct validity analysis.

Table 4. Descriptive statistics of the scale and Cronbach's alpha (α) and omega coefficient (ω) reliability analysis (n=250)

Subdimensions	Items	Mean ± SD	Corrected item-total correlation	Cronbach's alpha if item deleted	Std Cronbach's alpha	Omega coefficient (ω)
Physical activity	A1	1.14±0.69	0.38	0.81	0.81	0.87
	A2	1.38±0.62	0.67	0.76		
	A3	1.41±0.60	0.62	0.77		
	A4	1.66±0.63	0.54	0.78		
	A5	1.50±0.6	0.66	0.76		
	A6	1.22±0.61	0.39	0.80		
	A7	1.21±0.50	0.57	0.78		
	A8	0.87±0.55	0.35	0.81		
Pain	P9	1.55±0.85	0.70	0.95	0.95	0.97
	P10	1.47±0.98	0.81	0.94		
	P11	1.60±1.23	0.91	0.93		
	P12	1.67±1.0	0.84	0.94		
	P13	1.62± 1.16	0.78	0.94		
	P14	1.45±1.10	0.78	0.94		
	P15	1.73±1.13	0.90	0.93		
Feelings	F16	1.35±0.94	0.72	0.89	0.90	0.95
	F17	1.29±1.00	0.69	0.89		
	F18	1.18±1.04	0.84	0.87		
	F19	1.17±1.07	0.73	0.89		
	F20	1.08±0.92	0.72	0.89		
	F21	1.13±1.02	0.70	0.89		
Body image	B22	1.78±0.69	0.56	0.78	0.79	1.00
	B23	1.65±0.64	0.69	0.65		
	B24	1.59±0.62	0.64	0.70		
Physical health	H25	0.94±0.25	0.64	0.89	0.90	1.00
	H26	0.96±0.21	0.77	0.87		
	H27	0.95±0.22	0.78	0.87		
	H28	0.94±0.25	0.68	0.88		
	H29	0.95±0.22	0.76	0.87		
	H30	0.92±0.27	0.69	0.88		
	H31	0.93±0.25	0.59	0.89		
Sexual functioning	S32	0.93±0.25	0.93	0.81	0.90	1.00
	S33	0.90±0.30	0.87	0.83		
	S34	0.92±0.27	0.78	0.87		
	S35	0.87±0.34	0.54	0.85		
General health	G36	2.05±0.90	0.42	0.87	0.86	1.00
	G37	1.40±0.66	0.85	0.80		
	G38	1.39±0.69	0.74	0.82		
	G39	1.36±0.71	0.70	0.83		
	G40	1.51±0.67	0.73	0.82		
	G41	1.26±0.79	0.65	0.83		
	G42	0.68±0.48	0.32	0.88		
Relationship with other	R43	1.60±0.90	0.52	0.79	0.81	1.00
	R44	2.14±0.83	0.59	0.77		
	R45	2.20±0.82	0.58	0.77		
	R46	1.81±0.95	0.51	0.79		
	R47	1.89±0.94	0.51	0.79		
	R48	2.06±0.84	0.53	0.78		
	R49	1.94±0.86	0.54	0.78		
Total Cronbach's alpha					0.79	

SD: Standart deviation, Std: Standardized

Table 5. Test-retest statistical analysis results of the scale (n=75)

Subdimensions	Test Mean ± SD	Re-test Mean ± SD	Test value	p	r	p
Physical activity	11.51±2.38	11.47±2.61	0.115[t]	0.909	0.268[p]	0.020
Pain	6.75±4.22	6.81±4.07	-0.962[t]	0.339	0.990[p]	p<0.001
Feelings	6.21±3.88	6.27±3.88	-1.424[t]	0.159	0.997[p]	p<0.001
Body image	5.73±1.29	5.77±1.26	-1.000[t]	0.321	0.963[p]	p<0.001
Physical health	6.52±0.95	6.43±1.04	-1.097[w]	0.273	0.700[s]	p<0.001
Sexual functioning	3.56±0.93	3.57±0.93	-0.577[w]	0.564	0.971[s]	p<0.001
General health	8.88±4.05	8.80±4.00	1.621[t]	0.109	0.994[p]	p<0.001
Relationship with other	15.28±1.80	15.31±1.82	-0.497[t]	0.620	0.967[p]	p<0.001

[t]: Paired-samples t-test, [w]: Wilcoxon, [p]: Pearson correlation coefficient, [s]: Sperman correlation coefficient, SD: Standart deviation

CFA was used for construct validity in its adaptation to Turkish. CFA is performed to investigate the fit of an existing scale or model in a new data set, in other words, to test whether the factor structure is verified[18]. CFA was not performed at the original scale. However, model fit indices, which were not examined in the original article, were examined in our study. The fit of the data to the model is tested using chi-square fit statistics. If the Chi-square value (χ^2/df =1.064) of the BCSQOL scale is less than 2, this indicates that the model has an acceptable goodness of fit. In this scale adaptation study, the RMSEA score was below 0.05, indicating a perfect fit (Table 3). Among the other fit indices obtained as a result of CFA; CFI, AGFI, NFI, TLI, and IFI values were found to be a perfect fit of the model, while GFI and SRMR values were found to be an acceptable fit. In the study, the goodness of fit indices of the scale was found at the desired level, generally showing an excellent fit.

Reliability

Reliability is the power to obtain the same results when the scale measures the concept or dimension repeatedly[11]. The reliability of the scale adapted in the study was evaluated with test-retest reliability, Cronbach's alpha, and omega coefficients. Test-retest reliability is measured by comparing the results obtained from the same participants of the same scale at different times. In this study, the scale was re-applied to 75 individuals with an interval of 2 weeks to evaluate the test-retest reliability of the scale, and a significant positive correlation was found between them (p<0.001).

However, the test-retest correlation for the physical activity subdimensions was relatively low (r=0.268), indicating a weak positive relationship between the test and retest scores. Although this correlation was statistically significant, several factors may explain the low stability. The time interval between the test and retest could have contributed to genuine changes in participants' physical activity levels. Moreover, individual differences in postoperative recovery processes, including variations in physical limitations and pain levels, may have influenced physical activity behaviors during the study period. Psychosocial factors, particularly emotional states such as depression and anxiety, might also have limited participants' physical activity levels. These findings suggest that the low correlation reflects the natural variability in physical activity during recovery rather than a limitation of the scale itself. Future studies may benefit from exploring these factors in more detail to better understand their impact on the stability of physical activity measurements.

"Item-total score correlation analysis" is applied to determine how much the items in the scale are related to the measured theoretical structure following reliability analysis. In the literature, it is stated that the item-total score correlation coefficient values should be positive and above 0.30[19]. In the study, the values in which the item-total score correlation coefficients ranged from 0.32 to 0.93 indicate that the items in the scale are a reliable measurement tool for assessing the quality of life of individuals who had undergone surgery for breast cancer.

Cronbach's alpha coefficient measures the correlations between the items of the scale and evaluates the internal consistency of the scale. In scales with multifactorial items, the omega reliability coefficient is more recommended than the Cronbach's alpha value[20]. Like Cronbach's alpha coefficient, the omega coefficient measures the correlations between the items of the scale and gives information about the consistency[20]. In the study, the Cronbach's alpha coefficient of the subdimensions of the scale ranged from 0.79 to 0.95; the omega coefficient was found between 0.87 and 1.00. In addition, all corrected item correlation values for all subdimensions of the scale were positive as a result of the reliability analysis. There was no significant increase in the reliability coefficients for all eight subdimensions when the item was removed from the subdimensions. Finally, when

Cronbach alpha and omega coefficients of all subdimensions of the scale are evaluated, it can be said that the scale has "highly reliable" internal consistency.

The adaptation of the BCSQOL scale into Turkish has important clinical and academic implications. Clinically, having a culturally and language validity appropriate tool enables healthcare professionals to more accurately assess the specific physical and psychosocial challenges faced by patients undergoing breast cancer surgery in Türkiye. This, in turn, facilitates the development of individualized care plans and targeted interventions aimed at improving patients' quality of life during the postoperative period. Academically, the adapted scale provides researchers with a valid and reliable instrument to investigate quality of life outcomes in this population, supporting future studies and contributing to the national and international literature on breast cancer survivorship. Furthermore, the availability of this scale in Turkish may promote multicenter or cross-cultural research collaborations focused on enhancing the well-being of breast cancer patients.

Study Limitations

The limitations of the study include the failure to evaluate the quality of life of male patients who underwent surgery for breast cancer, as the sample consisted exclusively of female patients. Additionally, the inability to generalize the research findings is another limitation, as the data were collected from a single center. This single-center design may limit the representativeness of the sample, as it may not fully capture the diversity of patient populations in different geographic locations or healthcare settings. Therefore, the results may not be applicable to broader or more heterogeneous populations, and caution should be taken when attempting to apply the findings outside the context of the study. Future research with multi-center designs is needed to enhance the generalizability and external validity of the results.

CONCLUSION

As a result of the psychometric analysis, it was determined that the BCSQOL scale, consisting of 49 items and 8 subdimensions, adapted to Turkish, has adequate psychometric properties. Only a Turkish adaptation of the scale was carried out in this research. An adapted scale may be used in clinical settings and academic studies to evaluate the postoperative quality of life of women who had surgery for breast cancer.

This adapted scale has been specifically designed to assess the quality of life of patients undergoing breast cancer surgery. Using this scale enables healthcare professionals, particularly nurses, to gain a more accurate and comprehensive assessing of these patients' quality of life in the early postoperative period. This includes identifying factors such as pain, body image, physical activity and emotional distress that directly impact recovery and well-being. Furthermore, the obtained data can be used to plan interventions aimed at improving patients' quality of life. This facilitates the implementation of care strategies tailored to both physical and psychosocial needs. Consequently, using this scale enhances the quality of patient care, helps to manage the challenges encountered during recovery more effectively, and ultimately improves patient satisfaction.

Ethics

Ethics Committee Approval: Ethical approval for the study University of Health Sciences Türkiye Gülhane Faculty of Medicine Clinical Research Ethics Committee (decision no: 2022/36, date: 25.10.2022).

Informed Consent: Prior to data collection, all participants were informed about the purpose and procedures of the study, and both written and verbal informed consent were obtained.

Acknowledgment

We would like to thank all the patients who agreed to participate in the study.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Park JH, Jung YS, Kim JY, Bae SH. Trajectories of quality of life in breast cancer survivors during the first year after treatment: a longitudinal study. BMC Womens Health. 2023;23:12.
2. Yanardağ CH, Çürük GN, Karayurt Ö. Effects of selfcare and selected factors on the quality of life in women with breast cancer-related lymphedema. Support Care Cancer. 2023;31:1-7.
3. Ye XX, Ren ZY, Vafaei S, Zhang JM, Song Y, Wang YX, et al. Effectiveness of baduanjin exercise on quality of life and psychological health in postoperative patients with breast cancer: a systematic review and meta-analysis. Integr Cancer Ther. 2022;21:1-13.
4. Sel Z, Beydağ KD. The relationship between body perception, sexual quality of life and marital adjustment in women who underwent surgical operation for breast cancer. BANU Journal of Health Science and Research. 2022;4:119-27.
5. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71:209-49.

6. Sun FK, Lu CY, Yao Y, Chiang CY. Social functioning, depression, and quality of life among breast cancer patients: a path analysis. Eur J Oncol Nurs. 2023;62:102237.

7. Akezaki Y, Nakata E, Kikuuchi M, Tominaga R, Kurokawa H, Okamoto M, et al. Investigation of factors affecting early quality of life of patients after breast cancer surgery. Healthcare (Basel). 2021;9:213.

8. Aydin Özcan D, Uyar S, Kılıçoğlu GS, Kaya B, Yakut Y. Validity and reliability of the Turkish version of the breast cancer treatment outcome scale. JETR. 2021;8:187-9.

9. Yalçın Ö, Gürkan A. Validity and reliability of the functional evaluation of cancer treatment breast scale: a methodological study. Turkiye Klinikleri J Nurs Sci. 2022;14:974-82.

10. El Farrah, K. Breast cancer surgery QOL scale: development, validity, and reliability (master's thesis). McGill University, Montreal, Quebec, Canada, 2003.

11. Esin NM. Veri toplama yöntem ve araçları ve veri toplama araçlarının güvenirlik ve geçerliği. Nobel Tıp Kitabevleri; 2014;193-233.

12. R Core Team. (2021). R: A language and environment for statistical computing. Vienna, Austria: r Foundation for Statistical Computing. https://www.r-project.org

13. George D, Mallery M. SPSS for Windows step by step: a simple guide and reference, 17.0 update. Boston: Pearson; 2010.

14. Rosseel Y. lavaan: an R package for structural equation modeling. Journal of Statistical Software. 2012;48:1-36.

15. Davis LL. Instrument review: getting the most from a panel of experts. Appl Nurs Res. 1992;5:194-7.

16. Yeşil R. Nitel ve nicel araştırma yöntemleri. Bilimsel Araştırma Yöntemleri. Nobel Yayınevi, 2014.

17. Alpar R. Uygulamalı istatistik ve geçerlilik-güvenirlik. Detay Yayıncılık, 2018.

18. Evci N, Aylar F. Use of confirmatory factor analysis in scale development studies. The Journal of Social Science. 2017;10:389-412.

19. Özdamar K. Ölçek ve test geliştirme yapısal eşitlik modellemesi. Nisan Kitabevi, 2016.

20. Kalkbrenner MT. Alpha, omega, and H internal consistency reliability estimates: reviewing these options and when to use them. Couns Outcome Res Eval. 2021;14:77-88.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.81804
Nam Kem Med J 2025;13(3):333-340

Unveiling Subclinical Arrhythmias After COVID-19: Insights from 24-Hour Holter Monitoring

COVID-19 Sonrası Subklinik Aritmilerin Ortaya Çıkışı: 24 Saatlik Holter Monitörizasyonundan Elde Edilen Bulgular

Çağlar KAYA[1], Mustafa EBİK[2], Fatih KARDAŞ[1], Nicholas KOUNIS[3], Servet ALTAY[1]

[1]*Trakya University Faculty of Medicine, Department of Cardiology, Edirne, Türkiye*
[2]*Sultan 1. Murat State Hospital, Clinic of Cardiology, Edirne, Türkiye*
[3]*School of Medicine University of Patras, Department of Cardiology, Patras, Greece*

ABSTRACT

Aim: Coronavirus disease 2019 (COVID-19) has been linked to various arrhythmias, including atrial and ventricular ectopy, atrial fibrillation, and conduction disturbances. Proposed mechanisms include myocardial injury, systemic inflammation, autonomic dysfunction, and endothelial damage. This study aimed to assess arrhythmic burden and its predictors in patients with prior COVID-19 infection using 24-hour Holter monitoring.

Materials and Methods: We retrospectively analyzed 153 patients who underwent Holter electrocardiography (ECG) between January 2021 and June 2023. Participants were divided into COVID-19 positive (n=62) and control (n=91) groups. Demographic, clinical, and laboratory characteristics were compared. Ventricular ectopic beats (VES) were quantified, and ROC analysis identified 571 VES/day as the threshold for high arrhythmic burden. Patients were stratified accordingly, and predictors of high burden were determined using multivariate logistic regression.

Results: VES counts were significantly higher in the COVID positive group (p<0.001). Patients with >571 VES had a higher frequency of fragmented QRS (75.0% vs. 7.6%) and prior COVID-19 infection (60.4% vs. 31.4%) (p<0.001 for both). In multivariate analysis, fragmented QRS [odds ratio (OR): 26.99, p<0.001] and COVID-19 history (OR: 10.30, p<0.001) independently predicted high arrhythmic burden. COVID-19 vaccination was associated with a reduced risk (OR: 0.13, p=0.006).

Conclusion: COVID-19 infection is significantly associated with increased ventricular ectopy and fragmented QRS on Holter ECG, indicating persistent arrhythmic risk. Fragmented QRS and COVID-19 history independently predict high arrhythmic burden, while vaccination appears protective. Post-COVID rhythm surveillance may help identify high-risk individuals.

Keywords: COVID-19, cardiovascular arrhythmia, ventricular extra systole

ÖZ

Amaç: Koronavirüs hastalığı-2019 (COVID-19), atriyal ve ventriküler ektopi, atriyal fibrilasyon ve ileti bozuklukları dahil olmak üzere çeşitli aritmilerle ilişkilendirilmiştir. Öne sürülen mekanizmalar arasında miyokardiyal hasar, sistemik enflamasyon, otonom disfonksiyon ve endotel hasarı yer almaktadır. Bu çalışma, COVID-19 enfeksiyonu öyküsü olan hastalarda 24 saatlik Holter monitörizasyonu kullanılarak aritmi yükünü ve bunun belirleyicilerini değerlendirmeyi amaçlamıştır.

Gereç ve Yöntem: Ocak 2021 ile Haziran 2023 tarihleri arasında Holter elektrokardiyografi (EKG) uygulanmış 153 hasta retrospektif olarak analiz edildi. Katılımcılar COVID-19 pozitif (COVID+) (n=62) ve kontrol (n=91) gruplarına ayrıldı. Demografik, klinik ve laboratuvar özellikleri karşılaştırıldı. Ventriküler ektopik atımlar (VES) sayıldı ve ROC analiziyle günlük 571 VES değeri yüksek aritmi yükü için eşik olarak belirlendi. Hastalar bu değere göre sınıflandırıldı ve yüksek aritmi yükünün öngörücüsü olan değişkenler çok değişkenli lojistik regresyonla analiz edildi.

Address for Correspondence: Çağlar KAYA MD, Trakya University Faculty of Medicine, Department of Cardiology, Edirne, Türkiye
E-mail: caglarkaya2626@gmail.com **ORCID ID:** orcid.org/0000-0002-2968-5352
Received: 06.05.2025 **Accepted:** 14.07.2025 **Publication Date:** 07.10.2025

Cite this article as: Kaya Ç, Ebik M, Kardaş F, Kounis N, Altay S. Unveiling subclinical arrhythmias after COVID-19: insights from 24-hour holter monitoring. Nam Kem Med J. 2025;13(3):333-340

Bulgular: VES sayıları COVID pozitif grubunda anlamlı derecede yüksekti (p<0,001). Günde 571'den fazla VES olan hastalarda, fragmented QRS (%75,0 vs. %7,6) ve COVID-19 öyküsü (%60,4 vs. %31,4) daha sık görüldü (her ikisi için p<0,001). Çok değişkenli analizde, fragmented QRS [olasılık oranı (OR): 26,99, p<0,001] ve COVID-19 öyküsü (OR: 10,30, p<0,001), bağımsız olarak yüksek aritmi yükünü öngördü. COVID-19 aşılaması, daha düşük risk ile ilişkilendirildi (OR: 0,13, p=0,006).

Sonuç: COVID-19 enfeksiyonu, Holter EKG'de artmış ventriküler ektopi ve fragmented QRS ile anlamlı şekilde ilişkilidir ve bu durum kalıcı aritmik riske işaret eder. Fragmented QRS ve COVID-19 öyküsü yüksek aritmi yükünün bağımsız belirleyicileridir; aşılanma ise koruyucu bir etkiye sahip olabilir. Post-COVID döneminde ritim izlemi, yüksek riskli bireylerin belirlenmesine katkı sağlayabilir.

Anahtar Kelimeler: COVID-19, kardiyovasküler aritmi, ventriküler ekstra sistol

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 infection has been associated with a wide range of cardiovascular complications, including arrhythmias[1]. Post-Coronavirus disease (COVID) cardiac manifestations are increasingly recognized, with ventricular arrhythmias being one of the concerning sequelae[2]. The potential mechanisms underlying this increased arrhythmic burden include direct myocardial injury, systemic inflammation, and autonomic dysfunction. Emerging evidence also suggests that survivors of Coronavirus disease 2019 (COVID-19) may develop autonomic nervous system disturbances for instance, persistent sympathetic activation and reduced heart rate variability have been observed – which could facilitate post-COVID arrhythmogenesis[3]. Identifying patients at higher risk of developing significant arrhythmias may help in optimizing post-COVID-19 cardiac care[4]. However, data regarding the long-term ventricular arrhythmic burden in relatively young, non-hospitalized post-COVID populations remain limited. This study aimed to evaluate the arrhythmic burden in patients with a history of COVID-19 infection and determine the clinical predictors of increased ventricular ectopic activity (VEA).

MATERIALS AND METHOD

This retrospective study included 153 patients who underwent 24-hour electrocardiography (ECG) Holter monitoring, classified into two groups: COVID-19 positive (COVID+) (n: 62) and control (n: 91). Between January 2021 and January 2022, patients were selected for inclusion if they had a confirmed COVID-19 diagnosis by positive reverse transcription polymerase chain reaction and subsequently presented with cardiac symptoms such as palpitations, presyncope, or chest pain warranting Holter monitoring. Symptoms such as presyncope, syncope, palpitations were recorded. None of the patients had complete syncope. Patients in the control group also underwent Holter monitoring for similar cardiac symptoms such as palpitations, syncope, or chest pain, despite having no history of COVID-19. Exclusion criteria comprised any history of structural heart disease, pacemaker implantation, or severe electrolyte imbalances, to avoid confounding factors that could independently affect arrhythmia propensity. Patients' vaccination history was recorded, however, due to the heterogeneity in vaccine types (inactivated and mRNA-based) and variable dose numbers (ranging from one to four doses), vaccination status was analyzed categorically as vaccinated or unvaccinated. Baseline clinical characteristics, laboratory parameters, and ECG Holter recordings were compared between groups. A ROC curve analysis was performed to determine the optimal threshold for high arrhythmic burden based on ventricular extra systole (VES) counts. A cut-off value of 571 VES was identified. Patients were then further categorized into two subgroups: VES >571 (n: 48) and VES ≤571 (n: 105). Univariate and multivariate logistic regression analyses were conducted to identify independent predictors of high VES burden. The study was carried out in compliance with the Declaration of Helsinki and received approval from the Institutional Committee on Human Research and Ethics. All participants provided written informed consent before enrollment. Ethical approval for this research was granted by The Ethics Committee for Scientific Research, Faculty of Medicine, Trakya University (decision no: 02/17, date: 18.12.2021).

Statistical Analysis

Statistical analyses were conducted using SPSS version 25.0 (SPSS, Chicago, IL). Continuous variables were summarized as mean ± standard deviation, while categorical variables were reported as frequencies and percentages. The Shapiro-Wilk test was applied to evaluate the normality of distributions. Group comparisons for continuous variables were performed using the independent t-test, whereas categorical variables were analyzed with the Pearson's chi-square test. Correlations were evaluated with spearman correlation test. ROC curve analysis was utilized to determine the area under the curve (AUC) and optimal cut-off values for diagnostic performance. Logistic regression analysis was employed to identify independent predictors. All statistical tests were two-sided, with a significance level set at p<0.05.

RESULTS

Baseline clinical and demographic characteristics are presented in Table 1. mean age was similar between the COVID+ and control groups (38.57±15.14 vs. 37.12±16.03 years, p: 0.903). Female patients were more prevalent in the COVID+ group (41.93% vs. 25.27%, p: 0.030). Thyroid dysfunction was significantly higher in the COVID+ group (24.19% vs. 7.69%, p: 0.030), while rates of hypertension (43.54% vs. 43.95%, p: 0.848) and diabetes mellitus (37.09% vs. 29.67%, p: 0.336) showed no significant differences (Table 1). No difference was found in terms of the medication (antihypertensive or antiarrhythmic) used by the patients (Table 1). COVID+ patients had significantly elevated high-sensitivity-C-reactive protein (hs-CRP) levels (2.47±3.33 mg/dL vs. 1.46±1.95 mg/dL, p: 0.002) and a higher incidence of fragmented QRS (46.77% vs. 10.98%, p<0.001). The COVID+ group also exhibited a significantly higher burden of VEA (VES count, p<0.001). The analysis of ECG Holter parameters is presented in Table 2. ROC curve analysis identified 571 VES as the optimal threshold for high arrhythmic burden (AUC: 0.82, 95% confidence interval: 0.76-0.88, p<0.001) (Table 3).

Figure 1. Patients with VES >571 had a higher prevalence of fragmented QRS (75.0% vs. 7.62%, p<0.001) and prior COVID-19 infection (60.41% vs. 31.42%, p<0.001). VEA, age, ejection fraction (EF) and hs-CRP were evaluated by correlation analyses (Figure 2). A weak positive correlation was found between VEA and advanced glycation end-products. There is a negative, weak and statistically significant correlation between VEA and EF (r=-0.2313, p=0.004). There was a negative, moderate and statistically significant correlation between CRP and EF variables (Figure 2). Vaccination rates were significantly lower in the high VES group (52.08% vs. 90.47%, p<0.001). Elevated hs-CRP levels correlated with high VES burden (p: 0.002). Although the EF was slightly lower in the high VES group (54.45% vs. 55.58%, p: 0.314), the difference was not statistically significant.

Table 1. The Clinical characteristics of patients of the study population in comparison with COVID-19 and Control group

Variables	COVID+ (n: 62)	Control (n: 91)	p
Age, years	38.57±15.14	37.12±16.03	0.903
Sex, female, n (%)	26 (41.93)	23 (25.27)	0.030
BMI kg/m^2	25.34±5.11	25.93±3.69	0.796
Smoker (no) n (%)	25 (40.32)	33 (36.26)	0.075
HT n (%)	27 (43.54)	40 (43.95)	0.848
DM n (%)	23 (37.09)	27 (29.67)	0.336
CAD n (%)	14 (22.58)	21 (23.07)	0.426
Thyroid dysfunction	15 (24.19)	7 (7.69)	0.03
Vaccination	39 (62.90)	81 (71.37)	<0.001
Neutrophil/lymphocyte ratio	4.02±2.65	3.62±1.99	0.302
Hemoglobin (g/dL)	11.05±1.65	11.69±1.72	0.962
Na (mmol/L)	138.51±2.77	139.75±5.13	0.084
K (mmol/L)	4.60±0.59	4.38±0.81	0.074
hs-CRP (mg/dL)	2.47±3.33	1.46±1.95	0.002
Creatinine (mg/dL)	0.93±0.29	0.85±0.59	0.258
Frag QRS	29 (46.77)	10 (10.98)	<0.001
Ejection fraction (%)	54.11±8.87	55.82±5.18	0.135
Presyncope (%)	3 (3.29)	3 (4.83)	0.630
Syncope	0	0	-
Palpitation	62 (100)	91 (100)	-
Hospitalization	0	0	-
Chest pain	7 (11.29)	11(12.08)	0.721
ACEi	21(33.87)	30 (32.96)	0.671
Beta blockers	10 (16.12)	17 (18.68)	0.272
CCB-dihydropyridine Non dihydropyridine	10 (16.12) 2 (3.22)	16 (17.58) 4 (4.39)	0.568 0.128

ACEi: Angiotensin converting enzyme inhibitor, CAD: Coronary artery disease, CCB: Ca canal blockers, DM: Diabetes mellitus, HT: Hypertension, hs-CRP: High-sensitive C-reactive protein, BMI: Body-mass index, Na: Sodium, K: Potassium, COVID-19: Coronavirus disease-19

Table 2. The ECG Holter parameters of patients of the study population in comparison with COVID-19 and Control group

Variables	COVID+ (n: 62)	Control (n: 91)	p
Atrial extra systole	103 (31-401)	80 (30-189)	0.001
Ventricular extra systole	796 (361-950)	110 (65-180)	<0.001
Non-sustained ventricular tachycardia	3 (4.83)	4 (4.39)	0.758
Supraventricular tachycardia	11 (17.74)	19 (19.78)	0.129
QRS (msn)	93.45±18.04	91.82±12.41	0.547
QT (msn)	373.67±35.77	391.54±42.44	0.012
QTc	441.14±24.05	430.47±24.50	0.020
QTd	52.58±16.68	40.87±16.78	<0.001
QTcd	50.50±11.85	46.29±17.45	0.003
QRS (msn)	93.82±18.51	90.89±12.04	0.641
MQTc	437.09±22.02	427±25.53	0.047

COVID-19: Coronavirus disease-19, ECG: Electrocardiography, COVID+: COVID-19 pozitif

Table 3. Diagnostic performance of VES for group of COVID+

Diagnostic Performance of VES for COVID+					
	Cut-off	AUC (95% CI)	Sensitivity	Specificity	p-value
VES	571	0.82 (0.76-0.88)	0.77 (0.65-0.86)	0.91 (0.83-0.95)	<0.001

VES: Ventricular extra systole, AUC: Area under the curve, CI: Confidence interval, COVID: Coronavirus disease

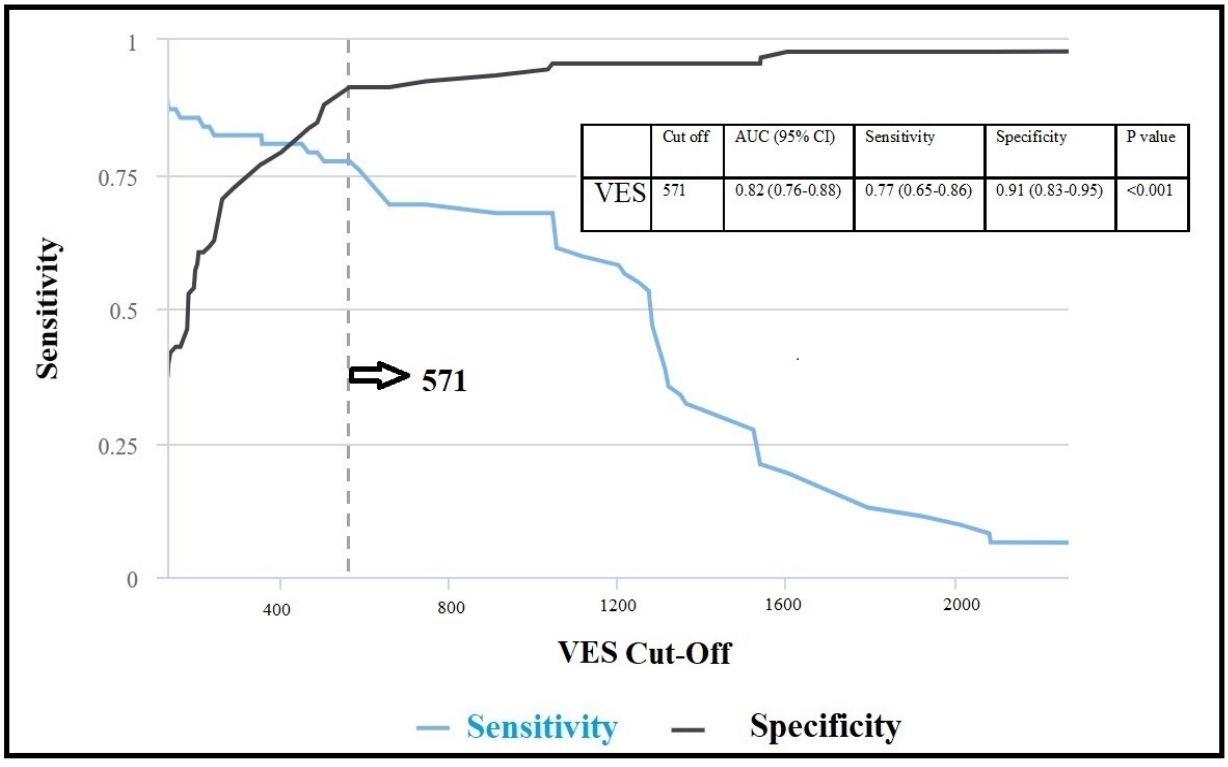


Figure 1. Sensitivity and specificity chart of VES

VES: Ventricular extra systole, AUC: Area under the curve, CI: Confidence interval

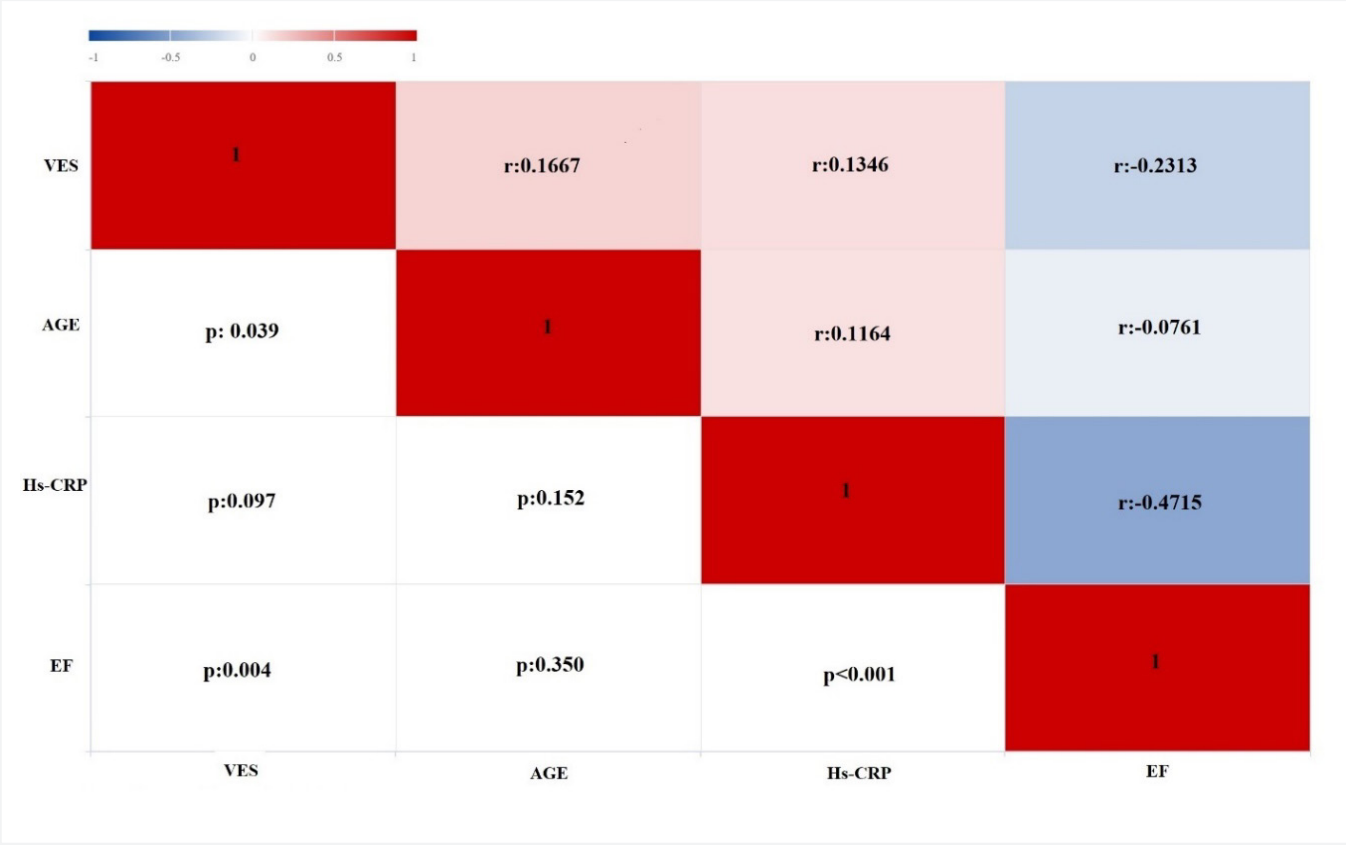


Figure 2. Correlogram table constructed as a result of correlation analyses

VES: Ventricular extra systole, AGE: Advanced glycation end-products, hs-CRP: High-sensitivity-C-reactive protein, EF: Ejection fraction

Table 4. The clinical characteristics of patients of the study population in comparison with VES count (571)

Değişkenler	VES >571 (n: 48)	VES <571 (n: 105)	p
Age	39.06±14.18	38.15±15.89	0.753
Gender, female n (%)	23 (47.92)	26 (24.76)	0.004
Smoker (no) n (%)	20 (41.66)	38 (36.19)	0.135
HT n (%)	17 (35.41)	36 (34.28)	0.297
DM n (%)	23 (37.09)	27 (29.67)	0.336
CAD n (%)	8 (16.66)	24 (22.85)	0.385
Thyroid dysfunction	10 (20.83)	9 (8.57)	0.041
COVID-19	29 (60.41)	33(31.42)	<0.001
Vaccination	25 (52.08)	95 (90.47)	<0.001
NLR	4.15±2.99	3.82±2.05	0.360
hs-CRP	2.17±2.75	1.73±2.58	0.343
Fragmented QRS	36 (75.00)	8 (7.62)	<0.001
Ejection fraction (%)	54.45±9.87	55.58±6.18	0.314

CAD: Coronary artery disease, DM: Diabetes mellitus, HT: Hypertension, hs-CRP: High-sensitive C-reactive protein, NLR: Neutrophil/lymphocyte ratio, COVID-19: Coronavirus disease-19, VES: Ventricular extra systole

Table 5. Univariate, multivariate and stepwise binary logistic regression analysis for the VEA >571

Log reg	Univariate model			Multivariate model		
	OR	95% CI	p	OR	95% CI	p
Fragmented QRS	36.37	13.74-96.23	<0.001	26.99	6.98-99.73	<0.001
Gender (female)	0.78	0.35-2.40	0.208	0.264	0.05-1.28	0.098
Thyroid dysfunction	2.80	1.05-7.44	0.038	0.578	0.05-5.44	0.632
COVID+	29.75	11.17-79.61	<0.001	10.30	3.07-34.51	<0.001
Vaccination	0.11	0.05-0.27	<0.001	0.13	0.03-0.57	0.006
Age	1.158	0.91-1.25	0.428	1.01	0.97-1.05	0.595

OR: Odds ratio, CI: Confidence interval, COVID: Coronavirus disease, VEA: Ventricular ectopic activity, COVID: Coronavirus disease, COVID+: COVID-19 pozitif

The clinical characteristics of the study population in comparison with VEA count (571) are presented in Table 4. ECG Holter analysis showed that COVID+ patients had a higher burden of atrial extrasystoles (987.22±208.85 vs. 463.55±221.33, p: 0.042), prolonged QTc intervals (430.47±24.50 ms vs. 441.14±24.05 ms, p: 0.020), and increased QT dispersion (40.87±6.78 ms vs. 52.58±16.68 ms, p<0.001) compared to controls. Multivariate logistic regression analysis identified fragmented QRS [odds ratio (OR): 26.99, p<0.001] and COVID-19 infection (OR: 10.30, p<0.001) as independent predictors of high VES burden (Table 5). Vaccination was associated with a significantly lower risk of arrhythmia (OR: 0.13, p: 0.006).

DISCUSSION

COVID-19 infection has been implicated not only in acute cardiac injury but also in long-term cardiac sequelae, including inflammatory and arrhythmogenic manifestations. In a previously published case, we reported the development of constrictive pericarditis in a young individual following COVID-19 and vaccination exposure, raising awareness of the virus's potential to induce delayed cardiac inflammation and remodeling, even in structurally normal hearts[5]. COVID-19 infection has been linked to various cardiac arrhythmias, including atrial and ventricular ectopy, atrial fibrillation, and conduction abnormalities. Proposed mechanisms include direct myocardial injury, systemic inflammation, autonomic dysfunction, and endothelial damage, all contributing to altered cardiac electrophysiology. Recent studies indicate that post-COVID-19 patients, especially those with severe illness, have an increased risk of ventricular arrhythmias and QT prolongation, highlighting the need for long-term cardiac surveillance[6]. In a recently published review, the relationship between COVID-19 and the cardiovascular system has been thoroughly evaluated. This study discusses the mechanisms by which COVID-19 induces arrhythmias, which are associated with hypoxia, myocarditis, and secondary causes. Specifically, it highlights that hypoxia can trigger anaerobic glycolysis through cellular damage, increase cytosolic calcium levels, and facilitate arrhythmogenesis by inducing early and late depolarizations[7]. Unlike prior studies focusing on hospitalized COVID-19 patients, our study uniquely highlights the persistent arrhythmic risk even in a relatively young outpatient population.

The findings of this study highlight three key results: 1. prior COVID-19 infection is significantly associated with increased VEA and fragmented QRS, suggesting a potential link between myocardial injury and arrhythmic risk. 2. hs-CRP levels were significantly elevated in patients with high arrhythmic burden, reinforcing the role of systemic inflammation in post-COVID arrhythmogenesis. 3. vaccination was associated with a significantly lower risk of high VES burden, indicating a potential protective effect against post-COVID cardiovascular complications.

The association between COVID-19 infection and increased ventricular arrhythmic burden has been supported by previous studies. For example, Turagam et al.[8] demonstrated that post-COVID patients, particularly those with severe disease, exhibited a higher frequency of ventricular ectopy. The f-QRS pattern may have significant prognostic implications in COVID-19 patients. There are publications in the literature suggesting that this condition may exhibit temporal and mechanistic variations, and our study supports these findings[9]. Our findings align with these studies, suggesting that myocardial injury and fibrosis, as reflected by fragmented QRS patterns, may contribute to sustained arrhythmic risk[10]. The presence of fragmented QRS has been well-documented as a marker of myocardial scarring, which predisposes patients to malignant arrhythmia.

However, it should be noted that truly malignant ventricular arrhythmias (such as sustained ventricular tachycardia or fibrillation) appear to be relatively infrequent in post-COVID patients without critical illness. In fact, clinical data indicate that ventricular tachyarrhythmias in COVID-19 are mainly observed in the presence of severe metabolic derangements suggesting that profound electrolyte imbalances or other

metabolic factors during acute illness are often required to precipitate life-threatening arrhythmias[11].

Elevated hs-CRP levels in patients with high VES burden further emphasize the role of systemic inflammation in arrhythmogenesis. Prior studies have shown that post-COVID inflammatory responses, including cytokine-mediated myocardial stress, can alter cardiac electrophysiology. In line with this, Marques et al.[12] reported that elevated inflammatory markers, particularly CRP and interleukin-6, were associated with an increased risk of ventricular arrhythmias post-COVID[13]. Our results support this concept, as heightened systemic inflammation may exacerbate myocardial excitability, leading to increased ectopic activity. The association between autonomic dysfunction and inflammation in long-COVID patients, as indicated by elevated CRP and impaired heart rate recovery, aligns with our findings of increased VEA and prolonged QT parameters in post-COVID individuals[14]. These results further support the hypothesis that persistent systemic inflammation and autonomic imbalance may contribute to heightened arrhythmic risk in COVID-19 survivors, underscoring the need for long-term cardiac monitoring

The observed protective effect of vaccination against high arrhythmic burden is an important finding. Previous research has demonstrated that COVID-19 vaccination reduces the severity of systemic inflammatory responses, which are known contributors to arrhythmia development. Studies such as Pari et al.[15] reported a lower incidence of post-COVID cardiovascular complications among vaccinated individuals, likely due to the mitigation of endothelial dysfunction and myocardial inflammation. Our findings reinforce these observations, suggesting that vaccination may play a role in reducing post-infectious arrhythmic risk.

Beyond these primary results, certain findings warrant further discussion. The significantly prolonged QTc interval and increased QT dispersion in the COVID+ group suggest potential autonomic dysregulation or direct myocardial electrophysiological alterations post-infection[16]. In addition to autonomic dysregulation, potential contributors to QTc prolongation may include subclinical electrolyte imbalances or the use of QT-prolonging medications, although major disturbances were part of our exclusion criteria. These parameters are known predictors of ventricular arrhythmias and sudden cardiac death, necessitating further research on long-term QT dynamics in post-COVID populations[17]. Additionally, the high prevalence of thyroid dysfunction in the COVID+ group raises questions regarding the interplay between endocrine abnormalities and arrhythmic risk, as thyroid dysfunction is a known contributor to ventricular ectopy. Although thyroid dysfunction was more prevalent in the COVID+ group, this association may be incidental, and no direct causal link can be established based on our data. Future prospective studies should investigate whether these abnormalities persist over time and their potential therapeutic implications in post-COVID cardiac care[18]. The observed increase in atrial extrasystoles, prolonged QTc intervals, and greater QT dispersion in the COVID+ group suggests a potential impact of COVID-19 on cardiac electrophysiology. QT dispersion, a known marker of ventricular repolarization heterogeneity, has been associated with an increased risk of malignant arrhythmias and sudden cardiac death in various clinical settings. Previous studies have reported that prolonged QT dispersion in post-viral myocarditis and systemic inflammatory states may contribute to long-term arrhythmic complications, emphasizing the need for close monitoring in post-COVID-19 patients[19].

Correlation analysis further supported these findings by revealing a weak but significant inverse relationship between VES burden and left ventricular EF, suggesting that even subtle reductions in systolic function may contribute to increased VEA. Additionally, the observed moderate inverse correlation between hs-CRP and EF highlights the interplay between systemic inflammation and myocardial performance, potentially linking inflammatory burden to both arrhythmic risk and subclinical cardiac dysfunction. Although the correlation between age and VES burden was weak, it may reflect an age-related increase in myocardial irritability or autonomic imbalance in post-COVID patients.

From a clinical standpoint, our results underscore the importance of vigilant cardiac follow-up in recovered COVID-19 patients. Implementing routine 24-hour Holter monitoring for post-COVID individuals with palpitations, syncope, or other worrisome cardiac symptoms could facilitate early detection of significant arrhythmias and guide timely intervention.

Study Limitations

Despite these insights, certain limitations must be acknowledged. This study is retrospective in nature, and the sample size remains relatively small. Furthermore, the long-term arrhythmic risk beyond the study period is unknown. Future research should focus on larger, prospective cohorts with longer follow-up durations to better elucidate the long-term cardiovascular effects of COVID-19. We could not classify COVID-19 positive patients according to clinical severity (e.g., asymptomatic, mild, moderate, severe) due to lack of standardized symptom documentation. This may have influenced arrhythmic outcomes. Future studies should consider subgroup analysis based on COVID-19 severity to better delineate its prognostic impact. Additionally, the exact time interval between COVID-19 infection and Holter monitoring could not be uniformly determined due to

retrospective data collection. This temporal uncertainty limits the ability to distinguish between transient and persistent arrhythmic patterns.

CONCLUSION

COVID-19 infection is significantly associated with an increased burden of VEA. Fragmented QRS and a history of COVID-19 infection are independent predictors of high arrhythmic burden, whereas vaccination appears to be protective. These findings highlight the need for close cardiac monitoring and risk stratification in post-COVID-19 patients, particularly those exhibiting fragmented QRS on ECG. Vaccination appears to have a protective effect against post-COVID arrhythmias, suggesting its potential role in reducing cardiovascular complications. Larger-scale studies with extended follow-up periods are needed to better understand the long-term clinical significance of these findings.

Ethics

Ethics Committee Approval: Ethical approval for this research was granted by The Ethics Committee for Scientific Research, Faculty of Medicine, Trakya University (decision no: 02/17, date: 18.12.2021).

Informed Consent: All participants provided written informed consent before enrollment.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Aktoz M, Altay H, Aslanger E, Atalar E, Atar İ, Aytekin V, et al. Türk Kardiyoloji Derneği uzlaşı raporu: COVID-19 pandemisi ve kardiyovasküler hastalıklar konusunda bilinmesi gerekenler. [Turkish Cardiology Association consensus report: COVID-19 pandemic and cardiovascular diseases. Turk Kardiyol Dern Ars. 2020;48:1-87.
2. Goerlich E, Chung TH, Hong GH, Metkus TS, Gilotra NA, Post WS, et al. Cardiovascular effects of the post-COVID-19 condition. Nat Cardiovasc Res. 2024;3:118-29.
3. Qin M, Lee K, Yoo SJ. The impact of long COVID on heart rate variability: a cross-sectional study. BMC Infect Dis. 2025;25:261.
4. Xiong TY, Redwood S, Prendergast B, Chen M. Coronaviruses and the cardiovascular system: acute and long-term implications. Eur Heart J. 2020;41:1798-800.
5. Kaya Ç, Yalta K. Is constrictive pericarditis associated with long COVID, its vaccine, or both? Balkan Med J. 2024;41:411-2.
6. Tanni SE, Tonon CR, Gatto M, Mota GAF, Okoshi MP. Post-COVID-19 syndrome: cardiovascular manifestations. Int J Cardiol. 2022;369:80-1.
7. Kounis NG, Gogos C, de Gregorio C, Hung M-Y, Kounis SN, Tsounis EP, et al. "When,""where," and "how" of SARS-CoV-2 infection affects the human cardiovascular system: a narrative review. Balkan Medical Journal. 2024;41:7.
8. Turagam MK, Musikantow D, Goldman ME, Bassily-Marcus A, Chu E, Shivamurthy P, et al. Malignant arrhythmias in patients with COVID-19: ıncidence, mechanisms, and outcomes. Circ Arrhythm Electrophysiol. 2020;13:e008920.
9. Taylan G, Öztürk C, Yalta K. Fragmented QRS pattern in patients with COVID-19: further ınsights into ıts temporal and mechanistic features. Anatol J Cardiol. 2022;26:239-40.
10. Katkat F, Kalyoncuoğlu M, Karahan S, Abanus H, Turhan Çağlar FN, Karabulut D, et al. Does coexistence of fragmented QRS and cardiovascular disease have the ability to predict the mortality in hospitalized, critically ill patients with COVID-19? Anatol J Cardiol. 2021;25:803-10.
11. Podestà MA, Valli F, Galassi A, Cassia MA, Ciceri P, Barbieri L, et al. COVID-19 in chronic kidney disease: the ımpact of old and novel cardiovascular risk factors. Blood Purif. 2021;50:740-9.
12. Marques KC, Quaresma JAS, Falcão LFM. Cardiovascular autonomic dysfunction in "long COVID": pathophysiology, heart rate variability, and inflammatory markers. Frontiers In Cardiovascular Medicine. 2023;10:1256512.
13. Koc M, Sumbul HE, Gulumsek E, Koca H, Bulut Y, Karakoc E, et al. Disease severity affects ventricular repolarization parameters in patients with COVID-19. Arq Bras Cardiol. 2020;115:907-13.
14. Çoksevim M, Yenerçağ M. Investigation of autonomic dysfunction following COVID-19: the role of heart rate recovery ındices. Kafkas Tıp Bil Derg. 2024;14:296-303.
15. Pari B, Babbili A, Kattubadi A, Thakre A, Thotamgari S, Gopinathannair R, et al. COVID-19 vaccination and cardiac arrhythmias: a review. Curr Cardiol Rep. 2023;25:925-40.
16. Kole C, Stefanou E, Karvelas N, Schizas D, Toutouzas KP. Acute and post-acute COVID-19 cardiovascular complications: a comprehensive review. Cardiovasc Drugs Ther. 2024;38:1017-32.
17. Lu D, Hu J, Qian J, Cheng F. Temporal variations in QTc interval during and after COVID-19 infection: a retrospective study. BMC Cardiovasc Disord. 2024;24:738.
18. Wu K, Van Name J, Xi L. Cardiovascular abnormalities of long-COVID syndrome: pathogenic basis and potential strategy for treatment and rehabilitation. Sports Med Health Sci. 2024;6:221-31.
19. Jang SW. QTc dispersion predicts prognosis in COVID-19 disease. Korean Circ J. 2021;51:863-5.

ORIGINAL ARTICLE

DOI: 10.4274/nkmj.galenos.2025.46074
Nam Kem Med J 2025;13(3):341-344

Smoking Patterns and Their Association with Histological Subtypes in Lung Cancer Patients

Akciğer Kanseri Hastalarında Sigara İçme Şekilleri ve Histolojik Alt Tiplerle İlişkileri

Nevin FAZLIOĞLU[1], Betül İNAN[2], Semira Sıla DEMİRER[2], Mehmet Kaan BELYURT[2], Celaleddin Arda BAĞ[2], Esila ARSLAN[2], Seyedehmahtab Mosanen TABATABAEI[2], Mithat FAZLIOĞLU[3]

[1]*Tekirdağ Namık Kemal University Faculty of Medicine, Department of Pulmonology, Tekirdağ, Türkiye*
[2]*Tekirdağ Namık Kemal University Faculty of Medicine, Department of General Medicine, Tekirdağ, Türkiye*
[3]*Tekirdağ Namık Kemal University Faculty of Medicine, Department of Thoracic Surgery, Tekirdağ, Türkiye*

ABSTRACT

Aim: Tobacco exposure remains the most significant risk factor in the development of lung cancer. Understanding detailed smoking characteristics in affected populations provides critical insight into disease etiology and progression. This study aimed to analyze smoking status, cumulative tobacco exposure, and their relationship with histological subtypes in a cohort of lung cancer patients.

Materials and Methods: A retrospective analysis was conducted on 539 patients diagnosed with lung cancer. Demographic data, smoking behavior (status and pack-year history), and histological classification were reviewed. Statistical comparisons assessed differences by gender and histological subtype.

Results: Active smoking was the most common status (56.4%) with substantial rates of ex-smoking (34.3%). Mean pack-year history was 49.2±1.03. Squamous cell carcinoma and small cell lung cancer were associated with heavier tobacco exposure. Adenocarcinoma was more common among never and passive smokers. Sex-based differences were significant: females had higher rates of never and passive smoking, while males had higher cumulative exposure.

Conclusion: Smoking characteristics differ markedly by gender and histological subtype in lung cancer patients. These findings underscore the need for personalized approaches to prevention, diagnosis, and public health policy.

Keywords: Lung cancer, smoking behavior, histological subtype

ÖZ

Amaç: Tütün maruziyeti, akciğer kanserinin gelişiminde en önemli risk faktörü olmaya devam etmektedir. Etkilenen popülasyonlardaki sigara içim özelliklerinin ayrıntılı olarak anlaşılması, hastalığın etiyolojisi ve ilerleyişi hakkında kritik bilgiler sunar. Bu çalışma, sigara içme durumu, kümülatif tütün maruziyeti ve bunların histolojik alt tiplerle ilişkisini analiz etmeyi amaçlamaktadır.

Gereç ve Yöntem: Akciğer kanseri tanısı almış 539 hasta retrospektif olarak analiz edilmiştir. Demografik veriler, sigara içme davranışı (içme durumu ve paket-yıl geçmişi) ile histolojik sınıflandırmalar incelenmiştir. Cinsiyete ve histolojik alt tipe göre farklılıklar istatistiksel olarak karşılaştırılmıştır.

Bulgular: Aktif sigara içimi en yaygın durumdu (%56,4) ve kayda değer oranda ex-smoker (%34,3) mevcuttu. Ortalama paket-yıl geçmişi 49,2±1,03 idi. Skuamöz hücreli karsinom ve küçük hücreli akciğer kanseri daha yoğun tütün maruziyetiyle ilişkiliydi. Adenokarsinom, sigara içmemiş ve pasif içici hastalarda daha sık görülüyordu. Cinsiyet temelli farklılıklar anlamlıydı: kadınlar daha yüksek oranda hiç sigara içmemiş veya pasif içici iken, erkekler daha yüksek kümülatif maruziyet gösterdi.

Sonuç: Sigara içim özellikleri, akciğer kanseri hastalarında cinsiyete ve histolojik alt tipe göre belirgin şekilde farklılık göstermektedir. Bu bulgular, önleme, tanı ve halk sağlığı politikalarında kişiselleştirilmiş yaklaşımların gerekliliğini vurgulamaktadır.

Anahtar Kelimeler: Akciğer kanseri, sigara içme davranışı, histolojik alt tıp

Address for Correspondence: Nevin FAZLIOĞLU MD, Tekirdağ Namık Kemal University Faculty of Medicine, Department of Pulmonology, Tekirdağ, Türkiye
E-mail: nfazlioglu@nku.edu.tr **ORCID ID:** orcid.org/0000-0001-5191-4636
Received: 01.06.2025 **Accepted:** 21.07.2025 **Publication Date:** 07.10.2025

Cite this article as: Fazlıoğlu N, Demirer SS, Belyurt MK, Bağ CA, Arslan E, Tabatabaei SM, et al. Smoking patterns and their association with histological subtypes in lung cancer patients. . Nam Kem Med J. 2025;13(3):341-344

INTRODUCTION

Lung cancer remains the leading cause of cancer-related mortality worldwide. Tobacco smoking is the most significant etiological factor, yet smoking patterns and their impact on lung cancer subtypes differ across genders. While men have traditionally shown higher smoking rates, recent shifts in social norms have led to increased female exposure both directly and passively[1-4]. Understanding how sex-specific smoking behaviors correlate with histological subtypes of lung cancer can provide valuable insights for risk stratification and targeted prevention strategies.

MATERIALS AND METHODS

Ethics approval for this study was obtained from the Non-Interventional Clinical Research Ethics Committee of Tekirdağ Namık Kemal University (desicion no: 2025.64.03.22, date: 25.03.2025). A retrospective cohort study was conducted including 539 patients diagnosed with lung cancer between 2017 and 2024. Demographic variables included age, gender, and survival status. Smoking-related variables included smoking status (never, active, passive, ex-smoker) and cumulative exposure (pack-year history). Histological classification was based on pathology reports: squamous cell carcinoma (SCC), adenocarcinoma, small cell lung cancer (SCLC), and non-SCLC not otherwise specified. Passive smoking was defined as regular exposure to cigarette smoke at home or in the workplace, as reported by the patient and documented in the anamnesis section of the medical records. The presence or absence of passive smoking was retrospectively extracted from the physician's notes. No objective quantification was available.

Statistical Analysis

Descriptive statistics were presented as means ± standard error of the mean for continuous variables and percentages for categorical data. ANOVA and chi-square tests were used to assess differences among groups. A p-value<0.05 was considered statistically significant.

RESULTS

Patient Characteristics

The study population had a mean age of 64.0±0.42 years. Males accounted for 84.8% (n=457) and females 15.2% (n=82). Mean age at diagnosis was 63.5±0.42 years, and the mean survival time following diagnosis was 11.4±1.53 months (Table 1).

Smoking Behavior by Sex

The overall smoking behavior differed significantly between the sexes. Active smoking was common in both groups (52.4% of females vs. 57.1% of males). However, females had markedly higher rates of never smoking (20.7% vs. 2.8%) and passive exposure (14.6% vs. 1.8%), while ex-smoking was more prevalent in males (38.3% vs. 12.2%) (Table 2). Mean pack-year exposure was significantly higher in males (49.92±1.09) than in females (43.19±3.23), despite the lower proportion of ex-smokers in the latter group.

Histological Subtype and Smoking Correlation

SCC was the most frequent subtype among males (38.6%) and showed the highest mean pack-year (53.01±1.75). In contrast, adenocarcinoma was slightly more common in females (36.6%) and had the lowest mean pack-year exposure (43.16±1.75) (Table 3 and 4).

ANOVA showed significant differences in mean pack-year history among subtypes (p=0.0016) (Table 4). The chi-square test confirmed a significant association between smoking status and histological subtype (p=0.0002).

DISCUSSION

This study reveals that sex differences in smoking behavior are not only statistically significant but also pathologically relevant, manifesting as distinct patterns in lung cancer histology.

Table 1. General characteristics of the study population

Characteristic	Value
Mean age	64.0±0.42
Sex	
Male	457 (84.8)
Female	82 (15.2)
Smoking status	
Never smokers	30 (5.6)
Active smokers	304 (56.4)
Passive smokers	20 (3.7)
Ex-smokers	185 (34.3)
The mean pack-year history	49.2±1.03
Histology	
SCC	197 (36.5)
Adenocarcinoma	164 (30.4)
SCLC	108 (20.0)
NSCLC (NOS)	69 (12.8)
Survival status	
Ex	300 (55.7)
Alive	239 (44.3)
Mean age at diagnosis	63.5±0.42
Mean survival time after diagnosis	11.4±1.53

ANOVA test was performed. A statistically significant difference in mean pack-year history was observed among histological subtypes (p=0.0016). NSCLC (NOS): Non-small cell lung cancer-not otherwise specified, SCC: Squamous cell carcinoma, SCLC: Small cell lung cancer

Table 2. Smoking status and pack-year history by sex

Sex	Never smoker (%)	Active smoker (%)	Passive smoker (%)	Ex-smoker (%)	Mean pack-years	SD	N
Female	20.7	52.4	14.6	12.2	43.19	3.23	53
Male	2.8	57.1	1.8	38.3	49.92	1.09	436

SD: Standard deviation, N: Number

Tablo 3. Cinsiyete göre histolojik alt tip dağılımı

Sex	Adenocarcinoma	NSCLC (NOS)	SCC	SCLC
Female	36.6	13.4	25.6	24.4
Male	29.4	12.7	38.6	19.3

NSCLC (NOS): Non-small cell lung cancer-not otherwise specified, SCC: Squamous cell carcinoma, SCLC: Small cell lung cancer

Table 4. Mean pack-year history by histological subtype

Histology	Mean pack-years	SD	N
Adenocarcinoma	43.16	1.75	138
NSCLC (NOS)	49.21	3.02	63
SCC	53.01	1.75	187
SCLC	50.36	2.16	101

ANOVA test was performed. A statistically significant difference in mean pack-year history was observed among histological subtypes (p=0.0016). NSCLC (NOS): Non-small cell lung cancer-not otherwise specified, SCC: Squamous cell carcinoma, SCLC: Small cell lung cancer, SD: Standard deviation

Women, although demonstrating markedly lower rates of active smoking compared to men, show significantly elevated levels of passive smoke exposure. This discrepancy suggests that indirect tobacco exposure may play a more pronounced etiological role in the female population than previously assumed[5]. Of particular concern is the disproportionately high incidence of SCLC among female patients a histological subtype recognized for its aggressive clinical course and poor prognosis. This observation raises the possibility that even non-direct smoking exposure may be sufficient to trigger the development of high-grade malignancies, particularly in biologically susceptible individuals.

Furthermore, the relatively higher prevalence of adenocarcinoma in women corresponds with findings from prior studies, which have consistently linked this histological subtype to lighter or indirect smoking habits. This pattern has led researchers to hypothesize that adenocarcinoma may arise through different carcinogenic pathways than SCC, including possible interactions with hormonal influences, such as estrogen receptors, or genetic predispositions unique to female patients[6-9].

In contrast, the data underscore that men, who have higher rates of both current and former smoking, also demonstrate significantly greater cumulative tobacco exposure as measured in pack-years. This increased exposure is paralleled by the predominance of SCC among male patients a subtype long known to be closely associated with heavy and prolonged smoking. The dose-response relationship observed here reinforces the carcinogenic potency of long-term tobacco consumption, particularly in the development of central airway tumors such as SCC[10-12].

Taken together, these findings highlight the critical importance of incorporating gender-specific smoking patterns into predictive models of lung cancer risk[7-9,13]. Public health policies and clinical screening protocols should not only continue to target active smoking but also increase focus on mitigating passive smoke exposure, particularly in women. In the era of personalized medicine, such stratified approaches could enhance both early detection and prevention efforts by tailoring them to the unique risk profiles shaped by gender and tobacco exposure dynamics.

Study Limitations

This study is limited by its retrospective design and reliance on medical records for smoking history, which may introduce recall bias. The relatively small female sample size also restricts the power of gender-based comparisons. Additionally, passive smoking exposure was not objectively measured, and potential confounders such as occupational and environmental factors were not accounted for.

CONCLUSION

Sex plays a critical role in shaping smoking exposure patterns and their oncological consequences. Recognition of these differences is crucial for developing more precise screening and prevention strategies in lung cancer care. Future studies should explore biological underpinnings and psychosocial determinants influencing gender disparities in lung carcinogenesis.

Ethics

Ethics Committee Approval: Ethics approval for this study was obtained from the Non-Interventional Clinical Research Ethics Committee of Tekirdağ Namık Kemal University (desicion no: 2025.64.03.22, date: 25.03.2025).

Informed Consent: A retrospective cohort study was conducted including 539 patients diagnosed with lung cancer between 2017 and 2024.

Footnotes

Concept: N.F., B.İ., M.K.B., E.A., M.F., Design: N.F., S.S.D., M.K.B., S.M.T., M.F., Data Collection or Processing: N.F., B.İ., S.S.D., M.K.B., C.A.B., E.A., S.M.T., Analysis or Interpretation: N.F., C.A.B., E.A., S.M.T., M.F., Literature Search: N.F., B.İ., S.S.D., M.K.B., E.A., Writing: N.F., B.İ., C.A.B., M.F.

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

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

REFERENCES

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin. 2023;73:17-48.
2. Wakelee HA, Chang ET, Gomez SL, Keegan TH, Feskanich D, Clarke CA, et al. Lung cancer incidence in never smokers. J Clin Oncol. 2007;25:472-8.
3. Couraud S, Zalcman G, Milleron B, Morin F, Souquet PJ. Lung cancer in never smokers--a review. Eur J Cancer. 2012;48:1299-311.
4. Thun MJ, Henley SJ, Burns D, Jemal A, Shanks TG, Calle EE. Lung cancer death rates in lifelong nonsmokers. J Natl Cancer Inst. 2006;98:691-9.
5. Wang YT, Hu KR, Zhao J, Ai FL, Shi YL, Wang XW, et al. The association between exposure to second-hand smoke and disease in the chinese population: a systematic review and meta-analysis. Biomed Environ Sci. 2023;36:24-37.
6. Hsu LH, Chu NM, Liu CC, Tsai SY, You DL, Ko JS, et al. Sex-associated differences in non-small cell lung cancer in the new era: is gender an independent prognostic factor? Lung Cancer. 2009;66:262-7.
7. Patel JD. Lung cancer in women. J Clin Oncol. 2005;23:3212-8.
8. Siegfried JM. Women and lung cancer: does oestrogen play a role? Lancet Oncol. 2001;2:506-13.
9. Risch HA, Howe GR, Jain M, Burch JD, Holowaty EJ, Miller AB. Are female smokers at higher risk for lung cancer than male smokers? A case-control analysis by histologic type. Am J Epidemiol. 1993;138:281-93.
10. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat Rev Cancer. 2003;3:733-44.
11. Wyss A, Hashibe M, Chuang SC, Lee YC, Zhang ZF, Yu GP, et al. Cigarette, cigar, and pipe smoking and the risk of head and neck cancers: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Am J Epidemiol. 2013;178:679-90.
12. Pesch B, Kendzia B, Gustavsson P, Jöckel KH, Johnen G, Pohlabeln H, et al. Cigarette smoking and lung cancer--relative risk estimates for the major histological types from a pooled analysis of case-control studies. Int J Cancer. 2012;131:1210-9.
13. LoMastro A, Maki RG. Sex and gender differences in cancer risk and treatment. J Clin Oncol. 2019;37:2446-54.

CASE REPORT

DOI: 10.4274/nkmj.galenos.2025.88942
Nam Kem Med J 2025;13(3):345-348

Diffuse Large B-Cell Lymphoma with Second Relapse in Leukemic Phase

İkinci Nüksünde Lösemik Faz ile Başvuran Yaygın Büyük B Hücreli Lenfoma

Sinan DEMİRCİOĞLU[1], İbrahim GÜRBÜZ[2], Naile KÖKBUDAK[3], Hacı Hasan ESEN[3], Atakan TEKİNALP[1]

[1]*Necmettin Erbakan University Faculty of Medicine, Department of Hematology, Konya, Türkiye*
[2]*Necmettin Erbakan University Faculty of Medicine, Konya, Türkiye*
[3]*Necmettin Erbakan University Faculty of Medicine, Department of Pathology, Konya, Türkiye*

ABSTRACT

The leukemic phase of diffuse large B-cell lymphoma (DLBCL) is a rare condition and can be difficult to distinguish from acute leukemia or the leukemic phases of other non-Hodgkin lymphoma subtypes. When intermediate-to-large abnormal lymphoid cells are observed in the peripheral blood, it should be considered as one of the differential diagnoses. Flow cytometry and immunohistochemical staining are helpful for definitive diagnosis. The leukemic phase typically occurs in the progressive phase or stage IV disease and is rare at the time of initial diagnosis in DLBCL. The development of a leukemic phase during the course of the disease is particularly associated with poor prognosis. There are no evidence-based treatment recommendations for this condition. In our case, we presented a patient with DLBCL who developed a leukemic phase upon second relapse.

Keywords: Diffuse large B-cell lymphoma, leukemic phase, prognosis

ÖZ

Yaygın diffüz büyük B hücreli lenfoma (DLBCL) lösemik fazı nadir görülen bir durumdur ve akut lösemiden veya diğer Hodgkin dışı lenfoma türlerinin lösemik fazlarından ayırt edilmesi zor olabilir. Periferik kanda orta ila büyük boyutlarda anormal lenfoid hücreler izlenmesi durumunda ayırıcı tanılardan biri olarak düşünülmelidir. Kesin tanı için flowsitometri ve immünohistokimyasal boyalar yardımcı olur. Lösemik faz genellikle ilerleyici bir faz veya evre IV hastalıkta görülür ve DLBCL tanı anında nadirdir. Özellikle hastalığın seyrinde lösemik faz gelişmesi kötü prognozla ilişkilidir. Tedavisi hakkında kanıta dayalı öneriler yoktur. Bizde DLBCL tanılı hastanın ikinci nüksünde lösemik faz başvurusunu sunduk.

Anahtar Kelimeler: Yaygın büyük B hücreli lenfoma, lösemik faz, prognoz

INTRODUCTION

Diffuse large B-cell lymphoma (DLBCL) is the most common subtype of non-Hodgkin lymphomas (NHLs) and accounts for approximately one-third of all types of NHL[1]. The presence of malignant lymphoma cells in peripheral blood is well recognized in mantle cell lymphoma, follicular lymphoma, anaplastic large cell lymphoma, and the terminal phases of all refractory lymphomas[2-5]. The leukemic phase of DLBCL is a rare condition and may be difficult to distinguish from acute leukemia or other types of NHL[6-9]. The leukemic phase usually occurs as a progressive phase or in stage IV disease and is rare at the time of diagnosis in DLBC[10,11].

Address for Correspondence: Sinan DEMİRCİOĞLU MD, Necmettin Erbakan University Faculty of Medicine, Department of Hematology, Konya, Türkiye
E-mail: sinandemircioglumd@gmail.com **ORCID ID:** orcid.org/0000-0003-1277-5105
Received: 03.03.2025 **Accepted:** 21.04.2025 **Publication Date:** 07.10.2025

Cite this article as: Demircioğlu S, Gürbüz İ, Kökbudak N, Esen HH, Tekinalp A. Diffuse large B-cell lymphoma with second relapse in leukemic phase. Nam Kem Med J. 2025;13(3):345-348

CASE REPORT

A 68-year-old female patient presented with complaints of neck swelling, weight loss, and night sweats. Complete blood count was normal. Superficial ultrasonography revealed multiple lymph nodes in bilateral cervical, inguinal, and axillary regions, the largest of which measured 20 mm in short axis, some with a reticular appearance, some with a compressed hilum, and thick cortex. An excisional biopsy was performed from the cervical region. The biopsy revealed tumor cells positive for immunohistochemical markers CD20, BCL-2, CD79A, and 50% positive for C-MYC, 1-2% positive for MUM-1, and negative for CD3, CD5, CD10, BCL-6, CD23, CD30, and Cyclin D1. The Ki-67 index was 80-90% (Figure 1). Diagnosis of non-germinal center type DLBCL was made. Bone marrow biopsy was also found to be consistent with DLBCL involvement. The patient with Stage 4 high International Prognostic index (IPI) score was given rituximab-gemcitabine-cyclophosphamide-vincristine-prednizon (R-GCVP) due to low ejection fraction. Prophylactic intrathecal (IT) methotrexate was administered as a result of the high probability of central nervous system (CNS) recurrence of lymphoma. After 3 cycles of R-GCVP, a partial metabolic response to treatment was observed with positron emission tomography-computed tomography. Continuation of R-GCVP treatment was planned but the patient developed strabismus during follow-up. Contrast brain magnetic resonance imaging (MRI) was performed to investigate CNS involvement. A mass was detected on contrasted pituitary MRI, which, when evaluated radiologically and clinically together, was considered as lymphoma infiltration. The patient was planned to start MATRix chemotherapy protocol. Three cycles were administered. After 3 cycles of MATRix chemotherapy, a follow-up contrasted pituitary MRI showed significant regression. Thereupon, the patient was planned for autologous hematopoietic stem cell transplantation (AHSCT). Stem cell mobilization was performed. When the patient was admitted for AHSCT one month later, she had complaints of fatigue and abdominal pain. Complete blood count revealed a white blood cell count of 16,760/mm^3, hemoglobin 8.6 g/dL, and platelet count of 169,000/mm^3. Peripheral smear showed leukocytosis and 34% medium- to large-sized atypical lymphoid cells with condensed nuclear chromatin and inconspicuous nucleoli (Figure 2). In the flow cytometry performed on peripheral blood, a lymphoid cell population of approximately 45% was observed, showing positive expression of CD19, CD20, CD22, CD45, CD79a and negative expression of CD3, CD5, CD7, CD23, CD34, CD56, TdT. Indistinct, scattered, slightly hyperintense nodular lesions with the largest measuring 2 cm were observed in the liver. Tru-cut needle biopsy was performed here. Liver biopsy was consistent with non-germinal center type DLBCL. During this period, the patient's leukocyte count rapidly increased to 50,000/mm^3, which was consistent with the leukemic phase of DLBCL. Since she was unfit and refractory to standard chemotherapy, rituximab-ibrutinib-lenalidomide treatment was started. There was no response to this treatment. The patient died due to renal failure and sepsis within a short period of 1 month after the leukemic phase diagnosis. A short survival of 11 months was achieved compared to the time of the first diagnosis.

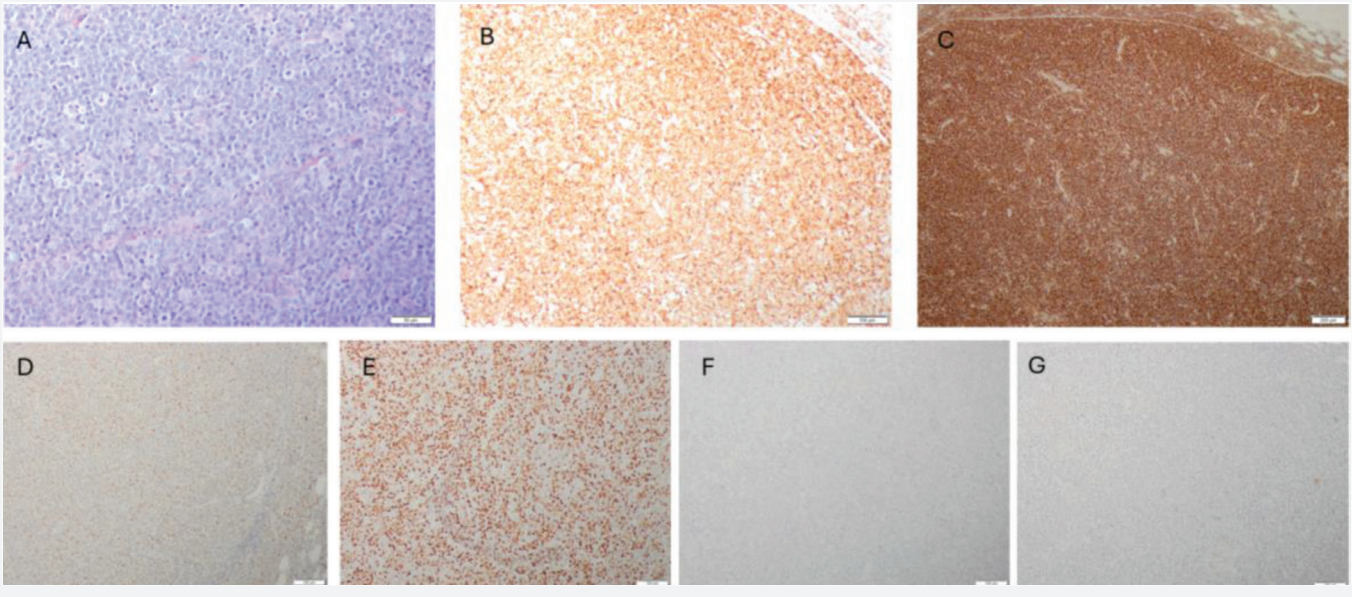


Figure 1. (A) Hematoxylin-eosin (50 µm), (B) BCL2 positive (scale bar: 100 µm), (C) CD20 positive (scale bar: 200 µm), (D) C-MYC 50% positive (scale bar: 100 µm), (E) Kİ-67 80-90% (scale bar: 100 µm) (F) BCL6: Negative (scale bar: 100 µm) (G) CD10 negative (scale bar:100 µm)

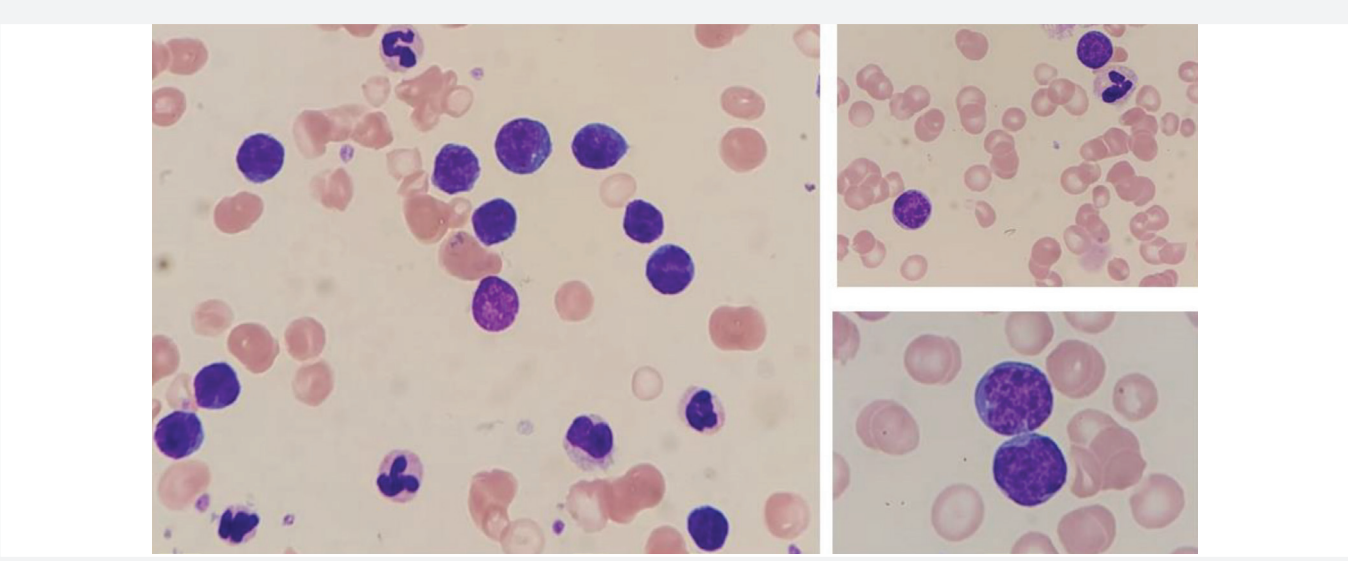


Figure 2. Peripheral smear shows large-sized atypical lymphoid cells with condensed nuclear chromatin and inconspicuous nucleoli

DISCUSSION

Lymphomas are diagnosed primarily based on histologic findings, although dissemination of these lymphoma cells into the circulation (leukemic phase) can be diagnosed based on cellular immunophenotypic analysis by flow cytometry[8]. Immunophenotypically, these cells show strong membrane positivity for B-cell lineage markers such as CD19, CD20 and follicular center markers such as CD10 (40%) and BCL6 (60%). Non-germinal center type DLBCL will show positivity for CD38 and MUM1[11,12].

IPI is the primary prognostic scoring system used for patients with DLBCL. IPI was developed to assess pre-treatment characteristics that predict outcomes in patients with aggressive non-Hodgkin lymphoma, including DLBCL. In patients receiving chemotherapy containing doxorubicin, overall survival and progression-free survival are associated with age, serum lactate dehydrogenase levels, performance status, clinical stage, and extranodal disease[13]. Following the introduction of rituximab, the IPI model was validated in patients treated with rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone (R-CHOP) and R-CHOP-like regimens[14,15]. Other studies have investigated prognostic markers such as neutrophil count, lymphocyte count, neutrophil/lymphocyte ratio, mean platelet volume, uric acid, and fibrinogen levels[16,17].

A study of 29 patients with leukemic phase DLBCL showed that all patients had extranodal involvement, high IPI, and poor performance status. Anthracycline and rituximab-based regimens R-CHOP or rituximab, hyperfractionated cyclophosphamide, vincristine, doxorubicin achieved an overall response of 88% and a complete response (CR) of 54%, with a 4-year survival of approximately 50%[12]. A patient diagnosed with leukemic phase DLBCL at the age of twenty-eight and given 6 cycles of DA-EPOCH (etoposide, prednisone, vincristine, cyclophosphamide, doxorubicin) chemotherapy followed by abdominal radiotherapy was reported to have relapsed in a short period of 2.5 months. The patient, who was subsequently given the R-BFM-90 protocol, died due to disease progression[7]. A 74-year-old female patient presented with DLBCL leukemic phase and conventional cytogenetic studies showed a complex karyotype with t (8;14) (q24q32). Although a CR was achieved with the modified hyper CVAD protocol, relapse occurred in a short time and permanent remission could not be achieved with various treatments[18]. One case of DLBCL in the leukemic phase that was CD19 negative was reported[19]. In another case report, a diagnosis of de novo DLBCL in the leukemic phase was made, positive for both CD5 and CD13, and the patient was treated with the recombinant human cartilage oligomeric matrix protein regimen and achieved complete remission but relapsed one month later. This was interpreted by the authors as an indicator of poor prognosis for DLBCL with CD5+ leukemic presentation[20]. A case of DLBCL with complex karyotype including TP53 deletion with leukemic phase and cerebrospinal fluid (CSF) involvement was given R-Hyper-CVAD chemotherapy with twice weekly IT therapy (alternative methotrexate and cytarabine) for two cycles. Due to persistent CSF involvement, his treatment was changed to rituximab and IT thiotepa was added to his regimen. Finally, after more than 2 months of systemic and CT chemotherapy, his CSF cleared. Then, consolidative craniospinal irradiation was performed before proceeding to allogeneic transplantation in the first remission. Unfortunately, despite aggressive measures, CNS relapsed on the 50th day post-allogeneic transplantation[21]. A 54-year-old patient with leukemic phase DLBCL was treated with a combination of rituximab, cyclophosphamide, R-CHOP and lenalidomide because of myelocytomatosis oncogene positivity, and achieved a CR[9].

Forty-five patients with relapsed and refractory DLBCL, who had received at least one prior treatment, were treated with ibrutinib + rituximab + lenalidomide. Of the patients, 51% had non-germinal center B-cell (non-GCB) like DLBCL, 33% had transformed DLBCL, 60% were refractory, and 27% had primary refractory disease. The overall response rate (ORR) was found to be 44% (CR: 28%); among these, in non-GCB patients, the ORR was 65% (CR: 41%), in relapsed patients (n=16), the ORR was 69%, and in secondary refractory patients (n=27), the ORR was 56%[22]. Our patient was relapsed/refractory after receiving two lines of treatment, and their current condition was not suitable for standard chemotherapy. Based on the study mentioned above, and considering the diagnosis of non-GCB DLBCL and performance status, treatment with ibrutinib + rituximab + lenalidomide was initiated.

CONCLUSION

Although the study by Muringampurath-John et. al[12] mentioned above found that the response rates and survival of DLBCL leukemic phase patients were similar to those of non-leukemic phase DLBCL, neither our case nor other cases presented in the literature support this. Our case survived only one month after entering the leukemic phase and the total survival was eleven months. There is insufficient evidence in the literature regarding the treatment of DLBCL leukemic phase. Since most of these patients may be primary refractory, new treatment methods are needed.

Ethics

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

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood. 2016;127:2375-90.
2. Morra E, Lazzarino M, Orlandi E, Pagnucco G, Brusamolino E, Castello A, Inverardi D, Bernasconi C. Leukemic phase of non-Hodgkin's lymphomas. Hematological features and prognostic significance. Haematologica 1984;69:15-29.
3. Berinstein NL, Reis MD, Ngan BY, Sawka CA, Jamal HH, Kuzniar B. Detection of occult lymphoma in the peripheral blood and bone marrow of patients with untreated early-stage and advanced-stage follicular lymphoma. J Clin Oncol. 1993;11:1344-52.
4. Criel A, Pittaluga S, Verhoef G, Wlodarska I, Meeus P, Mecucci C, et al. Small B cell NHL and their leukemic counterpart: differences in subtyping and assessment of leukemic spread. Leukemia. 1996;10:848-53.
5. Villamor N, Rozman M, Esteve J, Aymerich M, Colomer D, Aguilar JL, et al. Anaplastic large-cell lymphoma with rapid evolution to leukemic phase. Ann Hematol. 1999;78:478-82.
6. Bain BJ, Catovsky D. The leukaemic phase of non-Hodgkin's lymphoma. J Clin Pathol. 1995;48:189-93.
7. Balasubramanian P, Ramteke P, Mallick S, Kumar L, Tanwar P. Diffuse large B-cell lymphoma relapsing in leukaemic phase presenting as acute leukaemia. Clin Med Insights Blood Disord. 2019;12:1179545X18821160.
8. Suresh PK, Basavaiah SH, Goel M, Dsouza S, Rai S. Diffuse large B-cell lymphoma in blood and bone marrow: a rare presentation of leukemic phase at diagnosis. Hematol Transfus Cell Ther. 2020;42:180-3.
9. Bilgin YM, Castel R. MYC -Positive diffuse large B-cell lymphoma in leukemic phase at presentation: a diagnostic and therapeutic challenge. J Hematol 2021;10:80-2.
10. Wong E, Juneja S. Diffuse large B-cell lymphoma in leukemic phase with flower cell morphology. Blood. 2014;123:3378.
11. Pires PP, Kanegae MY, Rays J, Catania M, Lima FR, Noronha TR, et al. Diffuse large B-cell lymphoma presenting in the leukemic phase. Autops Case Rep. 2016;6:41-5.
12. Muringampurath-John D, Jaye DL, Flowers CR, Saxe D, Chen Z, Lechowicz MJ, et al. Characteristics and outcomes of diffuse large B-cell lymphoma presenting in leukaemic phase. Br J Haematol. 2012;158:608-14.
13. International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. N Engl J Med. 1993;329:987-94.
14. Ziepert M, Hasenclever D, Kuhnt E, Glass B, Schmitz N, Pfreundschuh M, et al. Standard International prognostic index remains a valid predictor of outcome for patients with aggressive CD20+ B-cell lymphoma in the rituximab era. J Clin Oncol. 2010;28:2373-80.
15. Salles G, de Jong D, Xie W, Rosenwald A, Chhanabhai M, Gaulard P, et al. Prognostic significance of immunohistochemical biomarkers in diffuse large B-cell lymphoma: a study from the Lunenburg Lymphoma Biomarker Consortium. Blood. 2011;117:7070-8.
16. Demircioğlu S, Doğan A, Ekinci Ö, Demir C. The association of NLR and MPV with treatment responses, disease stages, and international prognostic ındex scores in DLBCL. Eastern Journal Of Medicine. 2019;24:325-9.
17. Çiftçiler R, Ciftciler AE, Selim C. Evaluation of the association between serum danger molecules at diagnosis and survival outcomes in patients with diffuse large B cell lymphoma. Med Int (Lond). 2025;5:25.
18. Jain P, Sargent RL, Konoplev SN, Benjamini O, Kantarjian H, Kadia T. Multiple recurrent extra-medullary relapses of high-grade diffuse large B-cell lymphoma presenting in acute leukemic phase. Am J Hematol. 2013;88:433-4.
19. Ramalingam TR, Uthayasuriyan R, Prabhakar V, Vaidhyanathan L. CD19 negative diffuse large B-cell lymphoma presenting in leukemic phase. Cytometry B Clin Cytom. 2023;104:331-3.
20. Carulli G, Ciancia EM, Caracciolo F, Sammuri P, Domenichini C, Ferreri MI, et al. A rare case of de novo CD5+ diffuse large B-cell lymphoma in leukemic phase and positive for CD13. Hematol Rep. 2018;9:7437.
21. Sinkoff J, Olteanu H, Peterson JF, Shah NN. Leukemic phase and CSF involvement of diffuse large B-cell lymphoma with a complex karyotype including a TP53 deletion. Clin Case Rep. 2017;6:235-7.
22. Goy A, Ramchandren R, Ghosh N, Munoz J, Morgan DS, Dang NH, et al. Ibrutinib plus lenalidomide and rituximab has promising activity in relapsed/refractory non-germinal center B-cell-like DLBCL. Blood. 2019;134:1024-36.

CASE REPORT

DOI: 10.4274/nkmj.galenos.2025.90592
Nam Kem Med J 2025;13(3):349-351

Levodopa-Responsive Parkinsonism After Bilateral Hemorrhagic Stroke

İki Yanlı Hemorajik İnme Sonrası Gelişen Levodopa Yanıtlı Parkinsonizm Olgusu

Sadettin KALENDEROĞLU[1], Yüksel ERDAL[2]

[1]*University of Health Sciences Türkiye, İstanbul Training and Research Hospital, Clinic of Neurology, İstanbul, Türkiye*
[2]*Bezmialem University Faculty of Medicine, Department of Neurology, İstanbul, Türkiye*

ABSTRACT

Movement disorders after stroke are not uncommon and frequently manifest as chorea, ballismus, tremor, or parkinsonism. Post-stroke movement disorders may present as an acute symptom of the cerebrovascular event or, in certain instances, may emerge as a delayed sequela. This case report aimed to examine parkinsonism in a 55-year-old female patient who experienced two hemorrhagic strokes approximately five years apart and demonstrated a positive response to levodopa therapy. It is important to note that, in addition to pyramidal signs, extrapyramidal signs may also be observed subsequent to a cerebrovascular accident, and levodopa therapy can significantly ameliorate clinical manifestations.

Keywords: Hemorrhagic stroke, parkinsonism, movement disorders, levodopa, case report

ÖZ

İnme sonrası hareket bozuklukları nispeten yaygındır ve sıklıkla kore, ballismus, tremor, parkinonsizm şeklinde ortaya çıkmaktadır. İnme sonrası hareket bozuklukları kimi zaman inmenin akut semptomu olarak ortaya çıkarken bazı durumlarda ise gecikmiş bir sekel olarak ortaya çıkabilir. Bu olgu sunumunda, yaklaşık 5 yıl arayla iki kez hemorajik inme geçiren 55 yaşındaki bir kadın hastada gelişen ve levodopa tedavisine yanıt veren parkinsonizm ele alınmaktadır. İnme sonrası piramidal bulguların yanı sıra ekstrapiramidal bulguların da görülebileceği ve levodopa tedavisinin klinik bulguları anlamlı ölçüde düzeltebileceği akılda bulundurulmalıdır.

Anahtar Kelimeler: Hemorajik inme, parkinsonizm, hareket bozuklukları, levodopa, olgu sunumu

INTRODUCTION

Post-stroke movement disorders are one of the most common causes of secondary movement disorders[1]. The incidence is estimated to be approximately 1-4% among all strokes[2-4]. Post-stroke movement disorders may present as parkinsonism or a variety of hyperkinetic disorders such as chorea, ballism, athetosis, dystonia, tremor, myoclonus, stereotypy, and akathisia[5]. Such disorders can present as acute symptoms of the stroke or may emerge later as delayed sequelae[6,7]. Here, we aimed to present a case of post-stroke parkinsonism following a hemorrhagic stroke, which showed a favorable clinical response to levodopa therapy.

CASE REPORT

A 55-year-old female patient presented to our outpatient clinic with complaints of bradykinesia and speech disturbance following second episode of hemorrhagic stroke. Her medical history included hypertension and two episodes of hemorrhagic stroke. The patient experienced the first hemorrhagic stroke 5 years ago, and the second took place three months prior to presentation; both episodes required hospitalization. Neurological examination revealed dysarthric speech, hypomimia, bilateral wrist rigidity, and bilateral bradykinesia. Her gait was characterized by short steps with reduced

Address for Correspondence: Sadettin KALENDEROĞLU MD, University of Health Sciences Türkiye, İstanbul Training and Research Hospital, Clinic of Neurology, İstanbul, Türkiye
E-mail: sadettin.kalenderoglu@gmail.com **ORCID ID:** orcid.org/0000-0001-9697-9871
Received: 06.03.2025 **Accepted:** 27.04.2025 **Publication Date:** 07.10.2025

Cite this article as: Kalenderoğlu S, Erdal Y. Levodopa-responsive parkinsonism after bilateral hemorrhagic stroke. Nam Kem Med J. 2025;13(3):349-351

associated arm swing bilaterally. There were no complaints of constipation, rapid-eye-movement disorders or hyposmia as the non-motor symptom associated with idiopathic Parkinson's disease.

Brain magnetic resonance imaging demonstrated a T2-weighted fluid attenuated inversion recovery hyperintensity in the right putamen consistent with a subacute hematoma (Figure 1A.), as well as bilateral diffuse T1 hypointensities in the putaminal white matter (Figure 1B). Following the initiation of levodopa/benserazide (25/100 mg, three times daily), a significant improvement in parkinsonian symptoms was observed. This clinical improvement was sustained at both the 6-month and 12-month follow-up visits.

DISCUSSION

Parkinsonism can develop as part of a degenerative process or, as in our case, due to secondary causes[7,8]. In secondary parkinsonism, core features such as bradykinesia, rigidity, resting tremor, and postural instability, commonly seen in primary Parkinson's disease, are also present[4,7,9]. Additionally, dysarthria, hypomimia, and gait disturbances, as observed in our patient, may accompany the clinical presentation. However, unlike idiopathic Parkinson's disease, the findings in our case were bilateral, and no non-motor symptoms were observed. In stroke-related cases, coexisting neurological deficits may further complicate the clinical picture[6]. In our patient, who exhibited clinical signs of stroke, newly manifested extrapyramidal signs were addressed, and treatment for parkinsonism was initiated.

In the literature, hyperkinetic disorders such as chorea, ballism, or dystonia, which are more pronounced and easily recognized, are frequently reported early in the post-stroke period[3,5]. In contrast, hypokinetic disorders like parkinsonism often have a more insidious onset or may be masked by other neurological impairments, potentially delaying diagnosis[1,10]. Therefore, clinicians should consider parkinsonism in stroke survivors presenting with acute or gradually worsening bradykinesia, rigidity, or gait disturbances.

Classical vascular parkinsonism typically results from chronic or recurrent ischemic lesions due to small vessel disease, usually with a gradual onset and predominant lower-body involvement[11]. Hemorrhagic strokes, however, can acutely or subacutely involve the basal ganglia or deep gray matter, leading to parkinsonism[6]. In our case, the subacute recognition of parkinsonism and the patient's presentation with notable functional impairment suggest a clinical course that differs from "classical" vascular parkinsonism. While patients with vascular parkinsonism often do not benefit significantly from levodopa, this patient demonstrated marked improvement in parkinsonian symptoms with levodopa therapy despite bilateral striatal damage. Indeed, similar cases of levodopa-responsive parkinsonism have been reported in isolated bilateral putaminal hemorrhage though they are relatively rare[12].

CONCLUSION

Most cases of post-stroke movement disorders reported in the literature have been associated with lesions involving small vessel disease in the middle cerebral artery or posterior cerebral artery territories[11,13]. However, cases of parkinsonism

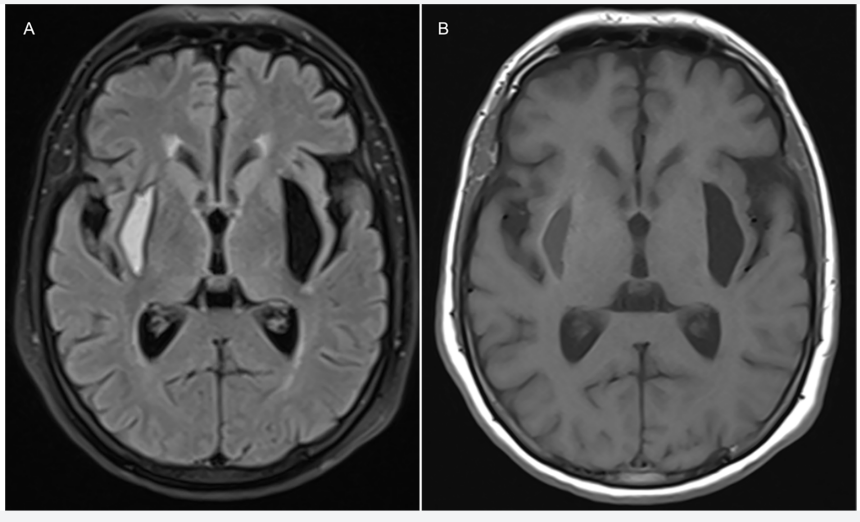


Figure 1. Brain magnetic resonance imaging demonstrated a T2-FLAIR hyperintensity in the right putamen consistent with a subacute hematoma (A), as well as bilateral diffuse T1 hypointensities in the putaminal white matter (B)

due to acute midbrain infarctions have also been reported[14]. These vascular territories encompass the basal ganglia structures, playing a critical role in the pathogenesis of post-stroke movement disorders. Furthermore, hemorrhagic strokes are emphasized in the literature to have a higher propensity to cause movement disorders compared with ischemic strokes likely due to more prominent tissue damage and inflammatory responses in the basal ganglia and extrapyramidal system[6]. In our patient, lesions in the bilateral putamen explained the acute-subacute course of parkinsonism.

In stroke survivors, it is crucial to recognize extrapyramidal involvement in addition to classic pyramidal signs. Hypokinetic movement disorders such as parkinsonism, which considerably impact quality of life, should not be overlooked. Treatment plans including both pharmacological interventions and rehabilitation strategies must be individualized to optimize patient outcomes.

Ethics

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

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Tater P, Pandey S. Post-stroke movement disorders: clinical spectrum, pathogenesis, and management. Neurol India. 2021;69:272-83.
2. Alarcón F, Zijlmans JC, Dueñas G, Cevallos N. Post-stroke movement disorders: report of 56 patients. J Neurol Neurosurg Psychiatry. 2004;75:1568-74.
3. Ghika-Schmid F, Ghika J, Regli F, Bogousslavsky J. Hyperkinetic movement disorders during and after acute stroke: the Lausanne Stroke Registry. J Neurol Sci. 1997;146:109-16.
4. Handley A, Medcalf P, Hellier K, Dutta D. Movement disorders after stroke. Age Ageing. 2009;38:260-6.
5. Mehanna R, Jankovic J. Movement disorders in cerebrovascular disease. Lancet Neurol. 2013;12:597-608.
6. Chuang C, Fahn S, Frucht SJ. The natural history and treatment of acquired hemidystonia: report of 33 cases and review of the literature. J Neurol Neurosurg Psychiatry. 2002;72:59-67.
7. Kim JS. Delayed onset mixed involuntary movements after thalamic stroke: clinical, radiological and pathophysiological findings. Brain. 2001;124:299-309.
8. Jankovic J. Parkinson's disease: clinical features and diagnosis. J Neurol Neurosurg Psychiatry. 2008;79:368-76.
9. Adams RD, Victor M, Ropper AH. Adams and Victor's Principles of Neurology. McGraw-Hill Education. 2019;11.
10. Freedman M, Sarkar G, Hill MD. Movement disorders following cerebrovascular events. Can J Neurol Sci. 2010;37:234-8.
11. Bansil S, Prakash N, Kaye J, Wrigley S, Manata C, Stevens-Haas C, et al. Movement disorders after stroke in adults: a review. Tremor Other Hyperkinet Mov (N Y). 2012;2:tre-02-42-195-1.
12. Hatano T, Kubo S, Niijima-Ishii Y, Hattori N, Sugita Y. Levodopa-responsive parkinsonism following bilateral putaminal hemorrhages. Parkinsonism Relat Disord. 2013;19:477-9.
13. Bogousslavsky, J. Post-stroke movement disorders: a clinical and imaging study. European Neurology. 2003;49:133-9.
14. Friedman-Korn T, Weill C, Ben-Haim S, Arkadir D. Delayed levodopa-responsive parkinsonism following acute midbrain injury. J Neurol Sci. 2024;459:122983.

ERRATUM

DOI: 10.4274/nkmj.galenos.2025.e001

Nam Kem Med J 2025;13(3):352

Sera ÇETİNGÖK, Hatice Selin IRMAK. Psychological Resilience and Adaptation Difficulties in Older Adults.

Nam Kem Med J 2022;10:392-400

The mistake has been made inadvertently by the author.

In the Abstract section on page 392 of the article, the Materials and Methods part has been revised by the author as follows.

The incorrect part of the Materials and Methods section

Gereç ve Yöntem: Araştırma, İstanbul'da yaşayan 60 yaş ve üzeri 200 kişi ile yapıldı.

Materials and Methods: The study was conducted with 200 individuals aged 60 years and above, and living in İstanbul.

Materials and Methods section correct version

Gereç ve Yöntem: Araştırma, İstanbul ve Afyon'da yaşayan 60 yaş ve üzeri 200 kişi ile yapıldı.

Materials and Methods: The study was conducted with 200 individuals aged 60 years and above, and living in İstanbul and Afyon.

The sub-section titled Study Sampling within the Materials and Methods section on page 393 has been revised by the author.

The incorrect part of the Study Sampling within the Materials and Methods section

Study Sampling

The study universe consisted of 200 individuals aged ≥60 years and living in İstanbul.

Study Sampling section correct version

Study Sampling

The research population consisted of individuals aged 60 and over living in the Çatalca district of Istanbul province and the Bolvadin district of Afyon province.

An Acknowledgments section has been added on page 399.

Acknowledgements

We would like to thank the Gerontology Department students who contributed to this study as part of their final assignment course.