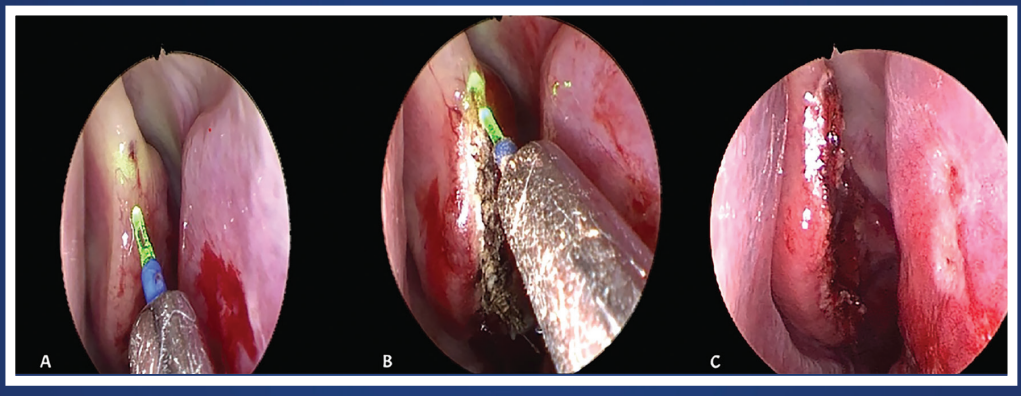
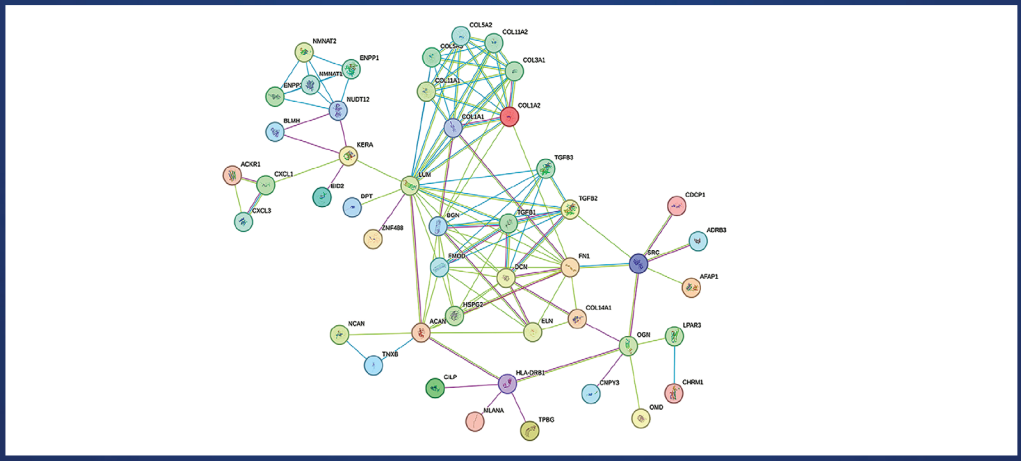
E-ISSN: 3023-6215

ANATOLIAN

JOURNAL of GENERAL MEDICAL RESEARCH

Volume/Cilt: 35 Issue/Sayı: 3 Aralık/December 2025





anatolianjmed.org

galenos yayınevi

ANATOLIAN
JOURNAL of GENERAL MEDICAL RESEARCH

anatolianjmed.org

EDITORIAL BOARD

The Anatolian Journal of General Medical Research is the official journal of the İzmir Tepecik Education and Research Hospital.

It is published three times a year (April, August, December).

Anatolian Journal of General Medical Research is an open Access, free and peer-reviewed journal and indexed in **Tubitak / ULAKBIM, EBSCO, Türk Medline** and **Türkiye Atıf Dizini**.

galenos yayınevi

Publisher Contact
Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1
34093 İstanbul, Türkiye
Phone: +90 (539) 307 32 03
E-mail: info@galenos.com.tr/yayin@galenos.com.tr
Web: www.galenos.com.tr **Publisher Certificate Number:** 14521

Online Publication Date: December 2025

E-ISSN: 3023-6215

International periodical journal published three times in a year.

ANATOLIAN
JOURNAL of GENERAL MEDICAL RESEARCH

ADVISORY BOARD

anatolianjmed.org

ANATOLIAN
JOURNAL of GENERAL MEDICAL RESEARCH

ADVISORY BOARD

anatolianjmed.org

ANATOLIAN
JOURNAL of GENERAL MEDICAL RESEARCH

ADVISORY BOARD

anatolianjmed.org

Please refer to the journal's webpage (https://tepecikdergisi.com/) for "Aims and Scope", "Instructions For Authors" and "Ethical Policy".

The Anatolian Journal of General Medical Research and/or its editors are members of ICMJE, COPE, WAME, CSE and EASE, and follow their recommendations.

The Anatolian Journal of General Medical Research is indexed in **Tubitak / ULAKBIM, EBSCO, Türk Medline, Türkiye Atıf Dizini**.

ANATOLIAN
JOURNAL of GENERAL MEDICAL RESEARCH

CONTENTS / İÇİNDEKİLER

anatolianjmed.org

CLINICAL RESEARCHES / KLİNİK ARAŞTIRMALAR

ANATOLIAN
JOURNAL of GENERAL
MEDICAL RESEARCH

CONTENTS / İÇİNDEKİLER

anatolianjmed.org

DOI: 10.4274/anatoljmed.2025.63825
Anatol J Gen Med Res 2025;35(3):247-252

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Neutrophil/Lymphocyte Ratio as a Marker of Infections in Patients Hospitalized in the Oncology Palliative Care Service

Onkoloji Palyatif Bakım Servisi'nde Yatan Hastalarda Enfeksiyon Belirteci Olarak Nötrofil/Lenfosit Oranı

Muhammed Mustafa Uzan, Umut Gök Balcı, Tevfik Tanju Yılmazer

University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Clinic of Family Medicine, İzmir, Türkiye

Cite as: Uzan MM, Gök Balcı U, Yılmazer TT. Neutrophil/lymphocyte ratio as a marker of infections in patients hospitalized in the oncology palliative care service. Anatol J Gen Med Res. 2025;35(3):247-252

Abstract

Objective: We aimed to evaluate the correlation of procalcitonin level with neutrophil to lymphocyte ratio (NLR) and to establish a cut-off value that can be used as an indicator parameter for oncological inpatient.

Methods: Study design is a cross sectional. The study was executed retrospectively on 640 patients with at least one oncological diagnosis. Demographic data, serum albumin, C-reactive protein (CRP), procalcitonin levels, and complete blood count data on the day of admission were attained.

Results: The average procalcitonin level was found to be 63.47% between 01/11/2021 and 30/01/2022 in the Clinic of Oncological Palliative Care, University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital. The NLR cut-off value is 6.97 with sensitivity and specificity of 68% and 64%, respectively. There was a medium-level positive correlation between procalcitonin with leukocyte and neutrophil number (r=0.419).

Conclusion: The NLR parameter, which can be determined through complete blood count, can be used nearly as effectively as albumin or CRP for the detection of infection. This approach might prevent unnecessary antibiotic usage, thereby affecting the cost-effectiveness issue positively.

Keywords: Procalcitonin, family-medicine, oncology

Öz

Amaç: Amacımız prokalsitonin düzeyinin nötrofil/lenfosit oranı (NLR) ile ilişkisini değerlendirmek ve onkolojik yatan hastalar için gösterge parametresi olarak kullanılabilecek bir kesim değeri oluşturmaktır.

Yöntem: Çalışma tasarımı kesitsel bir çalışmadır. Çalışma, en az bir onkolojik tanısı olan 640 hasta üzerinde retrospektif olarak gerçekleştirildi. Yatış günündeki demografik veriler, serum albümin, C-reaktif protein (CRP), prokalsitonin düzeyleri ve tam kan sayımı verileri elde edildi.

Bulgular: Sağlık Bilimleri Üniversitesi, İzmir Tepecik Eğitim ve Araştırma Hastanesi, Onkolojik Palyatif Kliniği'nde 01/11/2021 ile 30/01/2022 tarihleri arasında ortalama prokalsitonin düzeyi %63,47 idi. NLR kesme değeri 6,97 olup duyarlılık ve özgüllük sırasıyla %68 ve %64'tür. Prokalsitonin ile lökosit ve nötrofil sayısı arasında orta düzeyde pozitif bir korelasyon vardı (r=0,419).

Address for Correspondence/Yazışma Adresi: Muhammed Mustafa Uzan, MD, University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Clinic of Family Medicine, İzmir, Türkiye
E-mail: mustafauzan65@gmail.com
ORCID ID: orcid.org/0000-0002-2111-4520

Received/Geliş tarihi: 16.01.2025
Accepted/Kabul tarihi: 04.05.2025
Epub: 11.06.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

Sonuç: Tam kan sayımı ile belirlenebilen NLR parametresi, enfeksiyon tespitinde en az albumin veya CRP kadar etkin olarak kullanılabilir. Böylece gereksiz antibiyotik kullanımının önüne geçilebilir ve maliyet-etkinlik konusu olumlu yönde etkilenebilir.

Anahtar Kelimeler: Prokalsitonin, aile hekimliği, onkoloji

Introduction

Procalcitonin is an acute-phase protein that has been frequently used in recent years for the early diagnosis of bacterial infections, especially septicemia[1]. Procalcitonin comprises mainly the metabolization of thyroid hormone and secretion from adipose tissues and the lymphocytes.

Calcitonin is a peptide that is derived from a precursor protein. Serum levels start to rise within 24 hours during bacterial infections due to endotoxin production from parenchymal tissues. The level also rises post-operatively, after burns, heat stroke, and traumas[2]. Suggestive levels of septicemia are 2 ng/mL, and those of septic shock are 11.6 ng/mL. If the measurement is >5 ng/mL for patients who are known to be in shock status, the reason might be considered to result from an infection[3].

Procalcitonin might be accepted as an indication for the use of antibiotics in cases of bacteremia while waiting for the result of culture[4]. Laboratory parameters such as C-reactive protein (CRP), leucocyte, and neutrophil levels, can be evaluated jointly[5,6].

CRP is the prototype acute phase reactant of the pentraxin group. The serum levels might increase in 24 to 48 hours during inflammation or infection, and also decrease rapidly[7]. Normal levels of CRP are between 0.3 and 1.7 mg/L[8]. The increase in the serum level to 10 mg/dL is a sign of bacterial infection. The levels can also be increased in situations such as dental illnesses, hypertension, obesity, smoking, diabetes mellitus, exercise, cancer, chronic fatigue, senility, alcohol consumption, and depression[9]. The serum levels can be measured turbidometrically and nephelometrically[10].

Albumin, which constitutes 60% of plasma proteins, is a negative acute-phase protein. The serum and plasma levels decrease during infection, contrary to procalcitonin and CRP[11]. Its levels also decrease in situations like lasting starvation, liver dysfunction, and intestinal malabsorption. Albumin is responsible for providing plasma oncotic pressure. It can also be used to evaluate the nutritional status of individuals. Its normal serum levels are 3.5 to 5.0 g/dL[12,13].

Neutrophil to lymphocyte ratio (NLR) is one of the recently proposed parameters for early diagnosis of septicemia [5,14].

A complete blood count is a test that can be performed at almost any health facility, whereas a procalcitonin test cannot be implemented at every health institution due to its high cost.

The aim in this study was to evaluate the correlation of procalcitonin with NLR and also to establish a threshold level that can be used at early diagnosis during the courses of bacteremia and early phase septicemia. Additionally, the study aims to promote the NLR as an indicator parameter for inpatient clinics, especially for oncological ones where the risk for infection is increased.

Materials and Methods

Ethical approval was obtained from the Non-interventional Ethics Committee of University of Health Sciences of Türkiye, İzmir Tepecik Education and Research Hospital (decision no: 2021/10-28, date: 15.10.2021). The study retrospectively collected data on 640 patients treated at the care clinic in the hospital between November 2021 and January 2022. All the patients had at least one oncological diagnosis.

Demographic data, serum albumin, CRP, procalcitonin levels, and complete blood count data were obtained on the day of hospitalization. The cut-off levels of the biochemical and microbiological laboratories of the study center were procalcitonin: 0.06 ng/dL, albumin: 3.5 g/dL, white blood cell: 10.600/mm^3, neutrophil: 7.000/mm^3, CRP: 6.0 mg/L, platelet: 140.000/mm^3.

Statistical Analysis

Data were analyzed with SPSS 26. Continuous variables were presented as mean and standard deviation, and categorical variables were presented as numbers and percentages. Continuous variables were tested for suitability for normal distribution. The Q-Q plot method and sample size were taken into consideration, and it was observed that they did not comply with the normal distribution assumption. Correlation between variables was investigated using the Spearman's rho

method. Independent group comparisons were made using Kruskal-Wallis and Mann-Whitney U tests, with a Bonferroni correction applied for multiple comparisons. Receiver operating characteristic (ROC) analysis was performed on variables that showed significant differences, and the most appropriate cut-off values were determined according to the Youden index. Binary variables were created for these values. Distribution differences of categorical independent variables were compared with the chi-square test, while the univariate odds ratio was calculated. Type 1 error was evaluated at an a level of 0.05 with a two-tailed test.

Results

The study was carried out with 640 patients; of these, 44.53% (n=285) were male. The percentage of individuals who were 65 years of age and over was 50.78% (n=325). The most prevalent oncological disease of the patients was breast cancer with 15.75% (n=97); the second most prevalent was stomach cancer with 7.03% (n=45).

The average procalcitonin level of the patients was 63.47 ng/mL. The averages for white blood cell number were 9.780/mm^3; for neutrophil number, they were 7.830/mm^3; and lymphocyte numbers, they were 1.130/mm^3 (Table 1).

There was a medium-level positive correlation between procalcitonin and white blood cells and neutrophil number (r=0.419), a medium-level positive correlation between procalcitonin and NLR (r=0.365) (Table 2).

According to the ROC analysis, the area under the curve used to determine a cut-off value for NLR was calculated to be 0.697. Therefore, the cut-off was determined as 6.00. This condition was recorded with an ideal sensitivity of 68% and a specificity of 64% (Figure 1).

Discussion

Increases in procalcitonin levels, may be mistakenly interpreted as a positive sign of infection. Sometimes, overuse of antibiotics can mask an infection despite low procalcitonin levels. Therefore, it may be wise to set a cut-off value. Eraldemir[1] has intercrossed procalcitonin levels at cut-off values 0.1, 1, 2, and 5 ng/mL in separate groups with NLR, and he identified cut-off values for NLR to be 4.71, 6.44, 7.65, and 8.24 in the respective order. When 4.71 was used as the cut-off value for NLR, he determined the sensitivity to be 58.16% and the specificity to be 78.76%. Önmez et al.[15] have studied patients with pancreatitis and specified that the highest correlation of NLR parameter with the severity of pancreatitis was 68% in the ROC analysis, with an area under the ROC curve of 0.687 (confidence interval: 0.570-0.804). Furthermore, the sensitivity of this value was determined to be 61% and the cut-off value has been confirmed to be 6.0451[15]. Another study reported that NLR with a cut-off value of 4.68 can significantly differentiate between normal and inflamed appendices[16]. Sayah et al.[17] have analyzed the potential fatality of coronavirus disease 2019. According to this, interleukin-6 has been shown to have a positive correlation with NLR where the cut-off value was 7.4, sensitivity was 96.3%, and specificity was 70.5%. Çil et al.[18] have specified NLR to be 7.21 in patients with pneumonia who had high procalcitonin levels. Our study revealed the NLR cut-off value to be 6.97 using ROC analysis, with sensitivity and specificity at 68% and 64% respectively. Therefore, we

Table 1. Procalcitonin, albumin, CRP and complete blood count parameters

	n	Mean	Standard deviation	Median	Minimum	Maximum
PCT	640	3.58	11.74	0.11	0.01	75
WBC	640	9.78	8.18	7.5	0.1	84.4
NEU	640	7.83	7.51	5.4	0	77.7
LYM	640	1.13	1.1	0.9	0	10.6
N/L	640	11.09	13.72	6.43	0	121.5
HGB	640	10.32	1.94	10.2	2.6	15.8
PLT	640	235.45	137.4	218	5	943
MPV	640	8.24	1.12	8.1	5.5	13.3
CRP	640	47.03	77.35	12.05	0.01	514.2
ALB	640	2.9	0.75	2.8	0.9	4.73

PCT: Procalcitonin, WBC: White blood cell, NEU: Neutrophil, LYM: Lymphocyte, N/L: Neutrophile to lymphocyte ratio, HGB: Hemoglobin, PLT: Platelet, MPV: Mean platelet volume, CRP: C-reactive protein, ALB: Albumin

Table 2. Procalcitonin and its correlation with other parameters

		PCT	NEU	LYM	NLR	CRP	ALB
PCT	Correlation coefficient	-	0.436**	-0.052	0.367%*	0.667**	-0.722**
	Sig. (2-tailed)		<0.001	0.186	<0.001	<0.001	<0.001
NEU	Correlation coefficient	0.436**	-	0.175**	0.620**	0.353**	0.404**
	Sig. (2-tailed)	<0.001		<0.001	<0.001	<0.001	<0.001
LYM	Correlation coefficient	0.052	0.175**	-	-0.598**	-0.093*	0.046
	Sig. (2-tailed)	0.186	<0.001		<0.001	0.019	0.244
NLR	Correlation coefficient	0.367**	0.620**	-0.598**	-	0.351**	-0.344**
	Sig. (2-tailed)	<0.001	<0.001	<0.001		<0.001	<0.001
CRP	Correlation coefficient	0.667**	0.353**	-0.093	0.351**	-	-0.662**
	Sig. (2-tailed)	<0.001	<0.001	0.019	<0.001		<0.001
ALB	Correlation coefficient	-0.722**	-0.404**	0.046	-0.344**	-0.662**	-
	Sig. (2-tailed)	<0.001	<0.001	0.244	<0.001	<0.001	

*: Correlation is significant at the 0.05 level (2-tailed), **: Correlation is significant at the 0.01 level (2-tailed), PCT: Procalcitonin, NEU: Neutrophil, LYM: Lymphocyte, N/L: Neutrophile to lymphocyte ratio, CRP: C-reactive protein, ALB: Albumin

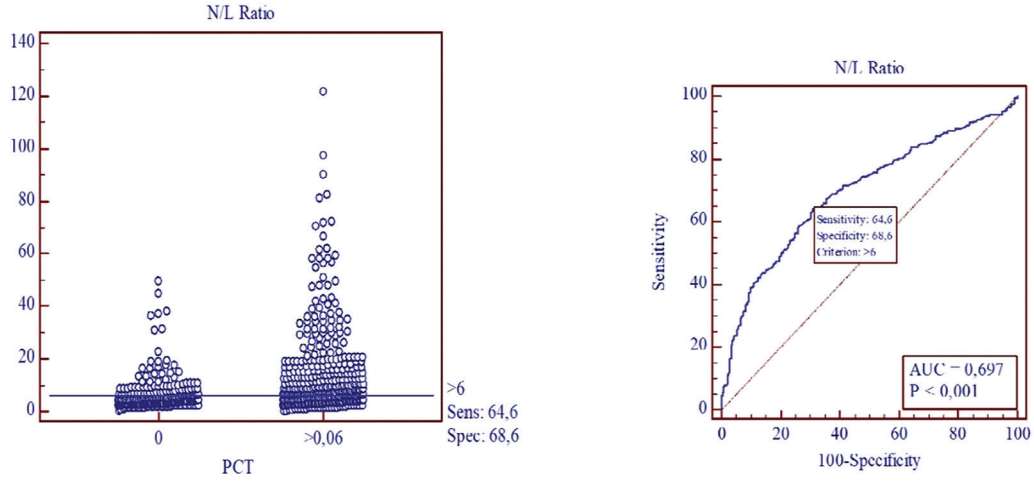


Figure 1. ROC curve cut-off value N/L ratio and associated criterion >6 for Youden index

ROC: Receiver operating characteristic, N/L: Neutrophile to lymphocyte ratio

consider these values to be self-consistent and to have high predictive value for septicemia.

The lymphocyte number decreases while neutrophils increase during systemic inflammatory response and septicemia, thereby providing credibility for NLR increase in its medical implementation[19]. de Jager et al.[20] revealed that NLR increase is more valuable in the early diagnosis of bacteremia at the emergency service compared to neutrophil and leucocyte increase. Beyazit et al.[21] revealed that NLR is superior to leucocytes in conditions of bacteremia and septicemia. We identified a medium-level positive correlation between procalcitonin and NLR. Therefore, we can consider that the implementation of NLR instead of procalcitonin, at least for oncological patients, might prevent unnecessary antibiotic usage and positively affect cost-effectiveness. One must bear in mind that a complete blood count provides an advantage as it is both fast and low-cost.

Tekin et al.[22] revealed that in patients presenting with high fever, CRP levels increase significantly with bacterial infection compared to those with viral. Yıldız Celtek et al.[23] found CRP levels to be significantly higher in patients. Ljungström et al.[5] determined that single use of NLR or procalcitonin has similar diagnostic power; nevertheless, using NLR with CRP acts as a confirmatory factor.

Arikan et al.[24] reported that there was a significant difference between all groups in NLR and platelet to lymphocyte ratio values[25]. In a study conducted in Wuhan, it was determined that high values of NLR had high sensitivity and specificity in predicting mortality[26]. There are studies in the literature in which NLR is successful in predicting sepsis and mortality. In this sense, this finding supports our study by corroborating our hypothesis.

Study Limitation

The most important limitation of our study is in investigating the presence of clinical variability changes, and it being conducted in a single center. Therefore, studies with larger samples are needed.

Conclusion

Both the cost and unattainability of procalcitonin at some health institutions directs clinicians to use alternative infection markers. We consider that the NLR parameter, which can be determined through complete blood count, can be used nearly as effectively as procalcitonin for the detection of infection. Further studies are needed to conduct more multicenter research, especially those without infection variability.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Non-interventional Ethics Committee of University of Health Sciences of Türkiye, İzmir Tepecik Education and Research Hospital (decision no: 2021/10-28, date: 15.10.2021).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

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

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

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

References

1. Eraldemir FC. Neutrophil/lymphocyte ratio instead of procalcitonin: is it an option? Türk Klinik Biyokimya Derg. 2018;16:11-6.
2. Lee H. Procalcitonin as a biomarker of infectious diseases. Korean J Intern Med. 2013;28:285-91.
3. Mitaka C. Clinical laboratory differentiation of infectious versus non-infectious systemic inflammatory response syndrome. Clin Chim Acta. 2005;351:17-29.
4. Günal Ö, Barut HS. Sepsis and procalcitonin. CMJ. 2009;31:502-12.
5. Ljungström L, Pernestig AK, Jacobsson G, Andersson R, Usener B, Tilevik D. Diagnostic accuracy of procalcitonin, neutrophil-lymphocyte count ratio, C-reactive protein, and lactate in patients with suspected bacterial sepsis. PLoS One. 2017;20:7.
6. Kim SY, Jeong TD, Lee W, Chun S, Min WK. Procalcitonin in the assessment of bacteremia in emergency department patients: results of a large retrospective study. Ann Clin Biochem. 2015;52:654-9.
7. Pepys MB. C-reactive protein fifty years on. Lancet. 1981;1:1653-7.
8. Baumann H, Gauldie J. The acute phase response. Immunol Today. 1994;15:74-80.
9. Hacimustafaoglu M. Acute phase reactants (erythrocyte sedimentation rate, CRP). J Pediatr Inf. 2017;11:53-5.
10. Büyükberber N, Sevinç A. İnfeksiyon ve kanserde C-reaktif protein. Arşiv. 2003;12:149.
11. Batırel A, Gençer S, Özer S. Enfeksiyon göstergesi olarak akut faz reaktanları: C-reaktif protein (CRP) ve serum amiloid A (SAA). KEAH. 2003;14:220-4.
12. Ceron JJ, Eckersall PD, Martinez-Subiela S. Acute phase proteins in dogs and cats: current knowledge and future perspectives. Vet Clin Pathol. 2005;34:85-99.
13. Georgieva TM, Andonova MJ, Slavov EP, Dzhelebov PV, Zapryanova DS, Georgiev IP. Blood serum protein profiles and lysozyme activity in dogs during experimental infection with staphylococcus intermedius. Revue Méd Vét. 2011;162:580-5.
14. Can E, Hamilcikan S, Can C. The value of neutrophil to lymphocyte ratio and platelet to lymphocyte ratio for detecting early-onset neonatal sepsis. J Pediatr Hematol Oncol. 2018;40:229-32.
15. Önmez A, Bilir E , Torun S. The relationship between acute pancreatitis severity of platelet lymphocyte ratio, neutrophil lymphocyte ratio, erythrocyte distribution width and mean platelet volume. Konuralp Medical Journal. 2019;11:24-9.
16. Kahramanca S, Ozgehan G, Seker D, et al. Neutrophil-to-lymphocyte ratio as a predictor of acute appendicitis. Ulus Travma Acil Cerrahi Derg. 2014;20:19-22.
17. Sayah W, Berkane I, Guermache I, et al. Interleukin-6, procalcitonin and neutrophil-to-lymphocyte ratio: potential immune-inflammatory parameters to identify severe and fatal forms of COVID-19. Cytokine. 2021;141:155428.
18. Çil E, Karadeniz G, Yenigün S, Çoraplı G. Evaluation of the relationships between procalcitonin and neutrophil/lymphocyte ratio and platelet/lymphocyte ratio in patients with pneumonia. Eur Rev Med Pharmacol Sci. 2022;26:3200-5.

19. Zahorec R. Ratio of neutrophil to lymphocyte counts--rapid and simple parameter of systemic inflammation and stress in critically ill. Bratisl Lek Listy. 2001;102:5-14.
20. de Jager CP, van Wijk PT, Mathoera RB, de Jongh-Leuvenink J, van der Poll T, Wever PC. Lymphocytopenia and neutrophil-lymphocyte count ratio predict bacteremia better than conventional infection markers in an emergency care unit. Crit Care. 2010;14:R192.
21. Beyazit Y, Sayilir A, Torun S, et al. Mean platelet volume as an indicator of disease severity in patients with acute pancreatitis. Clin Res Hepatol Gastroenterol. 2012;36:162-8.
22. Tekin M, Çalışkan M, Kayak D, Konca Ç. Is the C-reactive protein a reliable marker for the diagnosis of bacterial infection? ADYÜ Sağlık Bilimleri Derg. 2017;3:475-86.
23. Yıldız Celtek N, Ünlü U, Karaaslan E, Demir O. Can neutrophil lymphocyte ratio be an evaluation criterion for hospitalization in children with fever? Pediatr Pract Res. 2020;8:75-8.
24. Arikan MG, Akgul M, Akdeniz E, Iskan G, Arda E. The value of hematological inflammatory parameters in the differential diagnosis of testicular torsion and epididymorchitis. Am J Clin Exp Urol. 2021;9:96-100.
25. Yığman M, Ekenci BY, Durak HM, Sanci A. Predictive value of hematologic parameters in the differential diagnosis of testicular torsion and epididymo-orchitis. Androl Bul. 2024;26:167-72.
26. Wang X, Li X, Shang Y, et al. Ratios of neutrophil-to-lymphocyte and platelet-to-lymphocyte predict all-cause mortality in inpatients with coronavirus disease 2019 (COVID-19): a retrospective cohort study in a single medical centre. Epidemiol Infect. 2020;148:e211.

DOI: 10.4274/anatoljmed.2025.74946
Anatol J Gen Med Res 2025;35(3):253-259

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

The Long-term Follow-up of Axillary Lymphadenopathies Developed After COVID-19 Vaccination and Their Relationship with Types of Vaccines

COVID-19 Aşılaması Sonrası Gelişen Aksiller Lenfadenopatilerin Uzun Dönem Takibi ve Aşı Türleriyle İlişkisi

Sevgi Ünal[1], Merve Gürsoy[1], Özge Pasin[2]

[1]İzmir Atatürk Training and Research Hospital, Clinic of Radiology, İzmir, Türkiye
[2]University of Health Sciences Türkiye Hamidiye Faculty of Medicine, Department of Biostatistics, İstanbul, Türkiye

Cite as: Ünal S, Gürsoy M, Pasin Ö. The long-term follow-up of axillary lymphadenopathies developed after COVID-19 vaccination and their relationship with types of vaccines. Anatol J Gen Med Res. 2025;35(3):253-259

Abstract

Objective: To evaluate the long-term effects of the coronavirus disease-2019 vaccines on axillary lymph nodes (LN) and to examine the relationship between LN size, cortical thickness, and other morphological characteristics.

Methods: A total of 182 patients were included in this study. Data on the size and cortical thickness of axillary LN on both sides, the vaccinated arm, the number and type of vaccines, patient height and weight was recorded. The relationships between LN size, cortical thickness, additional variables, and time were analyzed.

Results: No significant relationship was found between the number of vaccine doses and LN size or cortical thickness. The vaccinated arm was the left arm. There was no significant relationship between the size of the left axillary LN and the right axillary LN. A significant increase in axillary LN size was observed correlated with body mass index ($p<0.0001$). When assessing the impact of the vaccine, a significant decrease in cortical thickness of the left axillary LN was observed in the patients presenting two weeks after vaccination ($p=0.041$). However, in the overall evaluation, the increase in left axillary LN cortical thickness was found to be significantly greater than the right axillary LN ($p=0.049$). When examining the time relationship (maximum 48 weeks), increase in left axillary LN cortex thickness was detected ($p=0.437$).

Conclusion: Our study demonstrated that cortical thickening of LNs persisted for an extended period (up to 48 weeks). This condition creates diagnostic challenges in patients with suspicious breast cancer findings, those under follow-up for a history of breast cancer, and in those who are newly diagnosed with breast cancer. Therefore, recording the vaccination history and comparing LN imaging before and after vaccination may help prevent unnecessary invasive procedures and misdiagnoses.

Keywords: Breast cancer, vaccine, aksillary lymphadenopathy

Öz

Amaç: Koronavirüs hastalığı-2019 aşısının aksiller lenf nodları (LN) üzerindeki uzun vadeli etkilerini değerlendirmek ve LN boyutu, korteks kalınlığı ve diğer morfolojik özellikler arasındaki ilişkiyi incelemektir.

Address for Correspondence/Yazışma Adresi: Sevgi Ünal, MD, İzmir Atatürk Training and Research Hospital, Clinic of Radiology, İzmir, Türkiye
E-mail: miraderinn2014@gmail.com
ORCID ID: orcid.org/0009-0000-0373-0904

Received/Geliş tarihi: 27.07.2025
Accepted/Kabul tarihi: 23.08.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

Yöntem: Bu çalışmaya toplam 182 hasta dahil edildi. Her iki taraftaki aksiller LN boyutu ve korteks kalınlığı, aşılanan kol, aşı sayısı ve türü, hasta boyu ve kilosu kaydedildi. LN boyutu, korteks kalınlığı, diğer değişkenler ve zaman arasındaki ilişkiler analiz edildi.

Bulgular: Aşı dozu sayısı ile LN boyutu veya korteks kalınlığı arasında anlamlı bir ilişki bulunmadı. Aşılanan kol sol koldu. Sol aksiller LN boyutu ile sağ aksiller LN boyutu arasında anlamlı bir ilişki bulunmadı. Aksiller LN boyutunda, vücut kitle indeksi ile korelasyon içinde anlamlı bir artış gözlendi (p<0,0001). Aşının etkisini değerlendirirken, aşılamadan iki hafta sonra başvuran hastalarda sol aksiller LN korteks kalınlığında anlamlı bir azalma gözlendi (p=0,041). Ancak genel değerlendirmede, sol aksiller LN korteks kalınlığındaki artışın sağ aksiller LN'ye kıyasla anlamlı derecede daha yüksek olduğu bulundu (p=0,049). Zaman ilişkisi incelendiğinde (maksimum 48 hafta), sol aksiller LN korteks kalınlığında anlamlı bir artış tespit edildi (p=0,437).

Sonuç: Çalışmamız, LN korteks kalınlaşmasının uzun bir süre (48 haftaya kadar) devam ettiğini göstermiştir. Bu durum, şüpheli meme kanseri bulguları olan hastalarda, meme kanseri öyküsü nedeniyle takip edilen hastalarda ve yeni tanı almış meme kanseri hastalarında tanı zorlukları yaratmaktadır. Bu nedenle, aşılama öyküsünün kaydedilmesi ve aşılama öncesi ve sonrası LN görüntülemelerinin karşılaştırılması, gereksiz invaziv prosedürlerin ve yanlış tanıların önlenmesine yardımcı olabilir.

Anahtar Kelimeler: Meme kanseri, aşı, aksiller lenfadenopati

Introduction

Pneumonia associated with coronavirus disease-2019 (COVID-19) was first reported in December 2019 in Wuhan, China[1]. The pandemic later reached a global scale and became the focus of worldwide attention. Various COVID-19 vaccine technologies, including mRNA (Moderna, Pfizer-BioNTech)[2], inactivated (Sinovac), recombinant protein (novavax), and vector-based (Janssen, Oxford-AstraZeneca) vaccines, were successfully developed and used during the COVID-19 pandemic[3]. Bar-On et al.[4] demonstrated the effectiveness of the third dose in protecting individuals aged 60 and above from COVID-19. Spitzer et al.[5] also reported the effectiveness of the third dose in individuals aged 18 and older. Therefore, the third dose of the vaccine became a routine practice in many countries worldwide. However, as vaccination rates increased, some vaccinated individuals experienced adverse reactions. One such reaction was ipsilateral axillary lymphadenopathy. Some researchers have stated that this condition is a short-term and harmless response, indicating vaccine effectiveness[6-8]. Axillary lymphadenopathy associated with vaccination can pose challenges in the treatment planning and prognosis assessment of patients diagnosed with or suspected of having breast cancer. In routine practice, this condition can also cause difficulties in the follow-up of patients with a history of breast cancer surgery, particularly if the lymph nodes (LN) size does not decrease over time or if signal enhancement persists on an 18-fluorodeoxyglucose (18F-FDG) positron emission tomography/computed tomography (PET/CT) scan. These issues lead to repeated interventional procedures to clarify the clinical picture. The aim of this study is to investigate the long-term effects of the vaccine on axillary LNs by correlating the type of vaccine administered and the time elapsed since vaccination.

Materials and Methods

Ethical approval for this retrospective study was obtained from the Non-Interventional Clinical Research Ethics Committee of İzmir Katip Çelebi University (approval no: 0043, date: 15.02.2024).

Patient Selection

Turkey initiated its COVID-19 vaccination program on January 13, 2021. Therefore, a total of 198 patients with a history of vaccination who presented to the breast imaging unit between January 2021 and March 2022 were reviewed. Two patients were excluded from the study due to newly detected malignant breast masses, and histopathologically confirmed axillary LN metastases. Fourteen patients were also excluded due to a left mastectomy and vaccination in the right arm. The remaining 182 patients who received the vaccine in the left arm were included in the study.

Radiological Evaluation

All patients included in the study underwent routine breast and axillary ultrasonography. The ultrasonographic examination was performed by a radiologist with five years of experience in breast imaging. Patient age, weight, and height were recorded. During the sonographic evaluation, the largest LNs in both axillae were measured in terms of long-axis diameter and cortical thickness. In follow-up ultrasonographic examinations, changes in LN size and

biopsy results if performed, were also assessed. Additionally, data on the number of vaccine doses received, the types of vaccines administered, and the arm in which the vaccine was given were recorded.

Statistical Analysis

Statistical analyses were performed using the SPSS (version 26) software package. Descriptive statistics for quantitative variables were presented as mean, standard deviation, median, minimum, and maximum values. The normality of quantitative variables was assessed using the Shapiro-Wilk test. The Wilcoxon signed-rank test was used to compare medians between two dependent groups. Pearson correlation analysis was used to evaluate relationships between quantitative variables. A p-value of <0.05 was considered statistically significant.

Results

Fine-needle aspiration biopsy (FNAB) was performed on the axillary LNs of eight out of 182 patients in the study group. Two patients were excluded from the study due to malignant histopathology findings. In the remaining six patients, histopathology results confirmed reactive LNs. One year after inclusion in the study, pathological sonographic findings persisted in the axillary LNs of three patients. Two of these patients had a known history of malignancy, and the vaccine had been administered in the contralateral axilla. Due to high standardized uptake value $(SUV)_{max}$ values in 18F-FDG PET/CT scans, FNAB was performed on both patients. Since FNAB results were benign and pathological LNs persisted in follow-up examinations, core biopsy was performed. The biopsy results confirmed benign/reactive LNs. The patients with increased LN size and SUV_{max} values in 18F-FDG PET/CT scans are still being monitored. In one patient without a history of malignancy, FNAB results were benign. However, the cortical thickening of the axillary LN persisted at the 18-month follow-up after the last vaccination, and the patient remains under surveillance. Patient data, including age, weight, body mass index (BMI), vaccine type, and number of doses, are presented in Table 1. Sonographic evaluation of LNs showed no statistically significant difference between the mean long-axis diameters of the right and left axillary LNs (p=0.109).

However, the mean cortical thickness of the left axillary LNs was significantly higher than that of the right axillary LNs (p=0.049) (Table 2).

Comparisons of LN size and cortical thickening among different vaccine types revealed no statistically significant differences between the right and left axillary LNs (Table 3). When assessing the relationship between LN size and other variables, BMI was found to be significantly correlated with the sizes of both right and left axillary LNs, with the right axillary LN (r=0.343; p<0.001) and the left axillary LN (r=0.343; p<0.001) (Table 3). Regarding the relationship between time elapsed since the last vaccine dose and sonographic findings, cortical thickening of the left axillary LN significantly decreased after the second week post-vaccination, (p=0.041) (Table 4). The measured parameters included time since the last vaccination (1 to 48 weeks), left axillary LN cortical thickness (0.6-6.6 mm), and right axillary LN cortical thickness (0.3-3.5 mm). The time elapsed since the last vaccination and the cortex thickness of the left axillary LN were found to be not significantly associated (p=0.437) (Table 5).

In daily clinical practice, we have observed cases of vaccine-induced lymphadenopathy in the contralateral axilla of patients with a history of breast cancer surgery (Figure 1A-D,

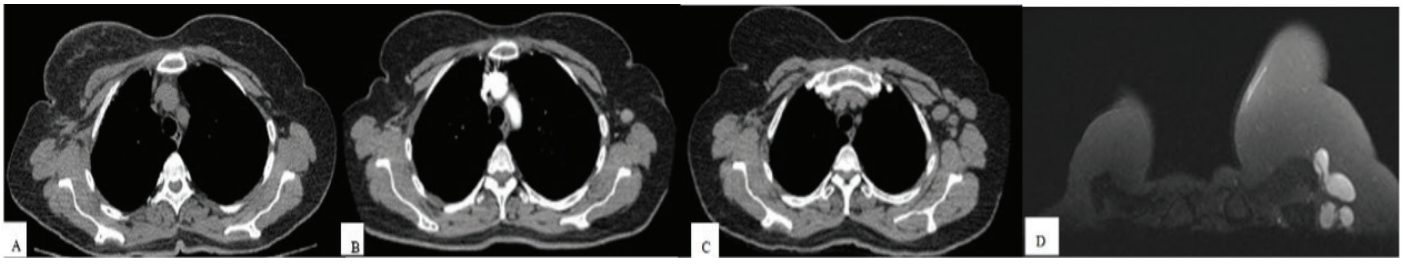


Figure 1. A) In a patient with a history of right partial mastectomy, no pathological lymph node was observed in the left axilla on the pre-COVID-19 vaccine thoracic CT scan (2020). **B)** In the post-vaccine thoracic CT scan (2022), a lymph node with cortical thickening was detected. FNAB was performed, and the result was benign. **C)** In 2023, an increase in the size of the lymph nodes was observed. **D)** In 2023, MRI was performed on the same patient. Due to the progression of lymphadenopathy, a tru-cut biopsy was performed, and the result was reported as benign

COVID-19: Coronavirus disease-2019, CT: Computed tomography, FNAB: Fine-needle aspiration biopsy, MRI: Magnetic resonance imaging

Figure 2A-D). These patients underwent long-term follow-up, multiple biopsies, and 18F-FDG PET/CT scans. Although these LNs exhibited pathological characteristics, repeated biopsies confirmed benign findings. This suggests that vaccines may cause prolonged lymphadenopathy in some patients. Our study supports this observation, as the cortical thickness of the left axillary LNs was significantly increased compared to the right, regardless of the time elapsed since the last vaccine dose. Overall, we found a significant reduction in the cortical thickness of the left axillary LN after

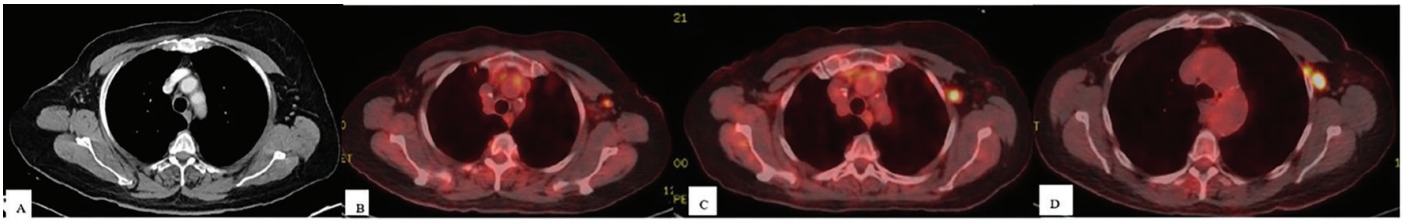


Figure 2. A) In a patient with a history of right mastectomy, reactive lymph nodes were observed in the left axilla in 2020. **B)** In 2022, after vaccination, a lymph node with an SUV_{max} of 3.6 was detected on 18F-FDG PET/CT. FNAB was performed, and the result was benign. **C)** In February 2023, FNAB was repeated for a lymph node with an SUV_{max} of 8.6. The result was benign. **D)** In September 2023, an increase in both size and number of lymph nodes was observed, with an SUV_{max} of 8.4. A tru-cut biopsy was performed and interpreted as reactive growth

SUV: Standardized uptake value, 18F-FDG: 18-fluorodeoxyglucose, PET: Positron emission tomography, CT: Computed tomography, FNAB: Fine-needle aspiration biopsy

Table 1. Patient data, including age, weight, BMI, vaccine type, and number of doses

	Age	Number of doses	Biontech	Sinovac	Weight	Height	BMI
Mean	50.21	3.21	2.27	0.87	70.33	7.6331	27.2863
Median	51.00	3.00	2.00	0.00	68.00	1.6000	26.4463
Standard deviation	13.731	1.123	1.185	1.315	14.564	79.78060	6.26935
Minimum	18	1	0	0	44	1.46	14.53
Maximum	80	6	5	5	130	1057.00	50.15

BMI: Body mass index

Table 2. Sonographic evaluation of lymph nodes showed no statistically significant difference between the mean long-axis diameters of the right and left axillary lymph nodes (p=0.109). However, the mean cortical thickness of the left axillary lymph nodes was significantly higher than that of the right axillary lymph nodes (p=0.049)

Measurement	Mean	Median	Standard deviation	Minimum	Maximum	p-value
Right axillary lymph node long-axis (mm)	18.392	16.7	6.9825	1.2	41.1	0.109
Left axillary lymph node long-axis (mm)	17.751	16.7	5.8354	1.1	38.5	
Right lymph node cortex thickness (mm)	1.6438	1.5	1.19506	0.3	13.0	0.049
Left lymph node cortex thickness (mm)	1.7562	1.6	1.19791	0.38	12.3	

Table 3. Comparisons of lymph node size and cortical thickening among different vaccine types revealed no statistically significant differences between the right and left axillary lymph nodes

Variable	Right size (r)	Right size (p)	Right cortex (r)	Right cortex (p)	Left size (r)	Left size (p)	Left cortex (r)	Left cortex (p)
Age	0.101	0.188	0.089	0.248	0.027	0.724	0.027	0.721
Weight	0.367	<0.001	0.078	0.316	0.285	<0.001	0.092	0.235
Height	-0.009	0.913	-0.032	0.683	0.001	0.986	-0.013	0.864
BMI	0.343	<0.001	0.086	0.271	0.273	<0.001	0.087	0.264

BMI: Body mass index

Table 4. The mean cortical thickness of the left lymph node in patients who were more than two weeks post-vaccination was significantly lower than in those who were two weeks or less post-vaccination (p=0.041)

Measurement	r	p-value
Right size	0.013	0.871
Right lymph node cortex thickness	0.047	0.543
Left size	0.003	0.965
Left lymph node cortex thickness	0.06	0.437

Table 5. The increase in cortex thickness of the left axillary lymph node was found to be significantly higher as more time elapsed since the last vaccination (p<0.437)

Measurement	Mean	Median	Minimum	Maximum	Standard deviation	p-value
Right size (≤2 weeks)	18.343	19.9	7.4	29.2	7.1246	0.838
Right size (>2 weeks)	18.425	16.7	1.2	41.1	7.0095	
Right lymph node cortex thickness ≤2 weeks)	2.0057	2.1	0.64	4.0	1.10675	0.263
Right lymph node cortex thickness (>2 weeks)	1.6285	1.5	0.3	13.0	1.20309	
Left size (≤2 weeks)	17.15	16.45	12.2	22.7	4.3876	0.880
Left size (>2 weeks)	17.764	16.6	1.1	38.5	5.9068	
Left lymph node cortex thickness (≤2 weeks)	2.2833	2.5	1.1	2.9	0.70828	0.041
Left lymph node cortex thickness (>2 weeks)	1.7374	1.5	0.38	12.3	1.2126	

two weeks. This indicates that findings are more pronounced during the first two weeks post-vaccination. Long-term follow-up suggests that COVID-19 vaccines can cause reactive lymphadenopathy in the vaccinated axilla. More extensive studies could help prevent unnecessary biopsies in these cases.

Discussion

With the start of the COVID-19 vaccination process in our country and around the world, many people presented to breast clinics with painful and enlarged axillary LNs after vaccination[9-11]. Managing a patient with only lymphadenopathy may be relatively easier. However, in patients with newly diagnosed breast cancer or a history of breast cancer surgery this condition after vaccination may create complexities regarding axillary metastasis. Axillary LN involvement in breast cancer is an important factor that negatively affects the prognosis of the disease and determines the clinical and surgical approach[12,13]. In breast cancer, cortical thickness greater than 3 mm is considered suspicious[12].

This study evaluated the association of COVID-19 vaccine types, number of doses, and the time elapsed since vaccination with ipsilateral axillary lymphadenopathy, and compared these findings with contralateral axillary LNs. There are studies in the literature evaluating the incidence of hypermetabolic axillary LNs detected by 18F-FDG PET/CT and factors affecting these LNs in oncology patients[14,15]. Especially after the second COVID-19 vaccine injection, an increase in metabolism was detected in ipsilateral axillary LNs[16]. In another study, no relationship was found between the number of doses and axillary lymphadenopathy, similar to the findings of our study[9].

In our study, no significant difference related to vaccination was detected in axillary LNs. There is no direct relationship between the size of LNs and malignancy; however, increased cortical thickness can be seen in cases of malignancy, systemic disease-related LN involvement, and vaccination[12].

The increase in size, thickness, and metabolic activity in axillary LNs, known to develop after vaccination, may pose challenges in breast cancer diagnosis and follow-up, leading to unnecessary invasive procedures. Therefore, the National Comprehensive Cancer Network and the Society of Breast Imaging recommend that breast imaging should be performed 4 to 6 weeks after the second COVID-19 vaccination[14]. In our study, the cortical thickness of the axillary LNs decreased over time.

Our study revealed a significant correlation between BMI and axillary LN size, indicating that patients with higher BMI tend to present with larger LNs. These findings imply that BMI may exert an influence on LN morphology. Consistent with our results, a previous investigation assessing the relationship between BMI and LN dimensions also reported a significant association between obesity and LN size[15].

In our study, LNs on the non-vaccinated side (contralateral) were also included, which is a distinction from other studies. No lymphadenopathy was detected in the contralateral axilla, thus reducing the likelihood of systemic disease-related LN involvement. In our study, it was demonstrated that there was a statistically significant increase in cortical thickness in the left axillary LN compared to the right axilla, in long-term follow-up. Increased cortical thickness in LNs can also be seen in metastasis or primary LN malignancies. Therefore, vaccination-related lymphadenopathy should be considered in the differential diagnosis of these patients.

COVID-19 vaccines such as Sinovac (inactivated) and BioNTech (mRNA) have been developed to prevent COVID-19. When comparing the side effects of both vaccine groups, it was found that people vaccinated with the mRNA group experienced more frequent and severe side effects. Swelling in the LNs was also more frequently observed after mRNA vaccinations[3]. In a study by Cohen et al.[16] ipsilateral axillary swelling/tenderness was seen in 11.6% and 16.0% of participants after the first and second doses, respectively, and was identified as the second most common local reaction following the Moderna (mRNA) COVID-19 vaccine. In the Moderna cohort, clinically detected axillary and supraclavicular lymphadenopathy was reported as an undesirable side effect in 1.1% of study participants within 2-4 days after vaccination[16]. In the Pfizer-BioNTech (mRNA) COVID-19 vaccine trial[8], the rate of ipsilateral axillary and supraclavicular lymphadenopathy was reported as 0.3%. In a study by B et al.[17] lymphadenopathy was most commonly observed in the first 2 weeks after mRNA vaccination; regardless of the vaccine type, it gradually decreased. These findings are similar to the results of our study. In a study involving 49 women with a 12-week follow-up ultrasonography (USG) after the first vaccination, persistent lymphadenopathy was observed in 25 of them (51%). It was found that in more than half of the patients with lymphadenopathy after the second vaccination, the cortical thickness returned to normal within a month[17]. Another study found that the regression time of reactive lymphadenopathy varied, and persistent axillary lymphadenopathy was detected up to 43 weeks after vaccination[18]. In our study, no significant difference was found between the effects of different vaccine types. There are also publications suggesting that human papilloma virus and influenza vaccines may cause similar effects[11]. Due to the different timing of these vaccinations, the incidence of such effects may be too low to be detected or reported. It should be kept in mind that similar results may be observed after future global vaccination campaigns. There are only a few studies in the literature that include long-term follow-up[18].

Although a study describing persistent lymphadenopathy exists, long-term data on COVID-19 vaccines are still lacking[17,18]. However, unexplained reactive lymphadenopathies continue to be observed in daily practice. Overall, our findings are particularly important for cancer patients. After COVID-19 vaccination, the imaging features of lymphadenopathy and its changes over time remain uncertain.

Study Limitations

There were several limitations of this study. Our study was conducted in a single institution. Different results may be obtained in different institutions. Secondly, the limited number of patients may limit the representativeness of the sample for the entire population. More valid results can be obtained with larger cohorts. Additionally, the USG examinations were evaluated by a single radiologist. Interobserver consistency could be calculated with evaluations by different radiology specialists.

Conclusion

In conclusion, COVID-19 vaccination is associated with ipsilateral axillary reactive lymphadenopathy. Consistent with the findings of our study, vaccination was associated with a significant long-term increase in cortical thickness of ipsilateral LNs compared to contralateral LNs. In breast cancer diagnosis and follow-up, querying the vaccination history in such cases may prevent unnecessary invasive procedures and misdiagnoses.

Ethics

Ethics Committee Approval: Ethical approval for this retrospective study was obtained from the Non-Interventional Clinical Research Ethics Committee of İzmir Katip Çelebi University (approval no: 0043, date: 15.02.2024).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practises: S.Ü., Concept: S.Ü., Design: S.Ü., Data Collection or Processing: S.Ü., Ö.P., Analysis or Interpretation: Ö.P., Literature Search: S.Ü., M.G., Writing: S.Ü., M.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. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382:727-33.
2. Polack FP, Thomas SJ, Kitchin N, et al. Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. N Engl J Med. 2020;383:2603-15.
3. Baden LR, El Sahly HM, Essink B, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med. 2021;384:403-16.
4. Bar-On YM, Goldberg Y, Mandel M, et al. Protection of BNT162b2 vaccine booster against COVID-19 in Israel. N Engl J Med. 2021;385: 1393-400.
5. Spitzer A, Angel Y, Marudi O, et al. Association of a third dose of BNT162b2 vaccine with incidence of SARS-CoV-2 infection among health care workers in Israel. JAMA. 2022;327:341-9.
6. Menni C, Klaser K, May A, et al. Vaccine side-effects and SARS-CoV-2 infection after vaccination in users of the covid symptom study app in the UK: a prospective observational study. Lancet Infect Dis. 2021;21:939-49.
7. Hiller N, Goldberg SN, Cohen-Cymberknoh M, Vainstein V, Simanovsky N. Lymphadenopathy associated with the COVID-19 vaccine. Cureus. 2021;13:e13524.
8. Schiaffino S, Pinker K, Magni V, et al. Axillary lymphadenopathy at the time of COVID-19 vaccination: ten recommendations from the European Society of Breast Imaging (EUSOBI). Insights Imaging. 2021;12:119.
9. Robinson KA, Maimone S, Gococo-Benore DA, Li Z, Advani PP, Chumsri S. Incidence of axillary adenopathy in breast imaging after COVID-19 vaccination. JAMA Oncol. 2021;7:1395-7.
10. Nguyen DL, Ambinder EB, Myers KS, Mullen LA, Panigrahi B, Oluyemi E. COVID-19 vaccine-related axillary adenopathy on breast imaging: follow-up recommendations and histopathologic findings. AJR Am J Roentgenol. 2022;218:997-8.
11. Özütemiz C, Krystosek LA, Church AL, et al. Lymphadenopathy in COVID-19 vaccine recipients: diagnostic dilemma in oncologic patients. Radiology. 2021;300:e296-300.
12. Dixon JM, Cartlidge CWJ. Twenty-five years of change in the management of the axilla in breast cancer. Breast J. 2020;26:22-6.
13. Dumitru D, Khan A, Catanuto G, Rocco N, Nava MB, Benson JR. Axillary surgery in breast cancer: the beginning of the end. Minerva Chir. 2018;73:314-21.
14. Bernstine H, Priss M, Anati T, et al. Axillary lymph nodes hypermetabolism after BNT162b2 mRNA COVID-19 vaccination in cancer patients undergoing 18F-FDG PET/CT: a cohort study. Clin Nucl Med. 2021;46:396-401.
15. diFlorio Alexander RM, Haider SJ, MacKenzie T, Goodrich ME, Weiss J, Onega T. Correlation between obesity and fat-infiltrated axillary lymph nodes visualized on mammography. Br J Radiol. 2018;91:20170110.
16. Cohen D, Krauthammer SH, Wolf I, Even-Sapir E. Hypermetabolic lymphadenopathy following administration of BNT162b2 mRNA COVID-19 vaccine: incidence assessed by [18F]FDG PET-CT and relevance to study interpretation. Eur J Nucl Med Mol Imaging. 2021;48:1854-63.
17. B SS, Patil KS, Kumar A. Unilateral axillary lymphadenopathy following COVID-19 vaccination. International Journal of Health Sciences. 2022;6:11158-65.
18. Wolfson S, Kim E, Plaunova A, et al. Axillary adenopathy after COVID-19 vaccine: no reason to delay screening mammogram. Radiology. 2022;304:e57.

DOI: 10.4274/anatoljmed.2025.65487
Anatol J Gen Med Res 2025;35(3):260-268

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Conventional Versus Holminum: YAG Laser-assisted Resection of Concha Bullosa: A Comparative Study

Konvansiyonel ve Holmium: YAG Lazer ile Yapılan Konka Bulloza Rezeksiyonlarının Karşılaştırılması

Elvin Alaskarov[1], Necdet Özçelik[1], Aynur Aliyeva[2]

[1]Medipol University Esenler Hospital, Clinic of Otorhinolaryngology, İstanbul, Türkiye
[2]The Catholic University of Korea, Seoul St. Mary's Hospital, Department of Otorhinolaryngology, Division of Head and Neck Surgery, Seoul, South Korea

Cite as: Alaskarov E, Özçelik N, Aliyeva A. Conventional versus holminum: YAG laser-assisted resection of concha bullosa: a comparative study. Anatol J Gen Med Res. 2025;35(3):260-268

Abstract

Objective: This study aims to compare the surgical outcomes of conventional concha bullosa resection and holmium: YAG laser-assisted lateral laminectomy, focusing on efficacy, safety, and postoperative recovery in patients with symptomatic concha bullosa.

Methods: A total of 46 patients (54 concha bullosa surgeries) were included. Patients were divided into two groups: group 1 underwent holmium: YAG laser-assisted lateral laminectomy (n=23), and group 2 underwent conventional lateral laminectomy (n=23). Primary outcomes included surgical duration, intraoperative bleeding, and postoperative complications. Symptom evaluation was conducted using the visual analog scale (VAS) for headaches and the sinonasal outcome test-22 (SNOT-22), recorded preoperatively and three months postoperatively.

Results: The mean surgical duration was significantly shorter in group 1 (15.0±5.7 minutes) compared to group 2 (28.0±7.5 minutes, $p<0.05$). Group 1 demonstrated a clearer surgical field with visibly reduced bleeding. Both groups exhibited statistically significant improvements in VAS and SNOT-22 scores ($p<0.05$); however, intergroup comparisons showed no significant differences. Crusting persisted longer in group 2, and one case of postoperative epistaxis requiring nasal packing was reported in this group. No major complications occurred in either group.

Conclusion: Holmium: YAG laser-assisted concha bullosa resection is a safe and effective alternative to conventional lateral laminectomy. It offers distinct intraoperative advantages, including shorter operative time, better visualization, and reduced bleeding. Despite similar symptomatic improvement in both groups, laser-assisted resection may enhance surgical efficiency and patient recovery. Further prospective studies with long-term follow-up are warranted to evaluate recurrence and delayed complications.

Keywords: Concha bullosa, endoscopic sinus surgery, holmium, laser therapy, visual analog scale

Öz

Giriş: Bu çalışmanın amacı, semptomatik konka bulloza hastalarında konvansiyonel konka bulloza rezeksiyonu ile holmium: YAG lazer destekli lateral laminektominin cerrahi sonuçlarını; etkinlik, güvenlik ve postoperatif iyileşme açısından karşılaştırmaktır.

Yöntem: Çalışmaya toplam 46 hasta (54 konka bulloza cerrahisi) dahil edilmiştir. Hastalar iki gruba ayrılmıştır: grup 1'de holmium: YAG lazer destekli lateral laminektomi (n=23), grup 2'de ise konvansiyonel lateral laminektomi (n=23) uygulanmıştır. Birincil sonuç ölçütleri arasında cerrahi süresi, intraoperatif

Address for Correspondence/Yazışma Adresi: Asst. Prof., Elvin Alaskarov, Medipol University Esenler Hospital, Clinic of Otorhinolaryngology, İstanbul, Türkiye
E-mail: ealaskarov@medipol.edu.tr
ORCID ID: orcid.org/0000-0002-3572-0062

Received/Geliş tarihi: 03.07.2025
Accepted/Kabul tarihi: 16.09.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

kanama ve postoperatif komplikasyonlar yer almıştır. Semptom değerlendirmesi baş ağrısı için görsel analog skala (GAS) ve sinonazal semptomlar için sinonazal sonuç testi-22 (SNOT-22) kullanılarak, ameliyat öncesi ve ameliyat sonrası üçüncü ayda yapılmıştır.

Bulgular: Ortalama cerrahi süresi grup 1'de anlamlı derecede daha kısa bulunmuştur (15,0±5,7 dakika) (grup 2: 28,0±7,5 dakika; p<0,05). Grup 1'de daha net bir cerrahi alan ve belirgin şekilde azalmış kanama gözlenmiştir. Her iki grupta da GAS ve SNOT-22 skorlarında istatistiksel olarak anlamlı düzelmeler elde edilmiştir (p<0,05); ancak gruplar arası karşılaştırmada anlamlı fark bulunmamıştır. Kabuklanma grup 2'de daha uzun sürmüş ve bu grupta bir hastada nazal tampon gerektiren postoperatif epistaksis gözlenmiştir. Her iki grupta da majör komplikasyon izlenmemiştir.

Sonuç: Holmium: YAG lazer destekli konka bulloza rezeksiyonu, konvansiyonel lateral laminektomiye güvenli ve etkili bir alternatiftir. Daha kısa operasyon süresi, daha iyi görüş alanı ve azalmış kanama gibi belirgin intraoperatif avantajlar sunmaktadır. Her iki yöntemde semptomatik iyileşme benzer olsa da lazer destekli rezeksiyon cerrahi etkinliği ve hasta iyileşmesini artırabilir. Nüks ve geç komplikasyonların değerlendirilmesi için uzun dönem takipli prospektif çalışmalara ihtiyaç vardır.

Anahtar Kelimeler: Konka bulloza, endoskopik sinüs cerrahisi, holmium, lazer tedavisi, görsel analog skala

Introduction

Concha bullosa, commonly defined as the pneumatization of the middle nasal turbinate, is reported in the literature as the most prevalent anatomical variation in the sinonasal region. The prevalence of concha bullosa ranges from 14% to 53%. Although the etiology of turbinate pneumatization remains uncertain, factors such as nasal septum deviation, mouth breathing, and trauma are thought to play a role[1-3].

Although most cases of concha bullosa are asymptomatic, the condition can obstruct the middle meatus by enlarging the middle turbinate, which may lead to blockage of the osteomeatal complex. This obstruction can impair the mucociliary drainage of the maxillary sinus, resulting in sinusitis. Symptomatic cases have also been linked to headaches, chronic sinusitis, and deviation of the nasal septum[4,5]. In such cases, surgical intervention becomes necessary to restore nasal airflow and improve sinus ventilation.

Several surgical techniques have been described for the treatment of concha bullosa. The conventional approach, which involves resecting the portion obstructing the nasal airway and sinus ventilation, is commonly preferred[6]. These techniques include lateral or medial laminectomy, crushing, turbinoplasty, and intrinsic stripping, each with its own risk-benefit profile depending on the extent and type of conchal pneumatization.

In recent years, laser-assisted procedures have gained popularity in otolaryngologic surgery due to their precision, hemostatic effect, and reduced tissue trauma. Among various laser modalities, the holmium: YAG laser offers specific advantages such as minimal penetration depth and strong water absorption, enabling controlled tissue ablation with limited collateral damage. However, its use in middle turbinate surgery, particularly for concha bullosa resection, remains underexplored in the literature.

In this study, we compared the outcomes of two surgical techniques: the conventional method and the holmium: YAG laser-assisted approach. Our aim was to evaluate their effectiveness, safety, and influence on intraoperative parameters and postoperative symptom improvement.

Materials and Methods

Study Design

This retrospective case-control study was conducted at the Clinic of Otorhinolaryngology of Medipol University Esenler Hospital between 2018 and 2023. The study protocol was approved by the İstanbul Medipol University Institutional Ethics Committee (approval no: 213, date: 20.02.2025) in accordance with the Declaration of Helsinki.

Patient Selection

Data from 54 concha bullosa surgeries performed on 46 patients were analyzed. The demographic data and primary complaints of the patients were evaluated. The most frequently reported symptoms were nasal obstruction and intermittent headaches. Classification of concha bullosa was based on nasal endoscopic examination. Preoperative computed tomography (CT) of the paranasal sinuses was performed using a multidetector CT unit.

Inclusion and Exclusion Criteria

Patients were included if they had symptomatic concha bullosa and underwent either laser-assisted or conventional resection. Exclusion criteria were as follows:

- Asymptomatic concha bullosa
- Age below 18 years
- Previous nasal surgery
- Concomitant nasal pathologies such as chronic rhinosinusitis with or without nasal polyps
- Incomplete follow-up data

Surgical Groups and Procedure

Patients were divided into two groups based on the surgical method used:

- Group 1: Holmium: YAG laser-assisted lateral laminectomy (n=23)
- Group 2: Conventional lateral laminectomy (n=23)

The surgical method was selected based on surgeon discretion and patient preference; randomization was not performed. All surgeries were performed by the same senior surgeon with assistance from the same team. Blinding was not applied during outcome assessments, as postoperative evaluations were performed by the operating surgeon.

Bilateral and unilateral concha bullosa cases were present in both groups. For resection time analysis, each concha was evaluated independently. In patients with bilateral concha bullosa, the resection times were summed and averaged per concha to maintain consistency.

Surgical and Clinical Assessment

All surgeries were performed under general anesthesia using endoscopic guidance with a 4 mm, 0° rigid telescope (Karl Storz, Tuttlingen, Germany). Intraoperative parameters included operative time, surgical field visibility, and bleeding. Because a significant number of patients underwent simultaneous septoplasty and/or inferior turbinate reduction, efforts were made to isolate the outcomes of concha bullosa resection. Therefore, intraoperative parameters such as operative time and bleeding were evaluated specifically during the concha bullosa phase of surgery. Septoplasty and inferior turbinate reduction durations were not included in the recorded concha bullosa resection time. The degree of bleeding was assessed subjectively by the surgeon using a binary descriptive scale: "less" or "more", based on visual clarity during the procedure. Simultaneous nasal procedures, including septoplasty and inferior turbinate reduction, were noted. Postoperative recovery and complications were recorded.

All patients underwent preoperative and postoperative evaluations using:

- Visual analog scale (VAS) for headache
- Sinonasal outcome test-22 (SNOT-22) for sinonasal symptom severity

Postoperative assessments were conducted at three months. Concha bullosa volume was assessed through nasal endoscopy and was scored using a 3-point scale (0-3) for size and prominence. This scale was adopted from previous studies and validated through consistent application by the same surgeon across all cases.

Statistical Analysis

Data analysis was performed using SPSS software (version 22.0, Chicago, IL, USA). A parametric paired-sample t-test was used to compare pre- and postoperative scores within groups. Independent-sample t-tests were applied for intergroup comparisons. Categorical variables were analyzed using the chi-square (χ^2) test. A p-value less than 0.05 was considered statistically significant.

Operative Technique

After the administration of general anesthesia, a local anesthetic solution consisting of lidocaine hydrochloride (2%) and adrenaline (0.1%) was injected submucosally into the anterior and posterior regions of the middle turbinate. In cases requiring simultaneous septoplasty or inferior turbinate reduction, these procedures were performed prior to concha bullosa intervention. Both groups were treated using a 0° rigid endoscope (Karl Storz, Tuttlingen, Germany).

In the holmium: YAG laser-assisted group, a 0.7 mm flexible quartz laser fiber (LISA, SPHINX® JR, Germany) was utilized by inserting it through an aspirator tube for maneuverability and visibility. Energy levels ranging from 0.5 to 1.5 J were applied at a frequency of 3 to 5 Hz. The mucosa was initially marked with laser shots to delineate the resection area, followed by precise ablation and resection of the bony conchal segment using direct laser contact. Subsequently, the complete mobilization of the lateral segment was verified

with the assistance of a Cottle elevator. Finally, the resected portion was removed from the nasal cavity using a punch (Figure 1).

In the conventional lateral laminectomy group, the concha was accessed using a hooked knife at the most ventilated area, followed by resection of the obstructing lateral portion with turbinate scissors. Hemostasis was successfully achieved in all cases using bipolar cautery (Valleylab Force 2 ESU, USA).

Patients were discharged the day after surgery and prescribed a one-week course of amoxicillin-clavulanate and analgesics. A Doyle splint tampon was applied at the end of the procedure and removed on postoperative day four. Saline irrigation and moisturizing nasal drops were prescribed three times daily for 2.5 weeks.

Results

A total of 46 patients were included in the study, with 23 patients in group 1 (laser-assisted) and 23 patients in group 2 (conventional). There was no statistically significant difference in gender distribution between the groups (p=0.750). The mean age was 39.56±13.35 years in group 1 and 37.44±11.64 years in group 2 (p=0.378) (Table 1).

Table 1. Patient characteristics and surgical details (n=46)

Characteristics	Group 1 (laser-assisted)	Group 2 (conventional)	p-value*
Gender (M/F)	8:15	10:13	0.750
Age (years, mean ± SD)	39.56±13.35	37.44±11.64	0.378
Side of concha bullosa, n (%)			0.686
Unilateral	18 (78%)	20 (87%)	
Bilateral	5 (22%)	3 (13%)	
Types of concha bullosa, n (%)			
Lamellar	4 (15%)	6 (24%)	
Bulbous	10 (35%)	8 (30%)	
Extensive	14 (50%)	12 (46%)	
Concurrent surgery, n			0.552
Septoplasty+reduction of inferior turbinates	19	21	
Reduction of inferior turbinates	4	2	

*: Statistical significance was set at p<0.05, SD: Standard deviation

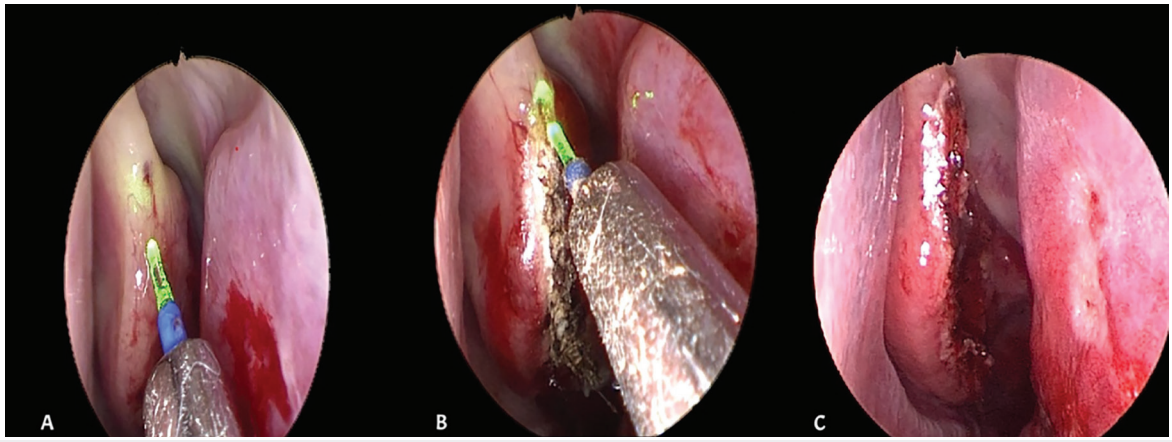


Figure 1. Intraoperative views of holmium: YAG laser-assisted resection of concha bullosa

(A) Initial laser marking of the mucosa to define the resection area. **(B)** Lateral laminectomy of the aerated middle turbinate resection using direct laser contact. **(C)** Post-resection view showing the treated area

Regarding laterality, unilateral concha bullosa was present in 78% of group 1 and 87% of group 2 patients, while bilateral involvement was seen in 22% and 13% of patients, respectively (p=0.686). The types of concha bullosa-lamellar, bulbous, and extensive-were similarly distributed between groups. Concurrent nasal surgeries were common in both groups, primarily septoplasty combined with inferior turbinate reduction, with no significant differences noted (p=0.552).

The mean follow-up duration was 6.5±1.8 months in group 1 and 6.6±1.5 months in group 2 (p>0.05). Mean resection time was significantly shorter in group 1 (15±5.7 minutes) compared to group 2 (28±7.5 minutes) (p<0.001). Additionally, the surgical field was consistently reported as clearer and less bloody in the laser-assisted group (Table 2). In subgroup analysis, no significant difference in resection time was observed between unilateral and bilateral cases when calculated per turbinate. The reported 28-minute average in the conventional group reflects cumulative intraoperative challenges such as intraoperative bleeding and limited visualization, rather than septoplasty or turbinate reduction time.

Endoscopic image classification showed significant postoperative reductions in concha bullosa scores in both groups, at the three-month follow-up (p<0.001 for each); however, the difference between groups was not statistically significant (Table 3).

Both groups demonstrated statistically significant improvement in VAS scores for headache and SNOT-22 scores for sinonasal symptoms (p<0.001 within groups). However, comparisons between groups revealed no significant intergroup differences in either VAS or SNOT-22 outcomes (Table 3). Figure 2 visually illustrates the pre- and postoperative improvements in both symptom domains across groups, confirming clinical benefit independent of the surgical technique.

Postoperative endoscopic follow-ups revealed temporary crusting and mucosal edema in both groups, more commonly persisting into the second week in group 2. One patient in group 2 experienced epistaxis on postoperative day five, which was successfully managed with bipolar cautery under local anesthesia. No further complications were reported. Importantly, there were no occurrences of cerebrospinal fluid (CSF) leakage, orbital injury, anosmia, synechiae, exposed bone, lateralized turbinates, or persistent symptoms in either group.

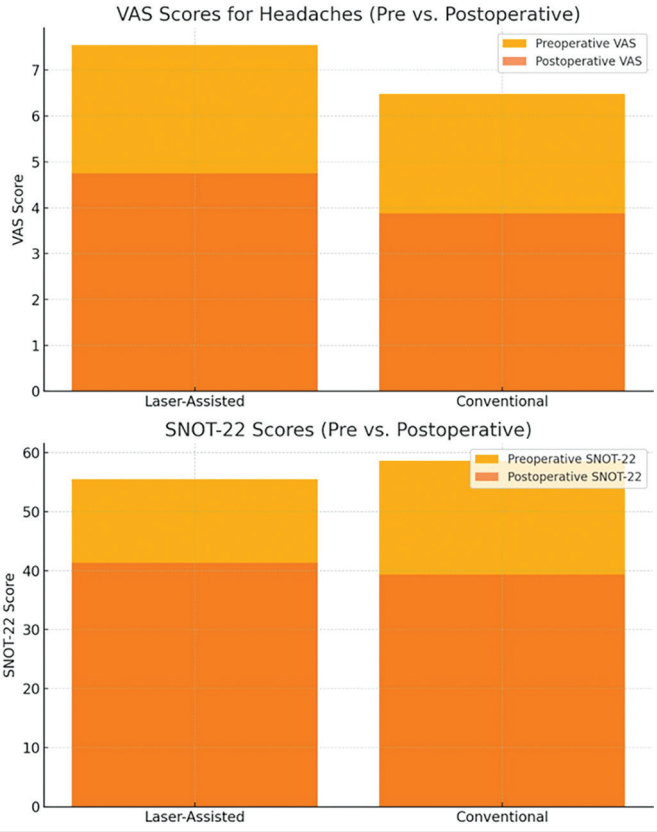


Figure 2. Comparison of preoperative and postoperative VAS and SNOT-22 scores between groups

VAS: Visual analog scale, SNOT-22: Sinonasal outcome test-22

Table 2. Comparison of the intraoperative and postoperative parameters of each group of patients

Parameter	Group 1 (laser-assisted)	Group 2 (conventional)	p-value
Mean resection time ± SD (minutes)	15±5.7	28±7.5	0.000*
Surgical field bleeding	Less	More	
Mean follow-up duration ± SD (months)	6.5±1.8	6.6±1.5	>0.05
Major complications	None	None	

*: p<0.05, independent samples t-test, SD: Standard deviation

Table 3. Comparative analysis of endoscopic classification scores, VAS, and SNOT-22 results in laser-assisted and conventional concha bullosa resection groups

The concha bullosa scores of endoscopic image classification			
	Endoscopy preoperative (mean ± SD)	**Endoscopy postoperative 3rd month (mean ± SD)**	**Paired student t-test (p-value)**
Group 1 (laser-assisted, n=23)	2.60±0.35	1.90±0.58	0.000*
Group 2 (conventional, n=23)	2.50±0.46	1.85±0.55	0.000*
Unpaired student t-test (p-value)	0.140	0.115	
Evaluation of VAS for headaches results in patients undergoing middle turbinate surgery			
	Preoperative VAS (mean ± SD)	**Postoperative VAS 3rd month (mean ± SD)**	**Paired student t-test (p-value)**
Group 1 (laser-assisted, n=23)	7.55±1.95	4.75±2.12	0.000*
Group 2 (conventional, n=23)	6.48±2.16	3.88±2.18	0.000*
Unpaired student t-test (p-value)	0.247	0.355	
Evaluation of SNOT-22 scores results in patients undergoing middle turbinate surgery			
	Preoperative SNOT-22 (mean ± SD)	**Postoperative SNOT-22 3rd month (mean ± SD)**	**Paired student t-test (p-value)**
Group 1 (laser-assisted, n=23)	55.45±12.95	41.35±13.72	0.004*
Group 2 (conventional, n=23)	58.64±11.16	39.29±15.18	0.000*
Unpaired student t-test (p-value)	0.557	0.672	

*: p-value considered significant at <0.05, SD: Standard deviation, VAS: Visual analog scale, SNOT-22: Sinonasal outcome test-22

Regression analysis (Table 4) identified that shorter resection time and higher preoperative SNOT-22 scores significantly predicted greater VAS improvement (β=-0.0297 and +0.0585, respectively; p<0.001). For SNOT-22 score improvement, significant predictors included longer resection time and higher baseline VAS scores (coefficients of β=0.4782 and 0.9033, respectively; p<0.001). These results emphasize the dual role of symptom severity and surgical dynamics in influencing postoperative recovery.

Figure 3 presents a correlation heatmap summarizing the strong associations between symptom severity, surgical time, and postoperative improvement. These patterns reinforce the internal consistency of clinical outcomes across both subjective and procedural variables.

Discussion

Although middle concha bullosa is a common anatomical variation, surgical intervention is not always recommended. Surgery is primarily indicated in symptomatic cases where concha bullosa contributes to the obstruction of the osteomeatal complex. Various surgical techniques have been described in the literature, including lateral or medial partial resection, total resection, turbinoplasty, crushing, and crushing with intrinsic stripping. However, there is no clear consensus on the optimal surgical technique yet[7,8].

In another study, the intrinsic stripping technique was shown to significantly reduce concha bullosa volume while also lowering recurrence rates[9]. No significant postoperative functional differences have been observed between medial and lateral partial resections, commonly referred to as classical resection. However, medial resection has been noted to offer advantages in terms of preventing synechiae in the lateral nasal wall-frontal recess area and facilitating frontal sinus drainage[10].

The crushing technique provides advantages such as a shorter operative time and the ability to be performed simultaneously with other nasal surgeries. However, it carries a risk of long-term recurrence, as noted in previous studies[11,12].

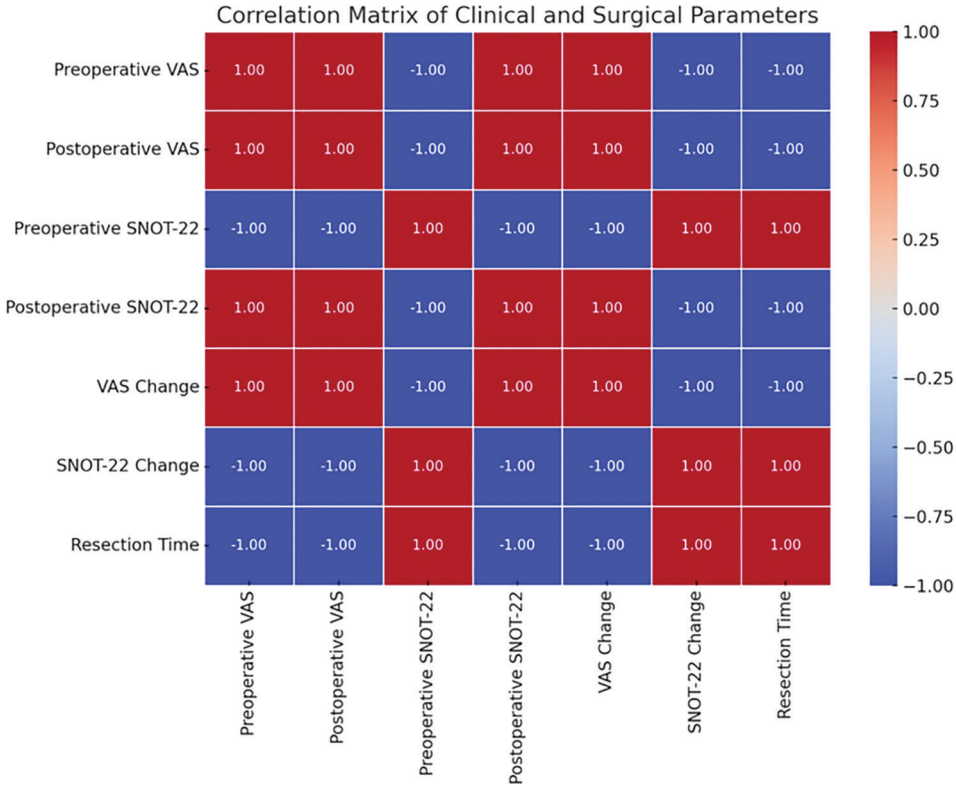


Figure 3. Correlation heatmap of symptom scores and surgical parameters in concha bullosa resection

VAS: Visual analog scale, SNOT-22: Sinonasal outcome test-22

Table 4. Regression analysis of predictors for VAS and SNOT-22 improvement

Outcome variable	Predictor	Beta coefficient (β)	p-value	Interpretation
VAS change	Resection time	−0.0297	<0.001	Shorter resection time is associated with greater VAS improvement.
VAS change	Preoperative SNOT-22	+0.0585	<0.001	Higher symptom burden predicts more headache relief.
SNOT-22 change	Resection time	+0.4782	<0.001	Longer surgery duration is associated with greater SNOT-22 improvement.
SNOT-22 change	Preoperative VAS	+0.9033	<0.001	More severe headache symptoms predict greater improvement in SNOT-22.

*: All models had R^2=1.000 due to dataset uniformity. Beta values indicate direction and strength of association, VAS: Visual analog scale, SNOT-22: Sinonasal outcome test-22

All concha bullosa cases were classified into three types: lamellar, bulbous, and extensive. A prospective study evaluating the crushing technique in all three types reported a significant reduction in concha volume on preoperative and postoperative tomography. Additionally, no cases of re-pneumatization were observed during long-term follow-up[13].

In our case series, the type and volume of concha bullosa were evaluated preoperatively through nasal endoscopic examination and paranasal sinus CT. In group 1, 16 bulbous and 7 extensive concha bullosa were observed, while in group 2, 14 bulbous, 8 extensive, and 1 lamellar type were identified. In the postoperative period, concha volume was assessed only through nasal endoscopic examination to avoid unnecessary radiation exposure.

A fundamental principle of surgery is minimizing tissue damage while effectively resolving the pathology in a short

time. This principle is particularly relevant in concha bullosa surgery, where limited and controlled resection of the affected portion is the primary goal[14,15].

The conventional resection technique, lateral excision of the middle turbinate, is the most commonly used approach for isolated concha bullosa. Since the medial part of the middle turbinate is connected to the skull base, lateral resection is considered safer in terms of preventing CSF fistula formation[16]. In our study, lateral resection was performed in both patient groups.

Ismail et al.[11] prospectively evaluated three surgical techniques for the treatment of symptomatic concha bullosa: crushing, lateral laminectomy with mucosal preservation, and lateral laminectomy without mucosal preservation. Postoperative headache assessment using the VAS showed a statistically significant improvement in patients who underwent lateral laminectomy (with and without mucosal preservation) compared to the crushing group. However, no significant difference was observed between the two lateral laminectomy groups[11]. In our study, in both groups undergoing lateral laminectomy, the resected bony segment was excised along with its inner and outer mucosal surfaces.

However, postoperative complications such as bleeding, synechiae, sinusitis, and olfactory dysfunction have been reported following middle turbinate surgery[17,18]. In our study, one patient from each group required hospitalization and re-tamponade due to nasal bleeding within the first postoperative week.

In some cases, the middle turbinate has a direct anatomical relationship with the skull base. Excessive mobilization (destabilization of the middle turbinate) or forceful maneuvers during surgery in such cases may lead to CSF leakage. This potential complication has been previously reported following septal and turbinate surgeries[19,20].

The use of a holmium: YAG laser in group 1 aimed to achieve sufficient resection without excessive manipulation of the middle turbinate. Compared to the conventional technique, laser-assisted resection provided a clearer surgical field with reduced intraoperative bleeding and a shorter resection time, which are significant advantages.

Winters and Worley[21] previously reported the use of the LightForce Gold Laser for concha bullosa resection, highlighting its ease of use and hemostatic. However, to our knowledge, no prior studies have investigated holmium: YAG laser-assisted concha bullosa resection. In our case series, holmium: YAG laser was utilized not only for concha bullosa resection but also for inferior turbinate reduction. The lower incidence of postoperative crusting in group 1 compared to group 2 may be attributed to the reduced need for bipolar cauterization for hemostasis in the laser-assisted technique.

Study Limitations

This study has several limitations. First, its retrospective design may introduce selection and information biases, as randomization and blinding were not applied. Second, although efforts were made to isolate the outcomes of concha bullosa resection, many patients underwent concurrent nasal procedures such as septoplasty and inferior turbinate reduction, which may have influenced postoperative symptom scores. Third, the relatively small sample size and short follow-up period (mean 6.5 months) limit the generalizability of the findings and prevent a comprehensive evaluation of long-term outcomes, such as recurrence or empty nose syndrome. Additionally, the subjective assessment of intraoperative bleeding and the absence of objective bleeding quantification may reduce the reproducibility of this outcome. Finally, the study did not include an economic analysis of the laser device's cost-effectiveness, which is an important consideration for wider clinical adoption.

Future prospective studies with larger cohorts and extended follow-up periods (>2 years) are warranted to better evaluate the long-term outcomes of these techniques. Furthermore, comparative studies involving other laser modalities, such as CO_2 or diode lasers, could provide further insights into optimizing surgical approaches for concha bullosa resection.

Conclusion

Holmium: YAG laser-assisted lateral laminectomy is a safe and effective alternative to conventional surgical techniques for the treatment of symptomatic concha bullosa. While both methods significantly improve headache and sinonasal symptoms, the laser-assisted approach offers advantages in terms of reduced operative time, improved surgical field visibility, and minimal intraoperative bleeding. These findings suggest that incorporating laser technology may enhance surgical efficiency without compromising clinical outcomes. Further studies with long-term follow-up and broader patient populations are recommended to validate these results.

Ethics

Ethics Committee Approval: The study protocol was approved by the İstanbul Medipol University Institutional Ethics Committee (approval no: 213, date: 20.02.2025) in accordance with the Declaration of Helsinki.

Informed Consent: This retrospective case-control study was conducted at the Clinic of Otorhinolaryngology of Medipol University Esenler Hospital between 2018 and 2023.

Footnotes

Authorship Contributions

Surgical and Medical Practises: E.A., Concept: N.Ö., A.A., Design: E.A., N.Ö., A.A., Data Collection or Processing: E.A., Analysis or Interpretation: N.Ö., Literature Search: E.A., A.A., Writing: E.A., N.Ö., A.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. Orlandi RR, Kingdom TT, Hwang PH, et al. International consensus statement on allergy and rhinology: rhinosinusitis. Int Forum Allergy Rhinol. 2016;6(Suppl 1):22-209.
2. Mehta KS, Yousuf A, Wazir IA, Sideeq K. Clinical benefits of surgical management of concha bullosa. Int J Otorhinolaryngol Head Neck Surg. 2017;3:833-6.
3. Kar M, Altıntaş M. The incidence of concha bullosa: a retrospective radiologic study. Eur Arch Otorhinolaryngol. 2023;280:731-5.
4. Tiwari R, Goyal R. Role of concha bullosa in chronic rhinosinusitis. Indian J Otolaryngol Head Neck Surg. 2019;71:128-31.
5. Kucybała I, Janik KA, Ciuk S, Storman D, Urbanik A. Nasal septal deviation and concha bullosa - do they have an impact on maxillary sinus volumes and prevalence of maxillary sinusitis? Pol J Radiol. 2017;82:62-8.
6. İsmi O, Özcan C. Concha bullosa and its surgery: review. Turk Rhinol J. 2014;3:117-22.
7. Woolford TJ, Jones NS. A concha bullosa crusher for use in endoscopic sinus surgery. J Laryngol Otol. 2000;114:224-5.
8. Mesbahi A, Movahhedian N, Akbarizadeh F, Hakimi AA, Khojastepour L. Assessing the efficacy of a modified crushing technique for the management of concha bullosa: a cone beam computer tomography study. Braz J Otorhinolaryngol. 2022;88:902-906.
9. Eren SB, Kocak I, Dogan R, Ozturan O, Yildirim YS, Tugrul S. A comparison of the long-term results of crushing and crushing with intrinsic stripping techniques in concha bullosa surgery. Int Forum Allergy Rhinol. 2014;4:777-82.
10. Kumral TL, Yıldırım G, Çakır O, et al. Comparison of two partial middle turbinectomy techniques for the treatment of a concha bullosa. Laryngoscope. 2015;125:1062-6.
11. Ismail M, Abdelhak B, Hamead K, Ebraheem RA, Abdelmoneim RA. Surgical treatment of concha bullosa: a comparison of the short-term results of crushing and lateral laminectomy with and without mucosal preservation. Indian J Otolaryngol Head Neck Surg. 2024;76:1949-58.
12. Tanyeri H, Aksoy EA, Serin GM, Polat S, Türk A, Önal ÜF. Will a crushed concha bullosa form again? Laryngoscope. 2012;122:1108-12.
13. Koçak İ, Gökler O, Doğan R. Is it effective to use the crushing technique in all types of concha bullosa? Eur Arch Otorhinolaryngol. 2016;273:3745-50.
14. Badran HS. Role of surgery in isolated concha bullosa. Clin Med Insights Ear Nose Throat. 2011;4:13-9.
15. Andaloro C, La Mantia I, Castro V, Grillo C. Comparison of nasal and olfactory functions between two surgical approaches for the treatment of concha bullosa: a randomised clinical trial. J Laryngol Otol. 2019;133:913-7.
16. Braun H, Stammberger H. Pneumatization of turbinates. Laryngoscope. 2003;113:668-72.
17. Ramadan HH, Allen GC. Complications of endoscopic sinus surgery in a residency training program. Laryngoscope. 1995;105:376-9.
18. Shih C, Chin G, Rice DH. Middle turbinate resection: impact on outcomes in endoscopic sinus surgery. Ear Nose Throat J. 2003;82:796-7.
19. Leong AC, Patel T, Rehman F, Oyarzabal M, Gluckman P. Cerebrospinal fluid rhinorrhea complicating septoplasty: a novel mechanism of injury. Ear Nose Throat J. 2010;89:27-9.
20. Soni RS, Choudhry OJ, Liu JK, Eloy JA. Postoperative cerebrospinal fluid leak after septoplasty: a potential complication of occult anterior skull base encephalocele. Allergy Rhinol (Providence). 2013;4:e28-31.
21. Winters R, Worley NK. Gold laser resection of the concha bullosa: description of a new technique. Otolaryngology. 2012;2:114.

DOI: 10.4274/anatoljmed.2025.09327
Anatol J Gen Med Res 2025;35(3):269-275

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Long-term Outcomes in Stage I Seminoma Patients: A Single-center Retrospective Study

Evre I Seminom Hastalarının Uzun Dönem Takip Sonuçları: Tek Merkez Deneyimi

Sercan Ön[1], Duygu Ayaz[2], Zeynep Altın[3], Mehmet Zeynel Keskin[4]

[1]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of Oncology, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of Pathology, İzmir, Türkiye
[3]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of Internal Medicine, İzmir, Türkiye
[4]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of Urology, İzmir, Türkiye

Cite as: Ön S, Ayaz D, Altın Z, Keskin MZ. Long-term outcomes in stage I seminoma patients: a single-center retrospective study. Anatol J Gen Med Res. 2025;35(3):269-275

Abstract

Objective: This study aims to evaluate long-term outcomes and relapse rates in patients with stage I seminoma treated at our center and to assess the effectiveness and safety of active surveillance and adjuvant therapy in clinical practice.

Methods: This retrospective study included 63 patients diagnosed with stage I seminoma at our institution between 2007 and 2024. All patients underwent radical inguinal orchiectomy. Patient demographics, tumors' pathological features, relapse rates, treatment-related adverse events, and long-term oncological outcomes were analyzed.

Results: The median age was 36 years, and the median follow-up was 50.5 months. Tumor size was ≥4 cm in 52.4% of cases, and rete testis invasion was present in 33.3%. Postoperative management included surveillance (n=43), adjuvant carboplatin (n=18), and radiotherapy (n=2). Adjuvant therapy was more frequently administered in patients with higher-risk features. Two patients (3.2%) relapsed during surveillance, both achieved a complete response with salvage chemotherapy. No cancer-related deaths occurred, and treatment-related toxicity was minimal, with only one case of infertility reported.

Conclusion: Our findings confirm that stage I seminoma has an excellent prognosis regardless of the initial post-orchiectomy management approach. Carboplatin is a safe adjuvant treatment option. Surveillance may prevent unnecessary adjuvant treatment. Decisions regarding adjuvant therapy should be patient-centered, based on the individual risk of relapse and the potential toxicity of therapy.

Keywords: Seminoma, testicular neoplasm, human chorionic gonadotropin, beta subunit, adjuvant chemotherapy, surveillance

Öz

Amaç: Bu çalışma ile evre I seminom hastalarında cerrahi sonrası uzun dönem takip sonuçlarını ve nüks oranlarını değerlendirmeyi, bunun yanı sıra klinik pratikte aktif izlem ve adjuvan tedavinin etkinliğini ve güvenilirliğini incelemeyi amaçladık.

Yöntem: 2007-2024 yılları arasında kurumumuzda evre I seminom tanısı almış 63 hastanın verileri retrospektif olarak incelenmiş ve çalışmaya dahil edilmiştir. Tüm hastalara radikal inguinal orşiektomi uygulanmıştır. Hastaların demografik özellikleri, tümörlerin patolojik özellikleri, nüks oranları, tedaviye bağlı yan etkiler ve uzun dönem onkolojik sonuçlar istatistiksel olarak analiz edilmiştir.

Bulgular: Ortanca yaş 36, ortanca takip süresi 50,5 aydır. Tümörlerin %52,4'ü ≥4 cm, %33,3'ünde rete testis invazyonu vardır. Postoperatif dönemde 43 hastada aktif izlem, 18 hastada adjuvan karboplatin, 2 hastada ise adjuvan radyoterapi uygulanmıştır. Yüksek risk özellikleri olan hastalarda adjuvan tedavi

Address for Correspondence/Yazışma Adresi: Sercan Ön, MD, University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of Oncology, İzmir, Türkiye
E-mail: dr.sercanon@gmail.com
ORCID ID: orcid.org/0000-0003-1461-7485

Received/Geliş tarihi: 03.05.2025
Accepted/Kabul tarihi: 22.09.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

daha sık tercih edilmiştir. İzlem grubunda iki hastada (%4,7) nüks gelişmiş ve her ikisi de salvage kemoterapi sonrası tam yanıt vermiştir. Testis kanserine bağlı ölüm görülmemiş, ciddi tedaviye bağlı toksisite sadece bir infertilite olgusu ile sınırlı kalmıştır.

Sonuç: Bulgularımız evre I seminonum cerrahi sonrası tedavi seçiminden bağımsız olarak mükemmel prognoza sahip olduğunu doğrulamaktadır. Karboplatin güvenli bir adjuvan tedavi seçeneğidir. Aktif izlem, gereksiz adjuvan tedavilerin önüne geçebilir ve güvenli görünmektedir. Adjuvan tedavi kararı, nüks riski ve tedaviye bağlı toksisite göz önünde bulundurularak hasta tercihleri de göz önüne alınarak verilmelidir.

Anahtar kelimeler: Seminoma, testis neoplasm, insan koryonik gonodotropin, beta subunit, adjuvan kemotereapi, surveyans

Introduction

Germ cell testicular tumors are among the most common solid tumors in men under the age of 40. Globally, approximately 75.000 men are diagnosed with testicular cancer each year[1]. In our country, the annual incidence is around 1.700 cases. The disease is more prevalent among Caucasian populations, particularly in Scandinavian countries, and its incidence has been increasing over time[2]. Histologically, germ cell tumors are classified into two main groups: seminomas and non-seminomas. Distinguishing between these subtypes is critical, as they differ in clinical course and systemic treatment approaches. Approximately half of all germ cell testicular tumors are of pure seminoma histology[3].

Patients most commonly present with a painless testicular mass. Physical examination and scrotal ultrasonography constitute the first-line diagnostic steps. Three serum tumor markers have established roles in the diagnosis and management of testicular cancer: alpha-fetoprotein (AFP), the beta subunit of human chorionic gonadotropin (β-hCG), and lactate dehydrogenase (LDH). While elevated AFP levels are not typically observed in pure seminomas, mild elevations in β-hCG and LDH levels may be present. Radical inguinal orchiectomy provides both a definitive histological diagnosis and local tumor control. Staging after orchiectomy is performed using radiological imaging and postoperative tumor marker levels. According to the eighth edition of the American Joint Committee on Cancer (AJCC) staging system, disease confined to the testis is classified as stage I[4]. In the presence of persistently elevated postoperative LDH or β-hCG, the disease is classified as stage IS, which represents a distinct entity from stage I, it suggests residual disease, and is managed similarly to systemic disease. Approximately 80% of seminomas present as stage I at diagnosis[5].

The prognosis of stage I seminoma is excellent. Following orchiectomy, recurrence rates range between 13% and 20%[6-8]. In the presence of risk factors such as tumor size, rete testis invasion, and lymphovascular invasion (LVI), the risk of relapse can exceed 30%[9,10]. Without these risk factors, the recurrence rate may be as low as 4%[11]. Seminoma is a radiosensitive and chemosensitive disease. In the adjuvant setting, radiotherapy and single-agent carboplatin have reduced relapse rates to below 10%[12]. However, due to the increased long-term risk of secondary malignancies, radiotherapy is no longer recommended. Although adjuvant carboplatin is generally well tolerated, its efficacy diminishes in the presence of multiple risk factors, with relapse rates approaching 10%; the estimated number needed to treat to prevent one relapse ranges between 15 and 20 patients. Moreover, adjuvant carboplatin has not demonstrated a survival benefit. Salvage treatment administered at relapse achieves nearly 100% survival[13]. As a result, international guidelines recommend active surveillance as the preferred management strategy for stage I seminoma[14].

However, risk-adapted adjuvant treatment remains a common approach in clinical practice. Considering the potential for recurrence and the higher toxicity associated with salvage treatment in the event of relapse, some patients may prefer to receive adjuvant carboplatin. This complicates the process of making clinical decisions. There is a need for more definitive prognostic risk factors, to better guide therapy. Furthermore, the literature lacks sufficient long-term data regarding the toxicity of adjuvant carboplatin and its impact on late relapse. Therefore, in this study, we aimed to evaluate the clinical characteristics of patients with stage I seminoma managed with surveillance or adjuvant therapy in our institution to identify factors influencing the decision to administer adjuvant carboplatin and to assess long-term follow-up outcomes.

Materials and Methods

Data on patients diagnosed with stage I seminoma who underwent surgery at our hospital between 2007 and 2024 and had at least one year of follow-up were retrospectively collected from the hospital's medical record system. Since

approximately 80% of relapses occur within the first year, only patients aged 18 years or older with a minimum of one-year follow-up were included in the study[7]. Postoperative cross-sectional imaging had to confirm stage I disease according to the 8th edition of the AJCC staging system[4]. Patients were included only if postoperative tumor markers (AFP, β-hCG, and LDH) were within normal limits. Patients with non-seminomatous germ cell components (i.e., mixed histology) were excluded. Postoperative follow-up, including tumor markers and imaging, was performed following contemporary guideline recommendations. Patients with missing follow-up data were excluded from the analysis. Patients lost to follow-up were included up to the last available follow-up point.

Data recorded included patient age, tumor size, histopathological features (rete testis invasion, LVI), preoperative serum β-hCG and LDH levels, adjuvant treatment status, presence of recurrence, and survival status. Acute and chronic toxicities related to adjuvant treatment, as well as any development of secondary malignancies, were recorded from patient files and national health system records. In the event of recurrence, the site of recurrence, salvage treatment administered, and treatment response were evaluated. Recurrence was defined as retroperitoneal lymph node involvement and/or distant lymph node or visceral organ metastasis. Development of a germ cell tumor in the contralateral testis was considered a second primary malignancy. This study was conducted in accordance with the Declaration of Helsinki and was approved by the Non-Interventional Research Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital (approval no: 2025/02-01, date: 10.03.2025).

Statistical Analysis

Statistical analyses were performed using SPSS version 27.0. Means and standard deviations (SD) were calculated for normally distributed numerical variables, medians and interquartile ranges (IQR) were reported for non-normally distributed variables. Frequencies and percentages were used for categorical variables. The chi-square test evaluated the differences in clinical and pathological characteristics between patients who did and did not receive adjuvant therapy. A p-value <0.05 was considered statistically significant. Disease-free survival (DFS) was estimated using the Kaplan-Meier method. DFS was defined as the time from diagnosis to the date of relapse, death, or last follow-up, whichever occurred first.

Results

Data from 82 patients diagnosed with seminoma at our hospital were reviewed. Four patients (4.8%) were excluded from the study due to unavailable follow-up data. Among the remaining patients, one (1.2%) had stage IS disease, 12 (14.6%) had stage II, and two (2.4%) had stage III disease. A total of 63 patients (76.8%) were diagnosed with stage I seminoma and included in the data analysis. The median age of these patients was 34 years (IQR: 30-41), and five patients (8%) were 50 years or older. All patients underwent radical inguinal orchiectomy. According to pathological findings, the mean tumor size was 4.4 cm (SD±2.3), and in 52.4% of cases, the tumor measured greater than 4 cm. Testicular hilar soft tissue invasion was reported in only 11 patients, with five of them (7.9%) having confirmed invasion. Rete testis invasion was identified in 21 patients (33.3%), and LVI was present in 17 patients (27%).

Every patient underwent thorough blood count and biochemical testing before and after surgery. In the preoperative period, 20 patients (31.7%) had elevated β-hCG levels. Among these patients, the median β-hCG value was 12.5 U/L (IQR: 5-79), indicating mild elevations in most cases. Preoperative LDH levels were elevated in 19 patients (30.2%), with a median LDH level of 338 U/L (IQR: 264-449). Patients with tumors larger than 4 cm exhibited significantly elevated levels of β-hCG and LDH (p=0.045 and p=0.003, respectively). Postoperatively, β-HCG and LDH levels normalized in all patients.

All clinical stage I seminoma patients were assessed in the oncology clinic. Surveillance was recommended for 43 patients (68.3%), while 20 (31.7%) received adjuvant therapy. Among those receiving adjuvant treatment, two patients (10%) received radiotherapy, 17 (85%) received a single cycle of carboplatin at a dose of 7 area under the curve, and one patient (5%) received two cycles of carboplatin. The characteristics of patients who received adjuvant therapy are shown in Table 1. In summary, adjuvant treatment was more frequently administered in patients with tumor size >4 cm (p<0.001), rete testis invasion (p=0.009), elevated preoperative β-hCG (p=0.039), and elevated LDH levels (p=0.040). The presence of LVI did not significantly differ between patients who received adjuvant therapy and those who did not (p=0.28).

During a median follow-up period of 50.5 months (IQR: 31.3 to 80.9 months), two patients (3.2%) experienced a relapse. Both incidents occurred in the surveillance group, resulting

Table 1. Patients' characteristics				
	All patients	**Active surveillance**	**Adjuvant treatment**	**p-value**
Age	36.2 (SD±9.9)	36.1 (SD±10.2)	35.7 (SD±9.6)	
Tumor size				
≤4 cm	30 (47.6%)	27 (62.8%)	3 (15%)	p<0.001
>4 cm	33 (52.4%)	16 (37.2%)	17 (85%)	
Rete testis invasion				
Absent	35 (55.6%)	28 (65.1%)	7 (35%)	
Present	21 (33.3%)	9 (20.9%)	12 (60%)	p=0.009
Unknown	7 (11.1%)	6 (14%)	1 (5%)	
LVI				
Absent	31 (49.2%)	23 (53.5%)	8 (40%)	
Present	17 (27%)	9 (20.9%)	8 (40%)	p=0.2
Unknown	15 (23.8%)	11 (25.6%)	4 (20%)	
Preoperative β-hCG				
Normal	37 (58.7%)	28 (65.1%)	9 (45%)	
Elevated	20 (31.7%)	10 (23.3%)	10 (50%)	p=0.039
Unknown	6 (9.5%)	5 (11.6%)	1 (5%)	
Preoperative LDH				
Normal	34 (54%)	25 (58.1%)	9 (45%)	
Elevated	19 (30.2%)	9 (20.9%)	10 (50%)	p=0.04
Unknown	10 (15.9%)	9 (20.9%)	1 (%5)	
Relaps				
No	61 (96.8%)	41 (95.3%)	20 (100%)	p=0.46
Yes	2 (3.2%)	2 (4.7%)	0	

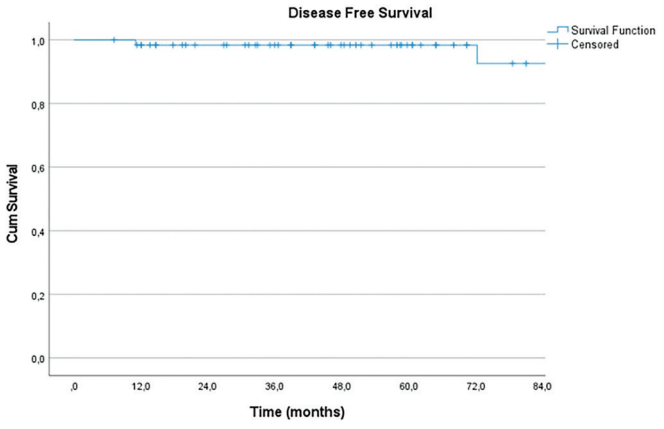
SD: Standard deviation, LVI: Lymphovascular invasion, β-hCG: Beta subunit of human chorionic gonadotropin, LDH: Lactate dehydrogenase

in a relapse rate of 4.7% within that group. One relapse took place at 12 months, while the other occurred at 74 months, with both relapses located in the retroperitoneal region (Graphic 1). Both patients were classified as good-risk and achieved remission following three cycles of bleomycin, etoposide, and cisplatin chemotherapy. Additionally, one patient who had undergone partial orchiectomy due to contralateral testicular atrophy, developed local recurrence and subsequently underwent radical orchiectomy. One patient was diagnosed with seminoma in the contralateral testis and subsequently received two cycles of adjuvant carboplatin following orchiectomy. There were no recorded deaths related to testicular cancer; however, one patient (1.6%) in the surveillance group died due to interstitial lung disease associated with scleroderma.

No secondary malignancies were identified as a result of adjuvant radiotherapy or chemotherapy. One patient who received a single cycle of carboplatin was diagnosed with infertility. Acute kidney injury was not observed in any patients treated with carboplatin. Grade 1-2 thrombocytopenia was reported in four patients, representing 22.2% of the cohort. All patients administered carboplatin also received granulocyte colony-stimulating factor support, and there were no hospitalizations due to febrile neutropenia. Furthermore, clinical side effects such as nausea, fatigue, or diarrhea were not documented in the patient records.

Discussion

In our tertiary referral center, where uro-oncological surgeries are commonly performed, 82 patients were diagnosed with seminoma over 17 years. Although seminomas are the most common testicular tumors in young men, they account for only about 1% of all solid tumors, making them relatively rare[1]. In comparison to non-seminomas, seminomas are typically diagnosed at a slightly older age, with peak incidence occurring between 35 and 39 years, whereas non-seminomas peak between 25 and



Grahpic 1. Disease-free survival

29 years. Notably, a second peak incidence for seminomas has been reported in individuals aged 50 to 55, particularly for the spermatocytic subtype[15]. The median age of our study population was consistent with the literature, and 8% of patients were over 50. However, we did not identify any cases of spermatocytic seminoma in the pathology records. Consistent with prior studies, approximately three-quarters of our patients presented with stage I disease.

Numerous studies have examined clinical and pathological factors that may predict relapse in stage I seminoma. Warde et al.[9] in a pooled analysis, identified two independent risk factors for relapse: tumor size greater than 4 cm and invasion of the rete testis. The prospective Swedish and Norwegian Testicular Cancer Group trial incorporated these factors and reported a relapse rate of 6.2% following a single cycle of adjuvant carboplatin. In that study, the relapse rate was 2% for patients without risk factors, compared to 9.2% for those with one or more risk factors[16]. Since that time, risk-adapted adjuvant therapy has been recognized as an option in clinical guidelines. In our study, we similarly demonstrated that considerations of tumor size and rete testis invasion were integral to the decision-making process regarding adjuvant treatment.

This study observed no relapses in patients who received adjuvant therapy, even with long-term follow-up. Although this rate is better than that commonly reported in the literature, it may be attributed to the small sample size of patients (n=20), who underwent adjuvant therapy. In stage I seminoma, adjuvant carboplatin may delay rather than prevent contralateral testicular tumors. A long-term observational study reported a higher incidence of such tumors in patients receiving carboplatin compared to the general population[17]. We did not find any evidence that adjuvant carboplatin increases the risk of late relapse in our data. However, consistent with the literature, the relapse rate among patients under active surveillance was 4.7%[16]. Notably, there were no deaths related to testicular cancer. Two patients who relapsed were successfully treated with salvage therapy. These findings align with existing literature and reinforce the value of a risk-adapted management approach[18].

Nevertheless, growing evidence suggests that current prognostic factors may not be entirely reliable. Earlier studies have been limited by selective cohorts, inconsistent pathological reporting, missing data, and lack of statistical power. Furthermore, few analyses have evaluated all possible risk factors simultaneously[19,20]. Since adjuvant therapy does not improve overall survival, surveillance remains the preferred approach. However, there is a compelling need for more reliable predictive markers of relapse. A recent large cohort study identified that testicular hilum invasion, LVI, and elevated preoperative levels of β-hCG and LDH can, either individually or in combination, serve as predictors of relapse. In this study, the relapse rate was recorded at 6% for patients without risk factors, escalating to 60% for those presenting all four indicators. Testicular hilum invasion emerged as the most significant prognostic factor[21]. In our study, testicular hilum invasion was reported in only a few cases and did not influence adjuvant therapy decisions. However, elevated preoperative tumor markers appeared to influence treatment decisions. Preoperative marker levels were also correlated with tumor size. Due to the small number of relapses and a limited number of patients receiving adjuvant therapy, statistical analysis of prognostic factors was not feasible in our study. Further prospective data are needed to assess whether testicular hilum invasion and preoperative marker elevation can guide adjuvant treatment decisions.

Before the 2000s, adjuvant radiotherapy targeting the retroperitoneal field was widely used in stage I seminoma. As research increasingly linked radiotherapy to secondary malignancies[22], the utilization of this treatment modality experienced a decline. In our cohort, only two patients received radiotherapy, both before 2005. No secondary malignancies were detected during long-term follow-up. Among patients treated with single-cycle carboplatin, no severe adverse effects were noted. Grade 1-2 thrombocytopenia was observed in four patients (22.2%), and no patient experienced neutropenic fever or required hospitalization. According to the literature, severe toxicity from adjuvant carboplatin is rare, with nausea, vomiting, fatigue, and hematologic toxicity being the most common side effects in approximately half of

the patients[12]. Due to the retrospective nature of our study, clinical toxicities were not systematically recorded and post-treatment blood counts were not routinely monitored, which may explain the lower reported rate of hematologic toxicity. During a median follow-up of 50 months, one patient developed infertility potentially related to carboplatin. Consistent with the literature, no association has been found between carboplatin and secondary malignancies[23]. Our findings support the short- and long-term safety of adjuvant carboplatin. Given the young age and future fertility expectations of most patients, sperm cryopreservation before chemotherapy remains essential.

Conclusion

In conclusion, our study provides evidence that risk-adapted adjuvant treatment is effective and safe. However, it is essential to note that some patients undergoing adjuvant therapy may have experienced overtreatment, as they might not have relapsed in the absence of such interventions. Conversely, although salvage therapy can yield high remission rates in patients who have relapsed, it typically involves more prolonged and toxic treatment regimens. These facts should be clearly communicated to patients, and adjuvant treatment decisions should be patient-centered. Our study had a long follow-up period, ensuring reliable long-term data, but the main limitation was underreporting of acute toxicity. Another limitation of this study includes the small sample size and potential for selection bias inherent to its retrospective, single-center design. Given the lack of national data on this subject in our country, multicenter studies with larger patient populations are needed.

Ethics

Ethics Committee Approval: This study was conducted in accordance with the Declaration of Helsinki and was approved by the Non-Interventional Research Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital (approval no: 2025/02-01, date: 10.03.2025).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Concept: S.Ö., D.A., Z.A., M.Z.K., Design: S.Ö., D.A., Z.A., M.Z.K., Data Collection or Processing: S.Ö., D.A., Analysis or Interpretation: S.Ö., D.A., Z.A., M.Z.K., Literature Search: S.Ö., Z.A., M.Z.K., Writing: S.Ö., D.A., Z.A., M.Z.K.

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

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

References

1. Global Cancer Observatory. International Agency for Research on Cancer. World Health Organization. Available from: https://gco.iarc.fr/ (Accessed: March 06, 2025).
2. McGlynn KA, Devesa SS, Sigurdson AJ, Brown LM, Tsao L, Tarone RE. Trends in the incidence of testicular germ cell tumors in the United States. Cancer. 2003;97:63-70.
3. Krag Jacobsen G, Barlebo H, Olsen J, et al. Testicular germ cell tumours in Denmark 1976-1980. Pathology of 1058 consecutive cases. Acta Radiol Oncol. 1984;23:239-47.
4. Brimo F, Srigley JR, Ryan CJ, et al. Testis. In: Amin MB, (editor). AJCC Cancer Staging Manual, 8th ed. New York, Springer; 2017:727.
5. Oldenburg J, Berney DM, Bokemeyer C, et al. Testicular seminoma and non-seminoma: ESMO-EURACAN Clinical Practice Guideline for diagnosis, treatment and follow-up. Ann Oncol. 2022;33:362-75.
6. Nayan M, Jewett MAS, Hosni A, et al. Conditional risk of relapse in surveillance for clinical stage I testicular cancer. Eur Urol. 2017;71:120-7.
7. Mortensen MS, Lauritsen J, Gundgaard MG, et al. A nationwide cohort study of stage I seminoma patients followed on a surveillance program. Eur Urol. 2014;66:1172-8.
8. Pierorazio PM, Albers P, Black PC, et al. Non-risk-adapted surveillance for stage I testicular cancer: critical review and summary. Eur Urol. 2018;73:899-907.
9. Warde P, Specht L, Horwich A, et al. Prognostic factors for relapse in stage I seminoma managed by surveillance: a pooled analysis. J Clin Oncol. 2002;20:4448-52.
10. Horwich A, Alsanjari N, A'Hern R, Nicholls J, Dearnaley DP, Fisher C. Surveillance following orchidectomy for stage I testicular seminoma. Br J Cancer. 1992;65:775-8.
11. Tandstad T, Ståhl O, Dahl O, et al. Treatment of stage I seminoma, with one course of adjuvant carboplatin or surveillance, risk-adapted recommendations implementing patient autonomy: a report from the Swedish and Norwegian Testicular Cancer Group (SWENOTECA). Ann Oncol. 2016;27:1299-304.
12. Oliver RT, Mead GM, Rustin GJ, et al. Randomized trial of carboplatin versus radiotherapy for stage I seminoma: mature results on relapse and contralateral testis cancer rates in MRC TE19/EORTC 30982 study (ISRCTN27163214). J Clin Oncol. 2011;29:957-62.
13. Chovanec M, Hanna N, Cary KC, Einhorn L, Albany C. Management of stage I testicular germ cell tumours. Nat Rev Urol. 2016;13:663-73.
14. NCCN Guidelines Version 2.2025 Testicular Cancer NCCN Guidelines Version 2.2025 Testicular Cancer-Pure Seminoma. Available from https://www.nccn.org/professionals/physician_gls/pdf/testicular.pdf (Accessed: April 20, 2025).
15. Park JS, Kim J, Elghiaty A, Ham WS. Recent global trends in testicular cancer incidence and mortality. Medicine (Baltimore). 2018;97:e12390.
16. Tandstad T, Ståhl O, Dahl O, et al. Treatment of stage I seminoma, with one course of adjuvant carboplatin or surveillance, risk-adapted

recommendations implementing patient autonomy: a report from the Swedish and Norwegian Testicular Cancer Group (SWENOTECA). Ann Oncol. 2016;27:1299-304.

17. Powles T, Robinson D, Shamash J, Moller H, Tranter N, Oliver T. The long-term risks of adjuvant carboplatin treatment for stage I seminoma of the testis. Ann Oncol. 2008;19:443-7.
18. Bumbasirevic U, Zivkovic M, Petrovic M, Coric V, Lisicic N, Bojanic N. Treatment options in stage I seminoma. Oncol Res. 2022;30:117-28.
19. Boormans JL, Mayor de Castro J, Marconi L, et al. Testicular tumour size and rete testis ınvasion as prognostic factors for the risk of relapse of clinical stage I seminoma testis patients under surveillance: a systematic review by the Testicular Cancer Guidelines Panel. Eur Urol. 2018;73:394-405.
20. Zengerling F, Kunath F, Jensen K, Ruf C, Schmidt S, Spek A. Prognostic factors for tumor recurrence in patients with clinical stage I seminoma undergoing surveillance-a systematic review. Urol Oncol. 2018;36:448-58.
21. Wagner T, Toft BG, Lauritsen J, et al. Prognostic factors for relapse in patients with clinical stage I testicular seminoma: a nationwide, population-based cohort study. J Clin Oncol. 2024;42:81-9.
22. van Leeuwen FE, Stiggelbout AM, van den Belt-Dusebout AW, et al. Second cancer risk following testicular cancer: a follow-up study of 1,909 patients. J Clin Oncol. 1993;11:415-24.
23. Dahl AA, Mykletun A, Fosså SD. Quality of life in survivors of testicular cancer. Urol Oncol. 2005;23:193-200.

DOI: 10.4274/anatoljmed.2025.18291
Anatol J Gen Med Res 2025;35(3):276-281

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Laryngeal Granulomas; Characteristics and Our Clinical Approach

Laringeal Granulomalar; Karakteristik Özellikleri ve Klinik Yaklaşımımız

Esma Altan[1], Dilara Söylemez[1], Elife Barmak[2]

[1]University of Health Sciences Türkiye, Ankara Etlik City Hospital, Department of Otorhinolaryngology, Ankara, Türkiye
[2]Ankara Yıldırım Beyazıt University Faculty of Health Sciences, Department of Speech and Language Therapy, Ankara, Türkiye

Cite as: Altan E, Söylemez D, Barmak E. Laryngeal granulomas; characteristics and our clinical approach. Anatol J Gen Med Res. 2025;35(3):276-281

Abstract

Objective: This study aims to present our clinical approach and treatment strategies for laryngeal granulomas, emphasizing etiological differentiation and individualized management.

Methods: This retrospective study included 65 patients diagnosed with laryngeal granulomas over a 3-year period. Patients were categorized into two main etiological groups: contact granulomas and intubation-induced granulomas. Clinical evaluations involved videolaryngostroboscopy, symptom assessment, and voice analysis when available. Treatment was tailored to etiology and included medical reflux management, voice therapy, surgical excision, and botulinum toxin injection in selected cases.

Results: Contact granulomas accounted for 90.8% of cases and were more prevalent in males. Laryngopharyngeal reflux emerged as the most common underlying factor in both subgroups. All patients received standardized anti-reflux therapy, which significantly reduced recurrence rates. In patients with persistent or bilateral lesions, botulinum toxin injection proved to be a valuable adjunctive treatment. Surgical excision was reserved for lesions causing airway compromise or not responding to conservative treatment, and was always combined with medical management to reduce recurrence risk. Among the 12 patients who underwent surgery, recurrence was more common in those who had excision alone than in those who also received botulinum toxin.

Conclusion: Our findings highlight the importance of identifying the underlying etiology in managing laryngeal granulomas. A multimodal, etiology-driven treatment strategy-particularly emphasizing reflux control-offers favorable outcomes. Botulinum toxin injection appears to enhance treatment success in refractory or recurrent cases and should be considered as part of individualized management protocols.

Keywords: Granuloma laryngeal, dysphonia, botulinum toxin

Öz

Amaç: Bu çalışma, laringeal granülomlara yönelik klinik yaklaşımımızı ve tedavi stratejilerimizi; etiyolojik ayrım ve bireyselleştirilmiş tedavi planlaması odağında sunmayı amaçlamaktadır.

Yöntem: Bu retrospektif çalışmada, 3 yıl içinde laringeal granülom tanısı alan 65 hasta değerlendirildi. Hastalar, kontakt granülom ve entübasyona bağlı granülom olmak üzere iki ana etiyolojik gruba ayrıldı. Klinik değerlendirmede videolaringostroboskopi, semptom sorgulaması ve mevcutsa ses analizleri kullanıldı. Tedavi planları etiyolojiye göre bireyselleştirildi; reflü tedavisi, ses terapisi, cerrahi eksizyon ve gerektiğinde botulinum toksin enjeksiyonu uygulandı.

Bulgular: Olguların %90,8'ini kontakt granülomlar oluşturdu ve erkeklerde daha sık görüldü. Her iki grupta da en yaygın etiyolojik neden laringofaringeal reflüydü. Tüm hastalara standart anti-reflü tedavisi başlandı ve bu tedavi rekürrens oranlarını anlamlı şekilde azalttı. Dirençli veya bilateral lezyonlarda

Address for Correspondence/Yazışma Adresi: Esma Altan, MD, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Department of Otorhinolaryngology, Ankara, Türkiye
E-mail: esmaaltan@gmail.com
ORCID ID: orcid.org/0000-0002-3080-3571

Received/Geliş tarihi: 28.08.2025
Accepted/Kabul tarihi: 22.09.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

botulinum toksin enjeksiyonu etkili bir tamamlayıcı yöntem olarak ön plana çıktı. Cerrahi eksizyon, hava yolunu daraltan veya konservatif tedaviye yanıt vermeyen hastalarda uygulandı ve her zaman medikal tedaviyle desteklendi. Sadece cerrahi uygulananlarda nüks oranı daha yüksekken, botulinum toksin ile kombine edilen olgularda başarı oranı daha yüksekti.

Sonuç: Laringeal granülomların tedavisinde altta yatan etiyolojinin belirlenmesi kritik öneme sahiptir. Özellikle reflü kontrolüne odaklanan multimodal ve bireyselleştirilmiş tedavi yaklaşımları daha başarılı sonuçlar sağlamaktadır. Botulinum toksin, dirençli ya da rekürren olgularda etkili bir seçenektir.

Anahtar Kelimeler: Laringeal granulomlar, disfoni, botulinum toksin

Introduction

Laryngeal granulomas are benign inflammatory masses typically located at the vocal process of the arytenoid cartilage[1]. They are also referred to by various names, including arytenoid granuloma and inflammatory polyp, vocal process granuloma, peptic granuloma, pyogenic granuloma, and laryngeal granuloma. Jackson[2] initially identified them as contact ulcers, while Clausen[3] later associated them with intubation-related trauma. Reported prevalence varies from 0.9% to 2.7% among adults with voice disorders, with a noted male predominance[4,5]. Vocal granulomas can be divided into two types: intubation granulomas and contact granulomas[6,7]. Despite similar histology, etiological distinctions exist, including associations with reflux and phonotraumatic behaviors[8–11]. Granulomas may occur with more than one underlying factor[12,13]. If there is a history of intubation, this supports the diagnosis of intubation granuloma. The most important underlying cause of contact granuloma are gastroesophageal reflux and laryngopharyngeal reflux[14,15]. Treatment resistance may arise from repeated irritation or unresolved underlying causes. Treatment options for contact granulomas include laser excision, microlaryngoscopic excision with a cold knife, and botulinum toxin injection; in addition to voice therapy, corticosteroid injection, inhaled corticosteroids, and zinc supplementation. Among these modalities, botulinum toxin injection has been widely regarded as the most effective approach, particularly in refractory cases[16,17]. While excision treatment is based on removal of the granuloma (especially in granulomas narrowing the airway); other treatments alleviate inflammation or reduce friction of the vocal cords during phonation.

The present study aims to present the descriptive characteristics, clinical course and treatment modalities of our patients with laryngeal granuloma in our clinical setting.

Materials and Methods

There is an ethics committee approval numbered 04/1206 on 17.04.2025 from Ankara Yıldırım Beyazıt University Faculty of Health Sciences, this study is conducted in University of Health Sciences Türkiye, Ankara Etlik City Hospital. Informed consent was obtained from all patients included in this study.

Exclusion Criteria

Patients were excluded from the study if they had any of the following:

- Incomplete clinical documentation.
- Follow-up duration shorter than 2 months.
- Coexisting laryngeal pathologies (e.g., vocal fold polyps, cysts, neoplastic lesions).
- History of prior laryngeal surgery unrelated to granuloma treatment.
- Non-compliance with anti-reflux or voice therapy that could not be verified.
- Patients, who were lost to follow-up during the observation period, were unaccounted for in the final analysis.
- Presence of concomitant vocal cord lesions.

A systematic evaluation was conducted on 65 patients aged 24-86 years were systematically evaluated.

1. Videolaryngostroboscopy (VLS) was employed to evaluate the closure pattern of the glottal cycle, vibration amplitude, mucosal fluctuation, presence of adynamic segments, vertical closure level, phase symmetry, and medial edge irregularity, supraglottic region, and posterior glottis.
2. Voice handicap index-10 (VHI-10), GRBAS and reflux symptom index (RSI) were used to evaluate voice perception.
3. The multi-dimensional voice program (MDVP) was employed for acoustic analysis of the voice: fundamental

frequency, intensity, perturbation measurements (Jitter, Shimmer), noise-to-harmonic ratio parameters were examined.

VHI, GRBAS, and voice analysis (MDVP) are not included in the findings of this study. Descriptive patient information and treatment methods are mentioned in this study.

At initial evaluation, VLS was performed in all patients. Although assessments such as VHI-10, GRBAS, RSI, and MDVP were indeed used during the clinical evaluation of some patients, the data were not equally available across the entire cohort due to the retrospective nature of the study and differences in clinical documentation over the 3-year period.

Surgical Intervention and Botulinum Toxin Application

Surgical excision of granulomas was performed under general anesthesia using direct microlaryngoscopy. A cold knife technique was employed in all cases, and excision was carried out with care to preserve the underlying vocal fold mucosa and arytenoid cartilage.

In selected patients, botulinum toxin type A (Botox®, Allergan Inc., Irvine, CA, USA) was administered in conjunction with excision. A dose of 2.5 units was injected directly into the thyroarytenoid muscle on the side of the lesion. In one patient with recurrence, a transcervical electromyography (EMG)-guided botulinum toxin injection was performed as an adjunctive procedure.

Antireflux Therapy

All patients received standardized antireflux therapy consisting of esomeprazole 40 mg twice daily and sucralfate 1 g three times daily, accompanied by dietary and lifestyle modifications. Treatment adherence and symptom control were assessed during follow-up visits at 2-month intervals.

All patients received postoperative anti-reflux therapy and voice hygiene counseling as part of a standardized treatment protocol.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 23.0 (IBM Corp., Armonk, NY). Numerical variables were presented as mean ± standard deviation, while categorical variables were presented as frequencies and percentages.

The Shapiro-Wilk test was used to assess the normality of numerical data. Homogeneity of variances was evaluated using Levene's test. Between-group comparisons of normally distributed variables were performed using the independent samples t-test. The chi-square test was used for comparing categorical data. A p-value of less than 0.05 was considered statistically significant.

Results

Our cohort included 65 patients aged between 24 and 86 years, with a mean age of 54.82±14.3 years. Of the patients, 21.5% were female and 78.5% were male (Table 1). Contact granulomas accounted for 90.8% of the cases, while 9.2% were intubation-related granulomas. Of the six patients with intubation granulomas, all were male (Table 1). In terms of localization, 49.2% (n=32) were located on the right vocal fold, 41.5% (n=27) on the left vocal fold, and 9.2% (n=6) were bilateral (Table 1).

Videostroboscopic examination revealed a decreased mucosal wave in all patients, while phase symmetry was preserved in the majority. No significant alterations were observed in the glottal closure pattern, vertical level of closure, amplitude of vibration, supraglottic activity, or the presence of adynamic segments.

The RSI revealed clinically significant reflux symptoms across the entire cohort, with a mean RSI score of 33.51±5.45. All patients received standardized antireflux therapy. Due to ethical considerations, a comparison group without antireflux therapy could not be created.

Among the five patients who developed recurrence, four had contact granulomas and were additionally referred for voice therapy due to persistent phonotraumatic behavior. One patient with an intubation granuloma was managed with surgical excision and pharmacologic therapy. All patients received voice hygiene counseling before initiating treatment.

Table 1. Demographic and clinical characteristics of patients with laryngeal granulomas

Variable	Category	n	%
Sex	Female	14	21.5%
Sex	Male	51	78.5%
Etiology	Contact	59	90.8%
Etiology	Intubation	6	9.2%
Side	Right	32	49.2%
Side	Left	27	41.5%
Side	Bilateral	6	9.2%

Surgical excision was indicated in cases where the granuloma posed a risk of airway compromise or was associated with clinically significant dyspnea, consistent with previously established criteria.

In total, 12 patients (18.46%) had a history of surgical intervention. Of these, five had previously undergone cold knife excision in an external center. Three of these patients experienced recurrence and were subsequently managed in our clinic with excision and botulinum toxin injection into the thyroarytenoid muscle (Table 2).

In our institution, seven primary patients underwent simultaneous cold knife granuloma excision and botulinum toxin injection in their initial session. Only two of these patients developed recurrence. The recurrence rate was 60% among patients treated with excision alone at external centers, compared to 28.5% in those treated with excision, plus botulinum toxin, at our clinic (Table 3).

Of the five patients requiring reintervention due to recurrence, four (80%) had contact granulomas and one (20%) had an intubation granuloma. The four recurrent contact granuloma patients included two males and one female; the recurrent intubation granuloma case was male. One patient required an additional EMG-guided transcervical botulinum toxin injection following relapse; the others responded to combined surgical and pharmacological management with complete resolution.

Patients were stratified into four follow-up categories: <3 months, 3-6 months, 6-12 months and >12 months. Follow-up durations were distributed as follows: 26.1% <3 months, 40% from 3-6 months, 8.45% from 6-12 months, and 8.45% >12 months (Table 4). Among patients with recurrent contact granulomas, three experienced recurrence at 4 months and one at 2 months. The patient with a recurrent intubation granuloma, experienced a recurrence at 4 months.

Discussion

Laryngeal granuloma should be evaluated according to its etiologic origin and treatment should be planned accordingly. The aim of this study was to classify the patients based on the presence of intubation-induced and contact granulomas and to evaluate the treatment response. In terms of clinical features, intubation granulomas are more common in women than men in the literature[18]; all 6 patients with intubation granulomas were men in our study. Contact granulomas showed a strong male predominance, consistent with global data on voice misuse and vocal process trauma in men[19]. In many studies in the literature, proton pump inhibitors are recommended as the first anti-reflux treatment[20]. All patients

Table 2. Detailed surgical patients data

Patient ID	Surgery history	Botulinum toxin during surgery	Recurrence	Granuloma type	Etiology
1	Only excision (external centre)	No	Yes	Contact	Voice misuse
2	Only excision (external centre)	No	Yes	Contact	Voice misuse
3	Only excision (external centre)	No	No	Contact	Voice misuse
4	Only excision (external centre)	No	No	Contact	Voice misuse
5	Only excision (external centre)	No	Yes	Intubation	Intubation trauma
6	Primary excision+botulinum toxin	Yes	Yes	Contact	Voice misuse
7	Primary excision+botulinum toxin	Yes	Yes	Contact	Voice misuse
8	Primary excision+botulinum toxin	Yes	No	Contact	Voice misuse
9	Primary excision+botulinum toxin	Yes	No	Contact	Voice misuse
10	Primary excision+botulinum toxin	Yes	No	Contact	Voice misuse
11	Primary excision+botulinum toxin	Yes	No	Contact	Voice misuse
12	Primary excision+botulinum toxin	Yes	No	Contact	Voice misuse

Table 3. Surgical treatments and recurrence

Patient group	Treatment	Number of patients	Recurrence (%)
Patients treated only in an external center	Cold knife granuloma excision only	5	3 (60%)
Primary patients treated in our clinic	Cold knife excision+botulinum toxin (1st session)	7	2 (28.5%)

Table 4. Follow-up time/duration		
Follow-up time	**Number of patients**	**Total (%)**
0-3 months	17	26.1
3-6 months	26	40
6-12 months	11	16.9
>12 months	11	16.9
Total	65	100

were started on esomeprazole 40 mg twice daily and sucralfate 1 g three times daily, accompanied by lifestyle and dietary modifications for reflux management. Follow-up evaluations were conducted at 2-month intervals, and treatment was continued or adjusted based on clinical improvement. Among the 59 patients with contact granulomas, 48 (81.3%) showed significant regression following 2 months of standardized anti-reflux treatment. Contact granulomas largely regressed within approximately 3 months-however, the resolution timelines (median 9-12 months) seen in other series[21] indicate variability. In the literature, steroid treatment is at the forefront intubation granulomas[22]. In contrast to contact granulomas, surgical intervention, including surgical excision and steroid injection alongside botox injection, was undertaken in only 1 of 6 patients with intubation granulomas, while the other 5 regressed with anti-reflux treatment and inhaled steroid[23,24]. Most of the granulomas in other patients with contact granuloma regressed after 3 months of anti-reflux treatment. 18.46% of the patients underwent surgical excision. Among the five patients who experienced recurrence, three had previously undergone excision alone in external centers without adjunctive botulinum toxin injection. Five patients who developed recurrence and underwent a second surgical intervention, of whom 4 (80%) had contact granulomas and 1 (20%) had an intubation-related granuloma. We believe that the continued mechanical trauma from the vocal process of the arytenoid may have contributed to the recurrence in the absence of muscle relaxation. In contrast, two of our primary cases experienced recurrence despite receiving both excision and botulinum toxin injection. However, due to the small sample size, we are unable to determine a clear causal factor in these cases, and further studies are warranted. Botox application, which is the gold standard treatment for granulomas, was first described by Fink et al.[25]. Damrose and Damrose[26] reported that transcervical Botox application is safe and effective in the treatment of granulomas. The mechanism of action of Botox treatment is to eliminate the effect of repeated trauma by preventing hard glottal closure with temporary paralysis of the thyroarytenoid muscle. It has been shown that this is the most effective treatment option in patients who do not respond to drug treatment[27,28]. In our clinical practice, we not only perform granuloma excision; if there is a granuloma that closes the posterior glottis and causes airway obstruction, botulinum toxin injection is planned together with surgical excision. 2.5 units of Botox were injected into the thyroarytenoid muscle on the side of the lesion. Of the 5 patients with recurrence, 3 underwent granuloma excision at an external center and 2 underwent granuloma excision and botox. In the follow-up after surgery for recurrence, only 1 of these patients experienced a recurrence of complaints and transcervical botox was applied to them under EMG guidance.

One of the main strengths of our study is the inclusion of both surgical and non-surgical treatment approaches, allowing for a more comprehensive understanding of clinical outcomes in laryngeal granuloma cases. The favorable response to anti-reflux therapy in non-surgical patients highlights the importance of medical management in selected cases. In addition, surgical excision demonstrated effective results in refractory granulomas, particularly when combined with postoperative reflux control. However, our study also has certain limitations. The retrospective design may introduce selection bias, and the lack of randomization limits causal inferences. Moreover, patient adherence to medical therapy was not objectively measured, which may have influenced the outcomes. Recurrence remains a concern in both treatment arms, emphasizing the chronic and multifactorial nature of laryngeal granulomas.

Study Limitations

This study has several limitations. First, its retrospective design may have introduced selection bias and limited the consistency of available data. Second, follow-up duration was relatively short in some patients, and pathology reports could not be accessed for all surgically treated cases, which restricted a comprehensive evaluation of outcomes. Third, perceptual and acoustic voice analyses (VHI-10, GRBAS, MDVP) were not consistently documented across the entire cohort, preventing a systematic analysis of correlations between clinical outcomes and voice measures. Future prospective studies with standardized data collection and long-term follow-up are warranted to overcome these limitations.

Although the RSI questionnaire was used as part of the clinical evaluation, pre- and post-treatment scores were not consistently recorded in all patients. Therefore, changes in RSI values could not be statistically evaluated. This represents a limitation of our retrospective data set.

Conclusion

Laryngeal granulomas should be managed based on their underlying etiology, as treatment responses may differ between contact and intubation-related granulomas. While most patients respond well to conservative therapy, including antireflux management and voice therapy, surgical excision combined with botulinum toxin injection may be required in recurrent or obstructive cases. Early identification of risk factors and adherence to individualized treatment protocols can improve patient outcomes and reduce recurrence rates.

Ethics

Ethics Committee Approval: There is an ethics committee approval numbered 04/1206 on 17.04.2025 from Ankara Yıldırım Beyazıt University Faculty of Health Sciences, this study is conducted in University of Health Sciences Türkiye, Ankara Etlik City Hospital.

Informed Consent: Informed consent was obtained from all patients included in this study.

Footnotes

Authorship Contributions

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

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

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

References

1. Sun GB, Sun N, Tang HH, Zhu QB, Wen W, Zheng HL. Zinc sulfate therapy of vocal process granuloma. Eur Arch Otorhinolaryngol. 2012;269:2087-90.
2. Jackson C. Contact ulcer granuloma and other laryngeal complications of endotracheal anesthesia. Anesthesiology. 1953;14:425-36.
3. Clausen RJ. Unusual sequela of tracheal intubation. Proc R Soc Med. 1932;25:1507.
4. Ylitalo R, Lindestad PA. Laryngeal findings in patients with contact granuloma: a long-term follow-up study. Acta Otolaryngol. 2000;120:655-9.
5. al-Dousary S. Vocal process granuloma. Ear Nose Throat J. 1997;76:382-6, 387.
6. Hong-Gang D, He-Juan J, Chun-Quan Z, Guo-Kang F. Surgery and proton pump inhibitors for treatment of vocal process granulomas. Eur Arch Otorhinolaryngol. 2013;270:2921-6.
7. Wang CT, Lai MS, Lo WC, Liao LJ, Cheng PW. Intralesional steroid injection: an alternative treatment option for vocal process granuloma in ten patients. Clin Otolaryngol. 2013;38:77-81.
8. Karkos PD, George M, Van Der Veen J, et al. Vocal process granulomas: a systematic review of treatment. Ann Otol Rhinol Laryngol. 2014;123:314-20.
9. Ylitalo R, Ramel S. Extraesophageal reflux in patients with contact granuloma: a prospective controlled study. Ann Otol Rhinol Laryngol. 2002;111:441-6.
10. Leonard R, Kendall K. Effects of voice therapy on vocal process granuloma: a phonoscopic approach. Am J Otolaryngol. 2005;26:101-7.
11. Martins RH, Branco A, Tavares EL, Iyomasa RM, Carvalho LR, Henry MA. Laryngeal and voice disorders in patients with gastroesophageal symptoms. Correlation with pH-monitoring. Acta Cir Bras. 2012;27:821-8.
12. Hillel AT, Lin LM, Samlan R, Starmer H, Leahy K, Flint PW. Inhaled triamcinolone with proton pump inhibitor for treatment of vocal process granulomas: a series of 67 granulomas. Ann Otol Rhinol Laryngol. 2010;119:325-30.
13. Rimoli CF, Martins RHG, Catâneo DC, Imamura R, Catâneo AJM. Treatment of post-intubation laryngeal granulomas: systematic review and proportional meta-analysis. Braz J Otorhinolaryngol. 2018;84:781-9.
14. Haggitt RC. Histopathology of reflux-induced esophageal and supraesophageal injuries. Am J Med. 2000;108(suppl 4a):s109-11.
15. Li Y, Xu G, Zhou B, et al. Effects of acids, pepsin, bile acids, and trypsin on laryngopharyngeal reflux diseases: physiopathology and therapeutic targets. Eur Arch Otorhinolaryngol. 2022;279:2743-52.
16. Carroll TL, Gartner-Schmidt J, Statham MM, Rosen CA. Vocal process granuloma and glottal insufficiency: an overlooked etiology? Laryngoscope. 2010;120:114-20.
17. Chen M, Chen J, Yang Y, Li CJ, Wu HT, Chen L. Conservative treatment versus surgery for laryngeal contact granuloma: a prospective study. ORL J Otorhinolaryngol Relat Spec. 2018;80:307-16.
18. Lemos EM, Sennes LU, Imamura R, Tsuji DH. Vocal process granuloma: clinical characterization, treatment and evolution. Braz J Otorhinolaryngol. 2015;71:494-8.
19. Yilmaz T, Süslü N, Atay G, Özer S, Günaydin RÖ, Bajin MD. Recurrent contact granuloma: experience with excision and botulinum toxin injection. JAMA Otolaryngol Head Neck Surg. 2013;139:579-83.
20. Havas TE, Priestley J, Lowinger DS. A management strategy for vocal process granulomas. Laryngoscope. 1999;109:301-6.
21. Rimoli CF, Martins RHG, Catâneo DC, et al. Treatment of post-intubation laryngeal granulomas: systematic review and proportional meta-analysis. Braz J Otorhinolaryngol. 2018;84:781-9.
22. Roh HJ, Goh EK, Chon KM, Wang SG. Topical inhalant steroid (budesonide, Pulmicort nasal) therapy in intubation granuloma. J Laryngol Otol. 1999;113:427-32.
23. Hillel AT, Lin LM, Samlan R, Starmer H, Leahy K, Flint PW. Inhaled triamcinolone with proton pump inhibitor for treatment of vocal process granulomas: a series of 67 granulomas. Ann Otol Rhinol Laryngol. 2010;119:325-30.
24. Rudman JR, McGee CS, Diaz J, Rosow DE. Assessing the utility of non-surgical treatments in the management of vocal process granulomas. J Laryngol Otol. 2020;134:68-73.
25. Fink DS, Achkar J, Franco RA, Song PC. Interarytenoid botulinum toxin injection for recalcitrant vocal process granuloma. Laryngoscope. 2013;123:3084-7.
26. Damrose EJ, Damrose JF. Botulinum toxin as adjunctive therapy in refractory laryngeal granuloma. J Laryngol Otol. 2008;122:824-8.
27. Ho CF, Lee YC, Hsin LJ, Lee LA, Li HY, Fang TJ. Low-dose LEMG-guided botulinum toxin type a injection for intractable vocal process granulomas. J Voice. 2022;36:277-82.
28. Tsai SW, Ma YF, Shih LC, Tsou YA, Sung CK. Operative and conservative management of laryngeal contact granuloma: a network analysis and systematic review. J Voice. 2021;35:300-6.

DOI: 10.4274/anatoljmed.2025.85619
Anatol J Gen Med Res 2025;35(3):282-287

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Surgical Perspective on Neoadjuvant Therapy in Resectable Gastric Cancer

Rezektabl Mide Kanserinde Neoadjuvan Tedaviye Cerrahi Bakış Açısı

Akay Edizsoy[1], Ogün Aydoğan[1], Ahmet Ege Sakur[2], Erdem Barış Cartı[2]

[1]Aydın Adnan Menderes University Faculty of Medicine, Department of General Surgery, Division of Surgical Oncology, Aydın, Türkiye
[2]Aydın Adnan Menderes University Faculty of Medicine, Department of General Surgery, Aydın, Türkiye

Cite as: Edizsoy A, Aydoğan O, Sakur AE, Cartı EB. Surgical perspective on neoadjuvant therapy in resectable gastric cancer. Anatol J Gen Med Res. 2025;35(3):282-287

Abstract

Objective: Perioperative chemotherapy is recommended in Western guidelines for resectable stage ≥T2 gastric cancer. However, its feasibility in real-world settings, particularly when upfront surgery is a viable option, remains uncertain.

Methods: A single-center, retrospective observational study was conducted that included 22 patients diagnosed with clinically resectable, non-obstructive, non-bleeding, locally advanced gastric cancer (T3-T4 and/or cN+). Resectability and clinical staging were assessed using computed tomography (CT), positron emission tomography-CT and endoscopic ultrasonography. Patients were categorized based on treatment completion and postoperative outcomes. Demographic and clinical variables were analyzed using descriptive statistical methods.

Results: Of the 22 patients, 5 (22.7%) were unable to complete neoadjuvant chemotherapy (NAC) due to tumor-related bleeding (n=4) or disease progression (n=1). Seventeen patients (77.3%) completed NAC and underwent surgery; among them, 4 (23.5%) experienced major postoperative complications, including anastomotic leak (n=1), duodenal stump leak (n=2), and pancreatic fistula (n=1). Overall, 7 patients (31.8%) developed serious complications, and 2 patients (9.1%) died of complete anastomotic dehiscence.

Conclusion: A significant proportion of patients did not complete NAC because of gastric cancer-related complications, which necessitated early surgery under suboptimal conditions. These patients experienced high rates of postoperative morbidity and mortality. Based on these findings and in accordance with East Asian Guidelines, our institution now favors upfront surgery followed by adjuvant chemotherapy for comparable cases. Future randomized trials should incorporate real-world experience to better inform patient selection and treatment planning.

Keywords: Gastric, cancer, resectability, perioperative, neoadjuvant, complications

Öz

Amaç: Rezektabl evre ≥T2 mide kanserinde perioperatif kemoterapi, Batı kılavuzlarında önerilmektedir. Ancak, özellikle doğrudan cerrahinin uygulanabilir bir seçenek olduğu durumlarda, gerçek yaşam koşullarındaki uygulanabilirliği belirsizliğini korumaktadır.

Yöntem: Klinik olarak rezektabl, obstrüksiyon ve aktif kanaması olmayan, lokal ileri mide kanseri (T3-T4 ve/veya cN+) tanısı konmuş 22 hastayı içeren tek merkezli, retrospektif gözlemsel bir çalışma gerçekleştirildi. Rezektabilite ve klinik evreleme; bilgisayarlı tomografi, pozitron emisyon tomografisi ve endoskopik ultrasonografi kullanılarak değerlendirildi. Hastalar, tedavi tamamlama durumu ve postoperatif sonuçlarına göre kategorize edildi. Demografik ve klinik değişkenler tanımlayıcı istatistiksel yöntemlerle analiz edildi.

Bulgular: Yirmi iki hastanın 5'i (%22,7), tümörle ilişkili kanama (n=4) veya hastalık progresyonu (n=1) nedeniyle neoadjuvan kemoterapiyi (NAK) tamamlayamadı. On yedi hasta (%77,3) NAK'ı tamamladı ve cerrahi uygulandı; bu hastaların 4'ünde (%23,5) anastomoz kaçağı (n=1), duodenum güdük

Address for Correspondence/Yazışma Adresi: Assoc. Prof., Erdem Barış Cartı, Aydın Adnan Menderes University Faculty of Medicine, Department of General Surgery, Aydın, Türkiye
E-mail: erdemcarti@yahoo.com
ORCID ID: orcid.org/0000-0002-7139-0339

Received/Geliş tarihi: 14.08.2025
Accepted/Kabul tarihi: 02.10.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

kaçağı (n=2) ve pankreas fistülü (n=1) gibi majör postoperatif komplikasyonlar gelişti. Genel olarak, 7 hastada (%31,8) ciddi komplikasyon gelişti ve 2 hasta (%9,1) tam anastomoz ayrışması nedeniyle hayatını kaybetti.

Sonuç: NAK'ı, mide kanserine bağlı komplikasyonlar nedeniyle tamamlayamayan önemli sayıda hasta, suboptimal koşullarda erken cerrahiye ihtiyaç duydu. Bu hastalarda postoperatif morbidite ve mortalite oranları yüksekti. Bu bulgular doğrultusunda ve Doğu Asya kılavuzlarıyla uyumlu olarak, kurumumuz benzer olgularda upfront cerrahi sonrası adjuvan kemoterapiyi tercih etmektedir. Gelecekteki randomize çalışmalar, hasta seçimi ve tedavi planlamasında gerçek yaşam deneyimlerini de dikkate almalıdır.

Anahtar Kelimeler: Mide, kanser, rezektabilite, perioperatif, neoadjuvan, komplikasyonlar

Introduction

Surgical approaches to gastric cancer have evolved in parallel with advancements in operative techniques and systemic treatment strategies. The concept of perioperative chemotherapy gained prominence in Europe through the MAGIC and FNCLCC/FFCD trials. The subsequent adoption of the 5-fluorouracil, leucovorin, oxaliplatin, docetaxel regimen improved treatment tolerability and completion rates, enabling more patients to proceed with adjuvant chemotherapy[1].

The MAGIC, ACCORD07, and CRITICS trials were among the first to demonstrate survival benefits of perioperative or neoadjuvant chemotherapy (NAC), particularly in non-East Asian populations. More recent trials, such as PRODIGY and ARTIST 2, further reinforced this benefit by showing significant tumor downstaging and improvements in disease-free survival[2-5]. According to current international guidelines, the indication for neoadjuvant or perioperative chemotherapy in gastric cancer varies by tumor stage and nodal status. The National Comprehensive Cancer Network (NCCN) 2024 and the European Society for Medical Oncology 2022 Guidelines recommend perioperative chemotherapy for patients with clinical stage T2 or higher or those with node-positive (cN+) disease. Although interest in neoadjuvant strategies continues to grow, their applicability across all patient groups remains uncertain. Tumor progression during treatment can preclude surgery, and the specific contribution of surgery-particularly its quality and timing-has yet to be clearly defined[6,7].

In East Asia, different strategies are employed to manage resectable T3-T4 and/or cN+ gastric cancer. The 2023 Chinese Society of Clinical Oncology (CSCO) Guidelines upgraded NAC to a grade I recommendation for clinical stage III (cT3-T4 Nany) gastric cancerbased on increasing evidence supporting improved R0 resection rates and survival outcomes[8]. In contrast, the 2024 Korean Gastric Cancer Association Guideline continues to prioritize upfront D2 surgery for resectable T3-T4 and/or cN+ disease and reserves NAC or perioperative therapy for high-risk or borderline patients[9]. The 2021 Japanese Gastric Cancer Treatment Guidelines similarly limit NAC to more advanced cases, such as cT4b tumors or bulky nodal disease. With the exception of the recent CSCO update, most Asian guidelines continue to favor initiating treatment with upfront surgery[10].

This variation underscores the importance of precise staging and individualized treatment planning, tailored to both tumor burden and regional practice standards. In this study, we evaluated the clinical course and outcomes of patients with resectable gastric cancer who were scheduled for neoadjuvant therapy at our institution, focusing specifically on patients who failed to complete treatment; the necessity and timing of surgery among such patients; and the postoperative outcomes of patients who completed NAC. Based on these findings, we seek to highlight the ongoing relevance of upfront surgery as a viable approach in appropriately selected patients.

Materials and Methods

Study Design and Patient Selection

This retrospective observational study was conducted at Aydın Adnan Menderes University Hospital and included patients diagnosed with resectable gastric adenocarcinoma between January 2022 and March 2024. A total of 38 patients were referred to the medical oncology department for NAC. Of these, 16 did not return to our institution to initiate treatment. Of these, 10 patients were later admitted to our center for adjuvant therapy, while 6 were lost to follow-up. Therefore, the final analysis included 22 patients who received NAC and underwent surgery at our institution.

All patients were clinically staged as T3-T4 and/or cN+, with no evidence of gastric outlet obstruction or active gastrointestinal bleeding. Resectability was assessed using contrast-enhanced thoracoabdominal computed

tomography (CT), positron emission tomography-CT (PET-CT), and endoscopic ultrasonography (EUS), in accordance with the NCCN 2024 Guidelines. All treatment decisions were made in multidisciplinary tumor board meetings.

Data Collection and Variables

Demographic and clinical data were retrieved from patient charts and the institutional electronic medical record system. Variables included age, sex, tumor location, clinical stage, NAC regimen, and completion status, surgical approach, postoperative complications, and treatment-related mortality. Surgical outcomes, including anastomotic leakage, duodenal stump leakage, and in-hospital death, were also recorded.

Patient Grouping

Based on clinical course, patients were divided into the following three groups:

1. Patients who failed to complete NAC due to bleeding or progression and underwent early surgery.

2. Patients who completed NAC, underwent surgery after 4-6 weeks, and experienced major postoperative complications.

3. Patients who completed NAC underwent surgery after 4-6 weeks and had an uneventful postoperative course.

Among the five patients who did not complete NAC, four required surgery due to gastrointestinal bleeding. All underwent two sessions of endoscopic hemostasis, which included argon plasma coagulation, cauterization, and epinephrine injection. Surgical intervention was performed when bleeding could not be controlled.

Inclusion and Exclusion Criteria

Patients aged 18 years or older with histologically confirmed gastric adenocarcinoma were eligible for inclusion. All patients had clinically staged T3 or T4 tumors and/or cN+ status, as determined by thoracoabdominal contrast-enhanced CT, PET-CT, or EUS. Only patients who were deemed resectable with curative intent according to the NCCN 2024 guidelines and who did not have gastric outlet obstruction or active upper gastrointestinal bleeding at the time of diagnosis were included. Furthermore, all patients were required to be referred for NAC and to subsequently undergo surgery at our institution.

Patients were excluded if they had distant metastases, synchronous malignancies, or had received definitive chemoradiotherapy without surgical intent. Those who underwent upfront surgery without neoadjuvant treatment or who had surgery performed at an external institution were also excluded. Incomplete or missing medical records required for analysis constituted an additional exclusion criterion.

Statistical Analysis

Descriptive statistical methods were used to summarize the data. Continuous variables were presented as mean ± standard deviation, and categorical variables as counts and percentages. Clinical outcomes and complication rates were compared across the three patient groups. Statistical analyses were performed using SPSS version 29 (IBM Corp., Armonk, NY, USA).

Ethics Statement

This study was approved by the Ethics Committee of Aydın Adnan Menderes University Faculty of Medicine (approval no: 04, date: 27.06.2025). Written informed consent was obtained from all participants prior to data collection. All procedures were conducted in accordance with the ethical standards of the institutional research committee and the Declaration of Helsinki.

Results

A total of 22 patients with clinically resectable, non-obstructive, non-bleeding, locally advanced gastric cancer (T3-T4 and/or cN+) were included in the final analysis. The mean age was 61.4±9.2 years; 68.2% (n=15) were male. At diagnosis, 54.5% (n=12) of patients had T3 tumors, 45.5% (n=10) had T4 tumors, and 81.8% (n=18) were clinically cN+.

Five patients (22.7%) were unable to complete NAC: four because of tumor bleeding and one because of disease progression leading to peritoneal carcinomatosis. Among these, two patients developed complete anastomotic dehiscence and died in the postoperative period. One patient developed a duodenal stump leak that resolved with conservative management, and another recovered uneventfully after emergency surgery. This subgroup had a postoperative morbidity rate of 60% and a mortality rate of 40%.

Seventeen patients (77.3%) completed neoadjuvant therapy and underwent surgery within 4-6 weeks. Of these, four (23.5%) developed postoperative complications: one patient had an anastomotic leak, two patients developed duodenal

stump leaks, and one patient who underwent distal pancreatectomy developed a pancreatic fistula confined to a large pseudocyst in the pancreatic bed. Duodenal stump leaks were resolved conservatively. The patient with the anastomotic leak was successfully treated with endoscopic stenting. The remaining 13 patients (76.5%) had uneventful postoperative recoveries.

Overall, major postoperative complications occurred in seven patients (31.8%) across all subgroups, and mortality occurred in two of 22 patients (9.1%). However, in patients who could not complete neoadjuvant therapy, morbidity and mortality rates were 75% and 50%, respectively (Table 1).

Discussion

While NAC is endorsed by Western guidelines and appears to yield better survival outcomes in selected patient groups, our study sought to emphasize two critical issues often overlooked: postoperative morbidity in patients who complete NAC, and increased morbidity and mortality in patients who fail to complete it.

In our series, 4 of 17 patients (23.5%) who completed NAC and underwent surgery after 4-6 weeks experienced major postoperative complications, including anastomotic leak, duodenal stump leak, and pancreatic fistula. This rate is comparable to the complication rates reported in OGSG1205 (26.3%)[11] and Wang et al.[12] (20.6%). Furthermore, this complication rate contrasts with findings from a recent study by Aydoğan et al.[13] conducted at our institution, which reported no worse survival outcomes for patients undergoing upfront surgery than for those receiving NAC. These observations suggest that completing NAC does not eliminate postoperative morbidity and that complication risk remains substantial.

Among the 5 patients who could not complete NAC due to bleeding or disease progression, 3 experienced major complications, and 2 died from anastomotic dehiscence. This corresponds to a morbidity rate of 60% and a mortality rate of 40; these rates are markedly higher than those in the group that completed NAC[14,15]. These findings underline the severe risks encountered during surgery among patients who are unable to complete chemotherapy. Dağıstanlı[16] reported postoperative morbidity and mortality rates of 21.7% and 4.5%, respectively, in patients with locally advanced gastric cancer who did not receive neoadjuvant therapy; additional organ resection was required in more than half of cases because of invasion of adjacent organs. Based on these findings, the authors suggested that the absence of neoadjuvant therapy might contribute to increased surgical complexity and morbidity. In contrast, our results indicate that high morbidity and mortality occurred particularly among patients who could not complete neoadjuvant therapy and required urgent surgery for tumor-related complications, whereas upfront surgery in selected patients without obstruction or active bleeding was not associated with worse outcomes.

In all such cases, chemotherapy discontinuation was due to tumor-related complications rather than systemic toxicity. Bleeding and progression disrupted the treatment plan and necessitated emergency (unplanned) surgery under suboptimal conditions. This highlights the tumor's clinical course and biological behavior as critical barriers to completion of NAC[17].

Table 1. Demographic and clinical characteristics according to neoadjuvant therapy completion and surgical outcomes

Variable	All patients (n=22)	Unable to complete NAC (n=5)	Completed NAC (n=17)
Age (mean ± SD)	61.4±9.2	62.6±8.4	60.0±9.3
Sex (male)	15 (68.2%)	3 (60.0%)	12 (70.6%)
T stage (T3)	12 (54.5%)	2 (40.0%)	10 (58.8%)
T stage (T4)	10 (45.5%)	3 (60.0%)	7 (41.2%)
Node-positive	18 (81.8%)	4 (80.0%)	14 (82.4%)
Anastomotic leak	3 (13.6%)	2 (40%)	1 (5.9%)
Duodenal stump leak	3 (13.6%)	1 (20%)	2 (11.8%)
Pancreatic leak	1 (4.5%)	0 (0%)	1 (5.9%)
Morbidity	9 (40.9%)	3 (60.0%)	4 (23.5%)
Mortality	2 (9.1%)	2 (40.0%)	0 (0%)

SD: Standard deviation, NAC: Neoadjuvant chemotherapy

One patient developed peritoneal carcinomatosis during NAC and became inoperable. This case exemplifies the risk of disease progression during chemotherapy, which can entirely eliminate the window for curative surgery, and is especially concerning for patients with a high tumor burden.

Unplanned surgeries in patients who failed to complete NAC led to poor outcomes, including mortality and serious morbidity. These results are likely attributable to the urgent timing of surgery and the compromised condition of the patient[18].

In patients with resectable T3-T4 gastric cancer, the possibility that early surgery will be required because of tumor-related complications undermines the rationale for offering NAC to patients with earlier-stage disease (e.g., borderline T2 or low-burden N+). Such an approach would conflict with the principles of oncologic safety[19].

Of the 38 patients initially planned for NAC in our center, 16 sought treatment elsewhere. Ten of these were confirmed to have undergone surgery at other institutions and later returned for adjuvant therapy, while no data could be obtained for the remaining six. These patterns suggest that NAC adoption may still be limited among surgeons in our region, potentially due to similar challenges encountered in clinical practice[20,21].

Although studies comparing NAC+surgery+adjuvant therapy to surgery+adjuvant therapy have demonstrated survival benefits, they often exclude patients who fail to complete NAC. This exclusion may lead to an overestimation of the real-world efficacy of neoadjuvant treatment. Accordingly, our institutional strategy aligns more closely with East Asian guidelines that recommend upfront surgery followed by adjuvant therapy in resectable, non-obstructive, non-bleeding gastric cancer cases[22].

Taken together, our findings indicate that, although NAC is supported by international guidelines, its applicability in real-world practice remains limited, particularly for specific patient subsets. Patients who fail to complete NAC face substantial risks of complications and mortality, and even among those who complete treatment, significant morbidity persists. These real-world challenges, such as low treatment completion rates and high rates of postoperative complications, should be carefully considered when assessing the feasibility of neoadjuvant strategies in the management of gastric cancer.

Study Limitations

This study has several limitations inherent to its retrospective and single-center design. The relatively small sample size limits the generalizability of the findings, and potential selection bias may have influenced outcomes, particularly since only patients who received both neoadjuvant therapy and surgery at our institution were included. Furthermore, excluding patients who received part of their treatment at external institutions may have introduced selection bias, as their outcomes could have differed from those treated entirely at our center. Because this was not a randomized controlled trial, causal relationships between completion of neoadjuvant therapy and surgical outcomes cannot be definitively established. Despite these limitations, the study provides valuable real-world insights into the feasibility and postoperative outcomes of neoadjuvant treatment in resectable gastric cancer.

Conclusion

This study demonstrates that a considerable proportion of patients with clinically resectable, non-obstructive, non-bleeding, locally advanced gastric cancer were unable to complete neoadjuvant therapy and subsequently experienced high postoperative morbidity and mortality. Based on these findings and aligned with East Asian treatment guidelines, our institutional approach has shifted to prioritize upfront surgery followed by adjuvant chemotherapy in patients with similar clinical profiles. Further randomized controlled trials incorporating real-world data and focusing specifically on this patient subgroup are warranted.

Ethics

Ethics Committee Approval: This study was approved by the Ethics Committee of Aydın Adnan Menderes University Faculty of Medicine (approval no: 04, date: 27.06.2025).

Informed Consent: Written informed consent was obtained from all participants prior to data collection.

Acknowledgments

The authors would like to thank the staff of the General Surgery and Medical Oncology departments for their contributions to patient care and data collection.

Footnotes

Authorship Contributions

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

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

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

References

1. Cunningham D, Allum WH, Stenning SP, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. N Engl J Med. 2006;355:11-20.
2. Kang YK, Yook JH, Park YK, et al. PRODIGY: a phase III study of neoadjuvant docetaxel, oxaliplatin, and S-1 plus surgery and adjuvant S-1 versus surgery and adjuvant S-1 for resectable advanced gastric cancer. J Clin Oncol. 2021;39:2903-13.
3. Park SH, Lim DH, Sohn TS, et al. A randomized phase III trial comparing adjuvant single-agent S1, S-1 with oxaliplatin, and postoperative chemoradiation with S-1 and oxaliplatin in patients with node-positive gastric cancer after D2 resection: the ARTIST 2 trial . 2021;32:368-74.
4. Boige V, Pignon JP, SaintAubert B, et al. Final results of a randomized trial comparing preoperative 5fluorouracil/cisplatin to surgery alone in resectable adenocarcinoma of the stomach and lower esophagus: FNLCC ACCORD07FFCD 9703 trial. J Clin Oncol. 2007;25(18 Suppl):4510-2.
5. Eom SS, Ryu KW, Han HS, Kong SH. A Comprehensive and comparative review of global gastric cancer treatment guidelines: 2024 update. J Gastric Cancer. 2025;25:153-76.
6. Ajani JA, D'Amico TA, Bentrem DJ, et al. Gastric cancer, version 2.2022, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2022;20:167-92.
7. Lordick F, Carneiro F, Cascinu S, et al. Gastric cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. Ann Oncol. 2022;33:1005-20.
8. Wang FH, Zhang XT, Tang L, et al. The Chinese Society of Clinical Oncology (CSCO): clinical guidelines for the diagnosis and treatment of gastric cancer, 2023. Cancer Commun (Lond). 2024;44:127-72.
9. Kim IH, Kang SJ, Choi W, et al. Korean Practice Guidelines for Gastric Cancer 2024: an evidence-based, multidisciplinary approach (update of 2022 guideline). J Gastric Cancer. 2025;25:5-114.
10. Japanese Gastric Cancer Association. Japanese Gastric Cancer Treatment Guidelines 2021 (6th edition). Gastric Cancer. 2023;26:1-25.
11. Shinkai M, Imano M, Yokokawa M, et al. Phase I/II study of neoadjuvant chemoradiotherapy consisting of S-1 and cisplatin for patients with clinically resectable type 4 or large type 3 gastric cancer (OGSG1205). Ann Surg Oncol. 2025;32:2651-61.
12. Wang J, Tong T, Zhang G, et al. Evaluation of neoadjuvant immunotherapy in resectable gastric/gastroesophageal junction tumors: a meta-analysis and systematic review. Front Immunol. 2024;15:1339757.
13. Aydoğan O, Coşkun MÇ, Şekerci UU, Cartı EB. Can proximal gastrectomy be an alternative to total gastrectomy due to its nutritional advantage? A retrospective cohort study. Ann Surg Treat Res. 2025;108:79-85.
14. Fong C, Johnston E, Starling N. Neoadjuvant and adjuvant therapy approaches to gastric cancer. Curr Treat Options Oncol. 2022;23:1247-68.
15. Hui C, Ewongwo A, Lau B, et al. Patterns of recurrence after poor response to neoadjuvant chemotherapy in gastric cancer and the role for adjuvant radiation. Ann Surg Oncol. 2024;31:413-20.
16. Dağıstanlı S. Factors affecting the morbidity and mortality of gastric cancer surgery. Anatol J Gen Med Res. 2022;32:257-61.
17. Claassen YHM, Hartgrink HH, Dikken JL, et al. Surgical morbidity and mortality after neoadjuvant chemotherapy in the CRITICS gastric cancer trial. Eur J Surg Oncol. 2018;44:613-9.
18. Ling Q, Huang ST, Yu TH, et al. Optimal timing of surgery for gastric cancer after neoadjuvant chemotherapy: a systematic review and meta-analysis. World J Surg Oncol. 2023;21:377.
19. Siegel JB, Mukherjee R, DeChamplain B, Sutton JM, Mahvi DM, Lancaster WP. Neoadjuvant chemotherapy is associated with decreased survival in early-stage gastric cancer. Am Surg. 2024;90:28-37.
20. Abboretti F, Lambert C, Schäfer M, et al. Neoadjuvant chemotherapy does not improve survival in cT2N0M0 gastric adenocarcinoma patients: a multicenter propensity score analysis. Ann Surg Oncol. 2024;31:5273-82.
21. Rausei S, Bali CD, Lianos GD. Neoadjuvant chemotherapy for gastric cancer: has the time to decelerate the enthusiasm passed us by? Semin Oncol. 2020;47:355-60.
22. Reddavid R, Sofia S, Chiaro P, et al. Neoadjuvant chemotherapy for gastric cancer: is it a must or a fake? World J Gastroenterol. 2018;24:274-89.

DOI: 10.4274/anatoljmed.2025.55265
Anatol J Gen Med Res 2025;35(3):288-293

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Electromyographic Findings in Patients Referred to İzmir Tepecik Education and Research Hospital as a Remote Center Following the February 6, 2023 Earthquake

6 Şubat 2023 Depremi Sonrası Uzak Bir Merkez Olarak İzmir Tepecik Eğitim ve Araştırma Hastanesine Sevkedilen Hastalarda Elektromiyografi Deneyimi

Aysel Çoban Taşkın[1], Ebru Bölük[2]

[1]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of Neurology, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir City Hospital, Department of Neurology, İzmir, Türkiye

Cite as: Çoban Taşkın A, Bölük E. Electromyographic findings in patients referred to İzmir Tepecik Education and Research Hospital as a remote center following the February 6, 2023 earthquake. Anatol J Gen Med Res. 2025;35(3):288-293

Abstract

Objective: On February 6, 2023, two powerful earthquakes struck our country, resulting in at least 50.000 fatalities and over 122.000 injuries, according to official reports. This study examines the demographic data, clinical features, and electromyography (EMG) findings of children and adults referred to the EMG laboratory at University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital for evaluation of suspected peripheral nerve damage sustained during the earthquake.

Methods: We analyzed the demographic, clinical, and EMG findings of patients who were referred to the EMG laboratory between February and July 2023 with a provisional diagnosis of peripheral nerve injury.

Results: A total of 45 patients (aged 3 to 62 years; mean age 22.1 years) were evaluated in the EMG laboratory between February and July 2023. Three of them sustained injuries after jumping from a height during the earthquake. The remaining 42 patients were rescued from beneath collapsed structures, with entrapment durations ranging from 2 to 105 h (mean: 22 h). The initial EMG examination was performed, on average, 32 days after the injury. At least two follow-up EMG examinations were conducted in six patients. No statistically significant difference was noted between the symptomatic and the injured sides. Electrophysiological evidence of peripheral nerve or plexus injuries was identified in 89% of patients, whereas 11% had normal EMG findings. Simultaneous involvement of multiple peripheral nerves or plexuses was detected in 22% of patients (n=10). A positive correlation was found between entrapment duration and the occurrence of crush/compartment syndrome.

Conclusion: EMG is a crucial diagnostic tool for evaluating traumatic peripheral nerve injuries. The EMG and laboratory findings reported in this study highlight the challenges faced by patients who survived two devastating earthquakes and were referred to our remote center for assessment of suspected traumatic nerve injuries.

Keywords: Electromyography, peripheral nerve damage, earthquake, crush syndrome

Address for Correspondence/Yazışma Adresi: Aysel Çoban Taşkın, MD, University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of Neurology, İzmir, Türkiye
E-mail: ayselcoban@hotmail.com
ORCID ID: orcid.org/0000-0002-2976-8545

This study was presented as an oral presentation at the 6th International Medical Congress of IMCIDU 2024 on 5-6 December 2024, Seferihisar, İzmir

Received/Geliş tarihi: 03.06.2025
Accepted/Kabul tarihi: 09.10.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

Amaç: 6 Şubat 2023'te ülkemizde iki büyük deprem yaşandı. Depremler sonucunda resmi rakamlara göre en az 50,000 kişi hayatını kaybetti ve 122,000'den fazla kişi ise yaralandı. Bu çalışma ile 6 Şubat depremi sonrası Sağlık Bilimleri Üniversitesi, İzmir Tepecik Eğitim ve Araştırma Hastanesi elektromiyografi (EMG) laboratuvarına, periferik sinir hasarlanması ön tanısı ile gönderilen, çocuk ve erişkin hastalara ait demografik veri, klinik ve muayene özellikleri ile EMG bulgularını derledik.

Yöntem: Şubat-Temmuz 2023 tarihleri arasında periferik sinir hasarı ön tanısıyla EMG laboratuvarına yönlendirilen çocuk ve erişkin hastalara ait demografik veri, klinik ve muayene özellikleri ile EMG bulguları incelendi.

Bulgular: Şubat-Temmuz 2023 tarihlerinde EMG laboratuvarına yaşları 3 ile 62 arasında değişen, yaş ortalaması 22,1 olan 45 hasta başvurdu. Hastaların 3'ü deprem sırasında kaçmaya çalışırken yüksekten atlama sonucu yaralanmıştı. Diğer hastalarda enkaz altında kalma süresi 2 ile 105 saat arasında ve ortalama 22 saatti. Laboratuvarımızdaki ilk EMG incelemesi yaralanmadan sonra ortalama ilk 32,13 günde yapılmıştı. Altı hastaya en az 2 takip EMG incelemesi yapılabilmişti. Hastaların yakınmaları olan ve travmaya uğradıkları taraflar arasında belirgin farklılık gözlenmedi, %89'unda bir periferik sinir ya da pleksus hasarını düşündüren elektrofizyolojik bulgular saptanırken, %11'inde inceleme normal sınırlardaydı. Aynı anda birden fazla periferik sinir ya da pleksus hasarı oranı ise %22'ydi (10 hasta). Enkaz altında kalma süresi ile crush/kompartman sendromu oluşumu arasında anlamlı bir pozitif korelasyon bulundu.

Sonuç: EMG incelemesi, travmatik periferik sinir yaralanmaları sonrasında sıklıkla başvurulan tanı yöntemlerindendir. Ülkemizin 11 ilini etkileyen iki büyük deprem sonrası kurtulan hastalardan travmatik periferik sinir yaralanması öntanısıyla bölgeye uzak bir merkez olan kliniğimize yönlendirilen hastalardan elde edilen EMG veri ve laboratuvar deneyimleri paylaşılmıştır.

Anahtar Kelimeler: Elektromiyografi, periferik sinir hasarı, deprem, crush sendrom

Introduction

On February 6, 2023, two powerful earthquakes struck our country, resulting in at least 50.000 fatalities and over 122.000 injuries, according to official reports[1]. Earthquakes are catastrophic natural disasters that result not only in extensive structural damage but also in significant physical and psychological trauma. Musculoskeletal injuries are the most frequently encountered conditions following an earthquake. These injuries include fractures, dislocations, crush injuries, compartment syndrome, and amputations[2]. Traumatic peripheral nerve injuries may also occur following an earthquake. Such injuries are frequently reported and occur secondary to mechanical trauma (e.g., musculoskeletal injuries) or to infection. The type and severity of nerve damage play a key role in determining prognosis and guiding treatment. Electrophysiological methods are critical for early diagnosis and follow-up[3]. Acute stress symptoms, such as anxiety and depression, are commonly reported among earthquake survivors. In the same way, children may develop psychiatric symptoms such as anxiety and depression[4]. Furthermore, short sleep duration, sleep insufficiency, poor sleep quality, and insomnia symptoms have been reported. These disturbances are often associated with financial hardship and disruptions to healthcare services[5]. This study examines the demographic data, clinical features, and electromyography (EMG) results of children and adults referred to the EMG laboratory at University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital with suspected peripheral nerve damage following the earthquake.

Materials and Methods

This retrospective study included 45 patients who were referred to the EMG laboratory between February and July 2023 with a provisional diagnosis of peripheral nerve injury. Eligible participants were those who sustained injuries during the earthquake and exhibited clinical signs or symptoms suggestive of at least one peripheral nerve injury. Patients were excluded if they had a known history or clinical evidence of pre-existing peripheral nerve injury, declined to undergo EMG, or were unable to undergo the procedure due to contraindications such as a cardiac pacemaker, limb amputation, or open wounds at the stimulation or recording sites.

Hospital records were retrospectively reviewed, and the following variables were recorded for each patient: age, sex, time spent under the rubble (in h), time interval from injury to the first EMG assessment (in days), affected extremity (upper or lower; right or left limb), specific injured peripheral nerves, and the presence of compartment syndrome or crush syndrome. Patients were stratified into two groups according to the presence or absence of compartment syndrome, and the time spent under the rubble was compared between groups.

Electrophysiological studies were performed using a Nihon Kohden MEB-9200K (2007) device. Motor and sensory nerve conduction studies were conducted using standardized protocols, with filter settings between 2 Hz and 20 kHz, stimulus durations of 0.05-0.1 ms, and a sweep speed

of 2 ms/division. Peripheral nerve injury was defined electrophysiologically by any of the following: reduced compound muscle action potential and/or sensory nerve action potential amplitudes; prolonged distal latency; decreased conduction velocity; conduction block; needle EMG evidence of denervation (fibrillation potentials, positive sharp waves); and neurogenic motor unit potential changes, including increased amplitude, prolonged duration, polyphasic morphology, and reduced recruitment.

All EMG evaluations were conducted by two clinical neurophysiology specialists simultaneously to ensure diagnostic accuracy. The primarily affected extremity was systematically assessed and, when the patient cooperated, the contralateral (unaffected) extremity was also examined for comparison.

This study was approved by the Non-Interventional Research Ethics Committee of the University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital (approval no: 2025/03-06, date: 10.04.2025). Written informed consent was obtained from all patients or their legal representatives before EMG testing. Consent also included permission for the use of anonymized data for research purposes.

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, New York, USA). Depending on the distribution of the data, descriptive statistics were reported as mean ± standard deviation, median (minimum-maximum), or frequency and percentage, as appropriate. Patients were divided into two groups based on the presence or absence of compartment syndrome. Differences in time spent under the rubble between these groups were assessed using the Mann-Whitney U test. Spearman correlation analysis was used to evaluate the relationship between time spent under the rubble and the likelihood of developing compartment syndrome. In addition, logistic regression analysis was conducted to assess the predictive effect of time spent under the rubble on the risk of developing compartment syndrome. A p-value of <0.05 was considered statistically significant.

Results

A total of 45 patients aged 3-62 years (mean age: 22.1 years) were included in the study. These patients were evaluated in the EMG laboratory between February 6, 2023 (the date of the earthquake) and July 2023. Among them, 25 (55.6%) were children (<18 years of age) and 20 (44.4%) were adults (≥18 years of age). Three patients sustained injuries after jumping from a height during the earthquake. The remaining 42 patients were rescued from beneath collapsed structures, with time spent under the rubble ranging from 2 to 105 h (mean: 22 h).

The initial EMG examination was performed in our laboratory on average 32 days after the injury. At least two follow-up EMG exams were conducted on six patients, all of whom were under the age of 18. During these evaluations, nerve regeneration was observed to varying degrees and at different rates. No statistically significant difference was noted between the symptomatic and injured sides. However, lower extremity involvement was more frequent than upper extremity involvement (Table 1).

Electrophysiological evidence of peripheral nerve or plexus injury was identified in 89% of patients, while 11% had normal EMG findings. Simultaneous involvement of multiple peripheral nerves or plexuses was present in 22% of patients (n=10). The distribution of affected nerves, ranging from mild or partial involvement to complete axonal loss, is shown in Figure 1. The peroneal and sciatic nerves were most frequently affected.

Crush syndrome, compartment syndrome, and/or fasciotomy were observed in 16 of the 42 patients who had been trapped under the rubble. The mean time spent under the rubble

Table 1. Demographic and clinical characteristics of earthquake survivors (n=45)

Characteristics		Number of patients
Sex	Male	21 (46.7%)
	Female	24 (53.3%)
	Total	45 (100.0)
Age (year)	Male	22±17.8
	Female	22.2±17.3
	Total	22.1±17.3
Time under the rubble (h)		21.9±25.2
Time of the first EMG test (days)		32.1±32.7
Affected extremity	Right	21 (47%)
	Left	19 (42%)
	Bilateral	5 (11%)
	Upper extremity	12 (27%)
	Lower extremity	27 (60%)
	Upper and lower extremities	6 (13%)

EMG: Electromyography

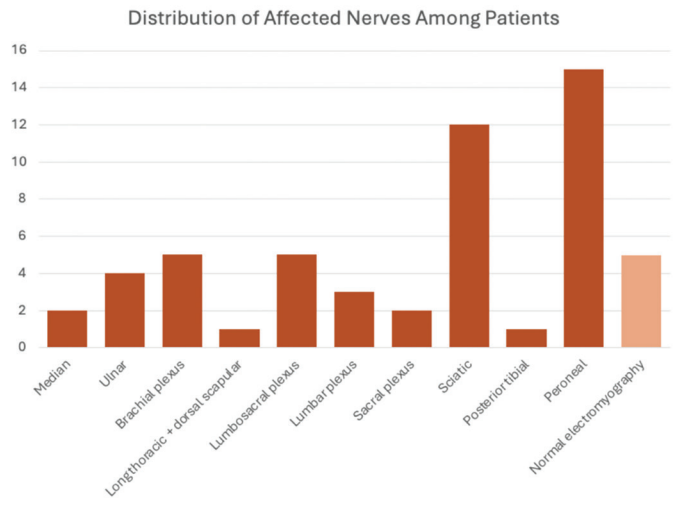


Figure 1. Distribution of affected nerves among patients. The peroneal and sciatic nerves were the most frequently affected; electromyographic findings were normal in five patients

was significantly longer in the group that developed crush or compartment syndrome (p=0.002).

There was a positive correlation between time spent under the rubble and occurrence of crush/compartment syndrome (r=0.480, p=0.001). Logistic regression indicated that the odds of developing this syndrome increased by 4% per additional hour under the rubble (odds ratio =1.040; 95% confidence interval: 1.007-1.075; p=0.018).

Discussion

Earthquakes are catastrophic natural events that often result in mortality and injuries. Peripheral nerve injuries, frequently resulting from prolonged compression during an earthquake, are a major cause of morbidity. These injuries can result from crush injuries, ischemia, bone fractures, and compartment syndrome; crush injuries are the most common type of peripheral nerve damage. Two main mechanisms are responsible for these injuries: mechanical crushing and ischemia. Short-term ischemic changes are generally reversible; however, with increasing duration and severity of compression, nerve damage worsens, recovery is delayed, and sustained compression may lead to irreversible damage[6].

This study analyzed data from 45 earthquake survivors who were evaluated in the EMG laboratory of our neurology clinic. Sex and age distributions in disaster-related injuries can vary. Although several studies suggest that females may be more frequently injured[7-10], a study investigating peripheral nerve injuries following the 1999 Marmara earthquake reported no significant sex difference (females, 52.5%; males, 47.5%) [11]. Similarly, we found no significant sex-related differences in the incidence of peripheral nerve injury (females, 53.3%; males, 46.7%), suggesting that sex does not appear to be a major risk factor in this context.

Extremity injuries are common following earthquakes, with the lower extremities more frequently affected than the upper extremities[12-14]. The Marmara earthquake study reported lower limb injuries as the most common, accounting for 47.5%, with no significant difference between the right and left sides[11]. Our finding that lower extremity injuries were more frequent (60%) is consistent with previous reports, such as Özdemir et al.[15] who reported 59.07%. This similarity supports the reproducibility of injury patterns in large-scale earthquakes.

Although our sample size was relatively small, this likely reflects the remote location of our center, which limited referrals despite a high number of injured individuals in the region. Nevertheless, our findings regarding the distribution of nerve injuries are consistent with previous studies from more centrally located tertiary hospitals[11,15,16]. This suggests that the injury pattern we observed is likely representative and not due to referral bias.

Electrophysiological examinations play a crucial role in identifying the distribution and extent of nerve damage. Follow-up assessments allow tracking of recovery and prognostication. In this study, the average time for the initial EMG was 32 days. Based on the electrophysiological findings, the sciatic nerve (24%) and the peroneal nerve (30%) were the most frequently affected. Previous studies of peripheral nerve injuries following earthquakes have also reported the sciatic nerve as the most frequently injured[16-18].

Two-year follow-up of patients with peripheral nerve injuries has demonstrated regeneration in 60%-70% of cases, indicating a favorable prognosis[11,19]. Although follow-up data were available for only six pediatric patients, these cases showed varying degrees of improvement and provided preliminary insight into prognosis.

The term crush denotes severe compression or crushing, whereas crush syndrome refers to the systemic consequences of such injuries. Crush syndrome was first described by Bywaters and Beall[20] in 1941. It results from muscle damage sustained during major earthquakes,

other disasters, or intense physical exertion. It includes medical and surgical complications such as acute renal failure, electrolyte imbalance, hemodynamic instability, respiratory failure, infection, compartment syndrome, and bleeding[20]. It is the second leading cause of disaster-related deaths, following fatalities from direct trauma. Compartment syndrome occurs when muscle edema raises intracompartmental pressure, impairing capillary perfusion. This increased pressure leads to ischemic damage to nerve and muscle cells and exacerbates rhabdomyolysis[21]. In this study, compartment syndrome was observed in 38% of earthquake survivors. Patients who developed compartment syndrome spent more time under the rubble than those who did not. Previous studies have reported a positive correlation between the time trapped under rubble and the severity of nerve damage[22,23], a finding that was also confirmed in our study.

Study Limitations

This study has several limitations. First, because our hospital is a referral center in a remote area, only a limited number of patients were included; this may restrict the generalizability of the findings. Second, not all patients can undergo follow-up EMG tests, limiting our ability to fully assess long-term recovery. Third, because the data were collected retrospectively from hospital records, some data may have been incomplete or missing, which could have affected certain analyses. Finally, psychological factors, such as posttraumatic stress or depression, which can significantly impact recovery, were not systematically evaluated in this study.

Conclusion

Earthquakes are unpredictable and devastating natural disasters that result in extensive human suffering and long-term health consequences. Among the most frequent and impactful outcomes in survivors are peripheral nerve injuries, which can lead to lasting neurological deficits and functional impairments, often necessitating prolonged rehabilitation. The risk and severity of these injuries increase proportionally with the time spent under the rubble, particularly when associated with compartment and crush syndromes, both of which are time-sensitive, life-threatening complications. Thus, timely and well-coordinated search and rescue operations are vital, not only for saving lives but also for minimizing long-term neurological damage.

Although peripheral nerve injuries are often considered a secondary concern in the post-earthquake period, they may in fact be a leading cause of long-term morbidity. Therefore, neurological evaluations should not be neglected during patient follow-up, and EMG should be used routinely for diagnostic and prognostic assessment. In our study, certain nerves were more frequently affected, suggesting that clinical evaluations in the field should prioritize these nerves. Furthermore, in these patients, physical trauma and psychological effects of prolonged time spent under the rubble may significantly influence rehabilitation and recovery processes. Finally, the diagnostic, therapeutic, and long-term rehabilitation needs associated with peripheral nerve injuries create a lasting and multifaceted burden on post-disaster healthcare systems. For these reasons, developing preventive and supportive strategies at both the individual and systemic levels is essential.

Ethics

Ethics Committee Approval: This study was approved by the Non-Interventional Research Ethics Committee of the University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital (approval no: 2025/03-06, date: 10.04.2025).

Informed Consent: Written informed consent was obtained from all patients or their legal representatives before EMG testing. Consent also included permission for the use of anonymized data for research purposes.

Footnotes

Authorship Contributions

Surgical and Medical Practises: A.Ç.T., E.B., Concept: A.Ç.T., E.B., Design: A.Ç.T., E.B., Data Collection or Processing: A.Ç.T., E.B., Analysis or Interpretation: A.Ç.T., Literature Search: A.Ç.T., E.B., Writing: A.Ç.T.

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

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

References

1. World Health Organization. Türkiye earthquake: external situation report no. 5, 13-19 March 2023. Copenhagen: WHO Regional Office for Europe; 2023. Available from: https://iris.who.int/bitstream/handle/10665/366587/WHO-EURO-2023-7145-46911-68823-eng.pdf?sequence=1
2. Uluöz M, Gökmen MY. The 2023 Turkey earthquake: management of 627 pediatric musculoskeletal injuries in the first month. Children (Basel). 2023;10:1733.

3. Uzunçakmak Uyanık H, Yıldız Sarıkaya FG. Periferik sinir yaralanmalarında elektrofizyolojik değerlendirme. İçinde: Gürler T, editör. Periferik sinir cerrahisinde güncel bilgiler. 1. baskı. Ankara: Türkiye klinikleri; 2022:9-13.
4. Özyurt G, Totur G, Üzüm Ö, et al. Psychiatric characteristics of child survivors after the Kahramanmaraş double earthquake. Pediatr Int. 2025;67:e15830.
5. Li X, Buxton OM, Hikichi H, et al. Predictors of persistent sleep problems among older disaster survivors: a natural experiment from the 2011 Great East Japan earthquake and tsunami. Sleep. 2018;41:zsy084.
6. Valone F 3rd, Lyon R, Lieberman J, Burch S. Efficacy of transcranial motor evoked potentials, mechanically elicited electromyography, and evoked electromyography to assess nerve root function during sustained compression in a porcine model. Spine (Phila Pa 1976). 2014;39:989-93.
7. Ellidokuz H, Ucku R, Aydin UY, Ellidokuz E. Risk factors for death and injuries in earthquake: cross-sectional study from Afyon, Turkey. Croat Med J. 2005;46:613-8.
8. Roy N, Shah H, Patel V, Coughlin RR. The Gujarat earthquake (2001) experience in a seismically unprepared area: community hospital medical response. Prehosp Disaster Med. 2002;17:186-95.
9. Tanida N. What happened to elderly people in the great Hanshin earthquake. BMJ.1996;313:1133-5.
10. Del Papa J, Vittorini P, D'Aloisio F, et al. Retrospective analysis of injuries and hospitalizations of patients following the 2009 earthquake of L'Aquila City. Int J Environ Res Public Health. 2019;16:1675.
11. Uzun N, Savrun Karaali F, Yazıcı S, Caksıbaeva C, Kızıltan ME. Experience obtained in an electromyography laboratory after the August 17, 1999 earthquake. Cerrahpaşa J Med. 2001;32:169-74.
12. Tang B, Chen Q, Chen X, et al. Earthquake-related injuries among survivors: a systematic review and quantitative synthesis of the literature. Int J Disaster Risk Reduct. 2017;21:159-67.
13. Peek-Asa C, Kraus JF, Bourque LB, Vimalachandra D, Yu J, Abrams J. Fatal and hospitalized injuries resulting from the 1994 Northridge earthquake. Int J Epidemiol. 1998;27:459-65.
14. Roces MC, White ME, Dayrit MM, Durkin ME. Risk factors for injuries due to the 1990 earthquake in Luzon, Philippines. Bull World Health Organ. 1992;70:509-14.
15. Özdemir G, Karlıdağ T, Bingöl O, et al. Systematic triage and treatment of earthquake victims: our experience in a tertiary hospital after the 2023 Kahramanmaras earthquake. Jt Dis Relat Surg. 2023;34:480-7.
16. Ahrari MN, Zangiabadi N, Asadi A, Sarafi Nejad A. Prevalence and distribution of peripheral nerve injuries in victims of Bam earthquake. Electromyogr Clin Neurophysiol. 2006;46:59-62.
17. Castillo-Galván ML, Martínez-Ruiz FM, de la Garza-Castro O, Elizondo-Omaña RE, Guzmán-López S. Estudio de la lesión nerviosa periférica en pacientes atendidos por traumatismos [Study of peripheral nerve injury in trauma patients]. Gac Med Mex. 2014;150:527-32.
18. Guner S, Guner SI, Isik Y, et al. Review of Van earthquakes form an orthopaedic perspective: a multicentre retrospective study. Int Orthop. 2013;37:119-24.
19. Yoshida T, Tada K, Uemura K, Yonenobu K. Peripheral nerve palsies in victims of the Hanshin-Awaji earthquake. Clin Orthop Relat Res. 1999;362:208-17.
20. Bywaters EG, Beall D. Crush injuries with impairment of renal function. Br Med J. 1941;1:427-32.
21. Better OS, Rubinstein I, Reis DN. Muscle crush compartment syndrome: fulminant local edema with threatening systemic effects. Kidney Int. 2003;63:1155-7.
22. Tahmasebi MN, Kiani K, Mazlouman SJ, et al. Musculoskeletal injuries associated with earthquake. A report of injuries of Iran's December 26, 2003 Bam earthquake casualties managed in tertiary referral centers. Injury. 2005;36:27-32.
23. Gesoğlu Demir T, Ethemoglu KB. Clinical, etiological, and electrophysiological characteristics of patients with peripheral nerve damage caused by the February 6 earthquake in Türkiye. Turk J Neurol. 2024;30:141-8.

DOI: 10.4274/anatoljmed.2025.28291
Anatol J Gen Med Res 2025;35(3):294-300

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Comparison of Clinicopathological Features in Gastric Cancer Patients with and without Bleeding at Initial Presentation

İlk Başvuru Anında Kanama Bulgusu Olan ve Olmayan Mide Kanseri Hastalarında Klinikopatolojik Özelliklerin Karşılaştırılması

Seval Akay[1], **Kamil Pehlivanoğlu**[2]

[1]İzmir City Hospital, Clinic of Medical Oncology, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of General Surgery, İzmir, Türkiye

Cite as: Akay S, Pehlivanoğlu K. Comparison of clinicopathological features in gastric cancer patients with and without bleeding at initial presentation. Anatol J Gen Med Res. 2025;35(3):294-300

Abstract

Objective: We aimed to investigate whether the presence of bleeding complaints at initial presentation affected laboratory and histopathological features in patients diagnosed with gastric cancer, dividing the patients into two groups: those with and those without such complaints.

Methods: Data from 148 patients diagnosed with gastric cancer were retrospectively analyzed using the hospital's computerized medical record system. Patients were grouped based on the presence or absence of hematemesis or endoscopic signs of bleeding, and compared with respect to age, sex, tumor location, disease stage, hemogram parameters, neutrophil-to-lymphocyte ratio (NLR), mean platelet volume (MPV)-to-platelet ratio (PLT), and blood group characteristics. Tumors located in the fundus or cardia were classified as proximal gastric tumors. Non-parametric variables were analyzed using the Mann-Whitney U test, parametric variables were analyzed using the independent samples t-test, and categorical variables were analyzed using the chi-square test. A p-value of ≤0.05 was considered statistically significant.

Results: Among the 148 patients included in the study (median age: 61.09 years; range: 29-87 years), 67 presented with bleeding symptoms, while 81 did not. There was no statistically significant difference in age or sex distribution between the two groups. Tumor localization, histological subtype, disease stage, and blood group characteristics were also similar between groups. Although the hemoglobin levels were lower in the bleeding group, this difference was not statistically significant (p=0.352). However, neutrophil counts were significantly higher (p=0.020) and lymphocyte counts significantly lower (p=0.007) in patients with bleeding. No significant differences were observed in platelet count, MPV, NLR, or MPV/PLT. The presence of bleeding was not associated with the need for neoadjuvant chemotherapy (p=0.155).

Conclusion: The presence of bleeding at initial presentation in patients with gastric cancer was not associated with differences in demographic features, tumor localization, histopathological subtype, or clinical stage. However, altered neutrophil and lymphocyte counts suggest that the bleeding may indicate an acute inflammatory response. These findings highlight the potential role of hematologic parameters as adjunctive markers in the initial assessment of gastric cancer patients presenting with gastrointestinal bleeding.

Keywords: Gastric cancer, upper gastrointestinal bleeding, hematologic parameters, tumor localization, neutrophil-to-lymphocyte ratio (NLR)

Address for Correspondence/Yazışma Adresi: Seval Akay, MD, İzmir City Hospital, Clinic of Medical Oncology, İzmir, Türkiye
E-mail: drsevalakay@hotmail.com
ORCID ID: orcid.org/0000-0002-1235-6739

Received/Geliş tarihi: 11.09.2025
Accepted/Kabul tarihi: 27.10.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

Amaç: Bu çalışmada, mide kanseri tanısı alan hastalarda ilk başvuruda kanama semptomlarının varlığının laboratuvar ve histopatolojik özellikler üzerine etkisini araştırmayı amaçladık.

Yöntem: Mide kanseri tanısı alan 148 hasta retrospektif olarak değerlendirildi. Hastalar hematemez veya endoskopik kanama bulgularının varlığına göre kanamalı ve kanamasız olmak üzere iki gruba ayrıldı. Yaş, cinsiyet, tümör lokalizasyonu, hastalık evresi, hemogram parametreleri, nötrofil-lenfosit oranı (NLR), ortalama trombosit hacmi (MPV)-trombosit oranı (PLT) ve kan grubu özellikleri açısından karşılaştırmalar yapıldı. Fundus ve kardiya yerleşimli tümörler proksimal mide tümörü olarak sınıflandırıldı. Parametrik olmayan değişkenler Mann-Whitney U testi, parametrik değişkenler bağımsız örneklem t-testi ve kategorik değişkenler ki-kare testi ile analiz edildi; p≤0,05 anlamlı kabul edildi.

Bulgular: Çalışmaya alınan 148 hastanın (medyan yaş: 61,09; aralık: 29-87) 67'sinde kanama semptomu mevcutken 81'inde yoktu. Gruplar arasında yaş, cinsiyet, tümör lokalizasyonu, histolojik tip, evre ve kan grubu açısından anlamlı fark saptanmadı. Kanamalı grupta hemoglobin düzeyleri daha düşük olmakla birlikte bu fark anlamlı değildi (p=0,352). Buna karşın nötrofil sayısı belirgin yüksek (p=0,020), lenfosit sayısı ise anlamlı düşük (p=0,007) bulundu. Trombosit sayısı, MPV, NLR ve MPV/PLT oranlarında anlamlı fark görülmedi. Kanama varlığı neoadjuvan kemoterapi ihtiyacıyla ilişkili değildi (p=0,155).

Sonuç: Mide kanserli hastalarda ilk başvuruda kanama varlığı; demografik özellikler, tümör lokalizasyonu, histopatolojik alt tip veya klinik evre ile ilişkili bulunmadı. Ancak nötrofil ve lenfosit düzeylerindeki değişiklikler kanamanın akut inflamatuvar yanıtı yansıtabileceğini göstermektedir. Bu bulgular, gastrointestinal kanama ile başvuran mide kanserli hastalarda hematolojik parametrelerin başlangıç değerlendirmesinde yardımcı biyobelirteçler olabileceğini düşündürmektedir.

Anahtar Kelimeler: Mide kanseri, üst gastrointestinal sistem kanaması, hematolojik parametreler, tümör lokalizasyonu, nötrofil-lenfosit oranı (NLR)

Introduction

Gastric cancer ranks as the fifth most common malignancy worldwide[1]. Its clinical presentation is highly variable, and a significant proportion of patients are diagnosed at an advanced stage. While most patients present with non-acute constitutional symptoms such as weight loss, abdominal pain, and fatigue, gastric cancer may also manifest with acute complications including hematemesis, tumor or organ perforation, or gastric outlet obstruction-conditions that have been associated with poor overall survival[2,3]. In fact, the overall survival of patients presenting with such acute symptoms has been reported to be as short as six months[4]. However, improved diagnostic tools and advancements in treatment have contributed to better overall survival rates in recent years.

Bleeding due to malignancy accounts for approximately 3% of upper gastrointestinal (GI) bleeding cases[5]. Regardless of etiology, upper GI bleeding is associated with a mortality rate ranging from 2% to 10%[6]. Although bleeding is a significant clinical feature of gastric cancer, it is not always evident. Some patients present with overt symptoms such as hematemesis or melena, whereas others have minor bleeding confirmed only by microcytic anemia or show no signs of bleeding.

This study aims to compare the clinical and pathological characteristics of gastric cancer patients based on the presence or absence of bleeding at the time of initial presentation. The findings are expected to enhance our understanding of the diverse clinical manifestations of gastric cancer and contribute to the development of improved diagnostic and therapeutic strategies.

Materials and Methods

The study was approved by the Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital (approval no: 2025/05-23, date: 12.06.2025). Given the retrospective design, the requirement for informed consent was waived.

This was a retrospective, single-center cohort study conducted between 2011 and 2023. A total of 148 consecutive patients diagnosed with gastric cancer during this period were included. Patients with prior gastric surgery affecting bleeding risk, hematologic disorders that could alter laboratory parameters, or incomplete clinical data were excluded.

Patients were stratified into two groups based on their bleeding status at initial presentation. The bleeding group included patients presenting with hematemesis, melena, or endoscopic evidence of bleeding. The non-bleeding group consisted of patients who had no clinical or endoscopic signs of upper GI bleeding at diagnosis.

Data were extracted from the hospital's electronic medical records, including several key variables. Demographic data included age, sex, and blood group. Tumor characteristics included location (categorized as proximal-cardia/fundus/corpus-, or antrum), histopathology, and clinical stage

according to the American Joint Committee on cancer tumor node metastasis classification. Laboratory parameters at presentation were recorded, including hemoglobin, leukocyte count, neutrophil count, lymphocyte count, platelet count, mean platelet volume (MPV), neutrophil-to-lymphocyte ratio (NLR), and MPV-to-platelet ratio (PLT). Treatment data included whether patients received neoadjuvant chemotherapy.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics, version 20 (IBM Corp., Armonk, NY, USA). Normality was assessed with the Kolmogorov-Smirnov test. Normally distributed variables were reported as mean ± standard deviation and compared using the independent samples t-test. Non-normally distributed variables were presented as medians (interquartile ranges) and compared with the Mann-Whitney U test. Categorical variables were analyzed using the chi-square test or Fisher's exact test as appropriate. A p-value ≤0.05 was considered statistically significant.

Results

A total of 148 patients aged 29-87 years were included in the study; the median age was 61.09 years. Among the 67 patients who presented with GI bleeding (age range 37-84 years), the mean age was 61.18 years; among the 80 patients without bleeding (age range 29-87 years), the mean age was 61.01 years. According to the Kolmogorov-Smirnov test, the age distributions in both groups were normal. When patients were stratified using age thresholds of ≥50, ≥60, and ≥70 years, no statistically significant association was found between age and bleeding at presentation (p=0.767, p=0.305, and p=0.540, respectively).

Of the 148 patients, 64% were male (n=95) and 36% were female (n=53). There was no significant difference in sex between patients with bleeding at diagnosis and those without (p=0.732). Tumor histopathology was also similar between the two groups (Table 1).

Table 1. Comparison of hematological parameters between gastric cancer patients with and without gastrointestinal bleeding at initial presentation

	Median	Standard deviation	p-value
Age			
*absent	61.01	11.87	0.930
*present	61.18	10.94	
Hemoglobin			
*absent	12.10	2.22	0.352
*present	11.07	2.34	
Leucocyte			
*absent	7760	2568	0.061
*present	8830	4219	
Neutrophyl			
*absent	5010	2426	0.020
*present	6230	3825	
Lymphocyte			
*absent	1940	802	0.007
*present	1590	732	
Platelet			
*absent	266260	103518	0.607
*present	276400	135491	
MPV			
*absent	8.60	1.06	0.274
*present	8.50	1.20	

Table 1. Continued

	Median	Standard deviation	p-value
Neutrophyl/lymphocyte			
*absent	3.92	6.77	0.118
*present	6.93	15.55	
MPV/platelet			
*absent	0.04	0.03	0.324
*present	0.03	0.01	

*: Refers to bleeding, MPV: Mean platelet volume

Regarding blood types, 35.2% of patients were blood group 0, 47.7% group A, 8.6% group B, and 8.6% group AB. No significant association was found between blood group and the presence of bleeding (p=0.556). Additionally, no significant differences were observed according to the presence of A or B antigens or antibodies (p=0.947 and p=0.179, respectively).

Tumor localization was as follows: corpus (41.2%), antrum (35.8%), and proximal stomach (fundus and/or cardia) (23%). Tumor location did not differ significantly between the bleeding and non-bleeding groups (p=0.403) (Figure1).

The presence of metastasis at diagnosis did not differ significantly between groups (p=0.056).

When hematological parameters were compared, the median leukocyte count was significantly higher in patients presenting with bleeding (p=0.001), and the neutrophil count was also significantly elevated in this group (p=0.035). The median lymphocyte count tended to be lower in bleeding patients (p=0.007). However, there was no significant difference in hemoglobin (p=0.601), platelet count (p=0.607), MPV (p=0.274), NLR (p=0.118), or MPV-to-PLT (p=0.324) between the two groups.

Lymphocyte count was normally distributed in non-bleeding patients, with a median of 1.90 x10^9/L; in bleeding patients it was non-normally distributed, with a median of 1.50 x10^9/L (p=0.025) (Table 2).

Histopathological evaluation of surgical specimens revealed no significant difference in tumor type between groups (p=0.877) (Table 3).

Among patients who received neoadjuvant chemotherapy, 45.5% were in the bleeding group and 54.5% in the non-bleeding group; this difference was not statistically significant (p=0.155).

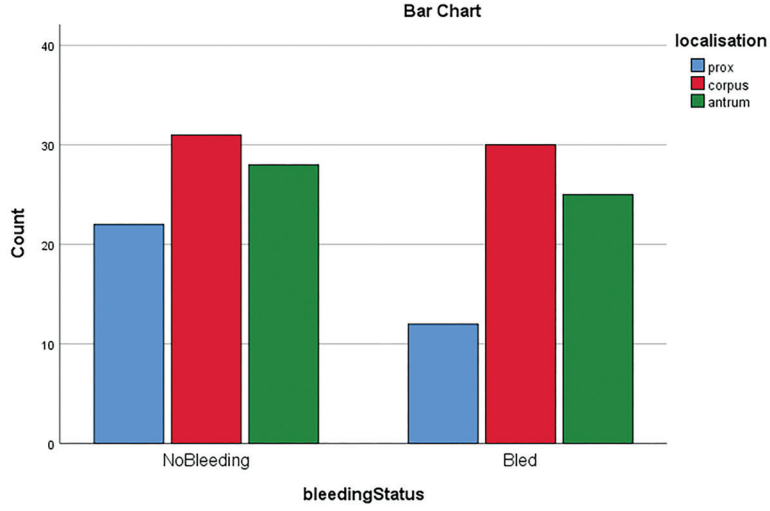


Figure 1. Distribution of tumor localization according to the presence or absence of gastrointestinal bleeding at initial presentation

Table 2. Comparison of clinicopathological and demographic characteristics between gastric cancer patients with and without gastrointestinal bleeding at initial presentation

Parameters	No bleeding	Bleeding	p-value
Sex			
Female	30	23	0.732
Male	51	44	
Location			
True gastric	63	57	0.259
Cardia	18	10	
Location			
Cardia	17	10	
Fundus	5	2	0.557
Corpus	31	30	
Antrum	28	25	
Blood group			
0	26	19	
A	34	27	0.566
B	4	7	
AB	5	6	
Group A			
Antigen-A	39	33	0.947
Antibody-A	30	26	
Group-B			
Antigen-B	9	13	0.179
Antibody-B	60	46	
Pathology			
Adenocarcinoma	69	54	
Neuroendocrine	7	8	0.877
GIST*	2	2	
Squamous cell	2	1	
Stage			
Stage 1	25	17	
Stage 2	14	15	0.826
Stage 3	19	16	
Stage 4	1	1	

GIST: Gastrointestinal stromal tumor

Table 3. Distribution of histopathological subtypes based on bleeding status at initial presentation in gastric cancer patients

Operation material	No bleeding (n)	Bleeding (n)	Total
Adenocarcinoma	69	54	123
Neuroendocrin carcinoma	7	8	15
GIST	2	2	4
Squamous cell carcinoma	2	1	3

GIST: Gastrointestinal stromal tumor

Discussion

In this retrospective study, clinical and laboratory parameters of patients diagnosed with gastric cancer were compared between those with and those without bleeding symptoms at initial presentation. Although hemoglobin levels were lower in patients presenting with bleeding, this difference did not reach statistical significance. This may be attributed to the acute or chronic nature of the bleeding, the hemodynamic stability of the patients, or the extent of blood loss. Additionally, the limited sample size may have contributed to the lack of statistical significance. While some studies have reported a higher incidence of GI bleeding among males and older patients, our findings did not confirm this association[7].

Inflammation is increasingly recognized as both an initiator and a promoter of cancer pathogenesis. Biological mediators released during the inflammatory process can cause structural damage to DNA, inhibit apoptosis, and facilitate tumor cell proliferation and metastasis. Consequently, inflammatory markers have been widely investigated in various types of cancer in recent years[8,9]. Among these, neutrophil, lymphocyte, and platelet counts, as well as ratios such as NLR and MPV-to-PLT, have been frequently evaluated.

In patients presenting to emergency departments with non-malignant upper GI bleeding, higher NLRs have been associated with increased mortality, whereas hemoglobin and hematocrit levels, leukocyte and platelet counts, and MPV were not significantly correlated with mortality[10]. However, no well-established thresholds or values have been defined specifically for bleeding attributable to gastric cancer. In our study, significantly elevated neutrophil counts and reduced lymphocyte levels in the bleeding group may reflect an acute inflammatory response to tumor-induced mucosal injury or suggest processes such as microbial translocation. Previous research has shown that lymphocyte counts-indicators of key players in the anti-tumor immune response-are often reduced in cancer patients[11].

Bleeding may result not only from tumor characteristics but also from acquired or iatrogenic causes. Major risk factors for GI bleeding include *Helicobacter pylori* infection, non-steroidal anti-inflammatory drug use, stress, and gastric acid secretion[12]. In this study, these factors were not assessed, and bleeding was analyzed only in relation to tumor histology and stage.

We found no significant association between bleeding and tumor localization, histological subtype, tumor stage, or demographic characteristics such as age and sex. This suggests that bleeding in gastric cancer may not depend solely on anatomical or histological features, but may involve more complex biological mechanisms. Previous studies have similarly reported no strong correlation between tumor location and bleeding in gastric cancer patients.

Study Limitations

This study has several limitations. Firstly, its retrospective, single-center design limits the generalizability of the findings. Furthermore, bleeding severity was not graded, and distinguishing acute from chronic bleeding was not possible, complicating the interpretation of the results. The standardization of laboratory tests and the timing of measurements, particularly the reliance on initial hemogram values at admission, may also have influenced the findings. Lastly, other genetic or iatrogenic factors that may predispose to bleeding were not assessed and could be explored in future research.

Conclusion

In this retrospective study, the presence of GI bleeding at initial presentation in patients with gastric cancer was not associated with significant differences in demographic characteristics, tumor location, histopathological subtype, or clinical stage. However, significantly elevated neutrophil counts and decreased lymphocyte levels in the bleeding group suggest an acute inflammatory response potentially linked to tumor-related mucosal injury. These findings underscore the potential utility of hematological parameters, particularly neutrophil and lymphocyte counts, as adjunctive markers in the early evaluation of gastric cancer patients presenting with bleeding. Further prospective studies are warranted to validate these associations and explore their prognostic implications.

Ethics

Ethics Committee Approval: The study was approved by the Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital (approval no: 2025/05-23, date: 12.06.2025).

Informed Consent: Given the retrospective design, the requirement for informed consent was waived.

Footnotes

Authorship Contributions

Surgical and Medical Practises: K.P., Concept: S.A., K.P., Design: S.A., K.P., Data Collection or Processing: S.A., Analysis or Interpretation: S.A., Literature Search: S.A., Writing: S.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. World Health Organization. International Agency for Research on Cancer: global cancer observatory. Available from: https://gco.iarc.who.int/media/globocan/factsheets/populations/900-world-fact-sheet.pdf
2. Kim HM, Kang D, Park JY, Cho YK, Choi MG, Park JM. Mortality risk scoring system in patients after bleeding from cancers in the upper gastrointestinal tract. Gut Liver. 2024;18:222-30.
3. Zhang J, Li K, Zhang Z, et al. Short-and long-term outcomes of one-stage versus two-stage gastrectomy for perforated gastric cancer: a multicenter retrospective propensity score-matched study. World J Surg Oncol. 2024;22:7.
4. Jeong SH, Park M, Seo KW, Min JS. Advances in surgical management of malignant gastric outlet obstruction. Cancers. 2025;17:2567.
5. Alali AA, Barkun AN. An update on the management of non-variceal upper gastrointestinal bleeding. Gastroenterol Rep (Oxf). 2023;11:goad011.
6. Antunes C, Tian C, Copelin II EL. Upper gastrointestinal bleeding. Treasure Island (FL): StatPearls Publishing; 2024.
7. van Leerdam ME, Vreeburg EM, Rauws EA, et al. Acute upper GI bleeding: did anything change? Time trend analysis of incidence and outcome of acute upper GI bleeding between 1993/1994 and 2000. Am J Gastroenterol. 2003;98:1494-9.
8. Çelik A, Armağan HH, Karaman K, et al. Retrospective review of hemogram parameters and vital findings in patients with non-variceal upper gastrointestinal tract bleeding. Med J SDU. 2024;31:45-51.
9. Dirican N, Anar C, Atalay Ş, et al. Effects on prognosis of hematologic parameters in patient with small cell lung cancer. Cukurova Med J. 2016;41:333-41.
10. Sargın M, Şahin M, Yalçın HR. Are hematological markers helpful in predicting the difference between endometrial intraepithelial neoplasia and grade 1 endometrioid endometrial carcinoma in the preoperative period? Jinekoloji Obstetrik Neonatol Tıp Dergisi. 2025;22:101-6.
11. Dertli R, Toka B, Asıl M, et al. Can neutrophil-lymphocyte ratio predict mortality in acute non-variceal upper gastrointestinal bleeding? Ulus Travma Acil Cerrahi Derg. 2022;28:626-33.
12. Dunn GP, Old LJ, Schreiber RD. The immunobiology of cancer immunosurveillance and immunoediting. Immunity. 2004;21:137-48.

DOI: 10.4274/anatoljmed.2025.70883
Anatol J Gen Med Res 2025;35(3):301-306

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Early Discharge in Breast Cancer Surgery: Safety and Patient Satisfaction with the Outpatient Approach

Meme Kanseri Cerrahisinde Erken Taburculuk: Günübirlik Cerrahi Yaklaşımının Güvenlik ve Hasta Memnuniyeti

Mahmut Onur Kültüroğlu, Ferit Aydın, Fatih Aslan, Mehmet Furkan Sağdıç, Bülent Aksel, Lütfi Doğan

University of Health Sciences Türkiye, Ankara Etlik City Hospital, Department of Surgical Oncology, Ankara, Türkiye

Cite as: Kültüroğlu MO, Aydın F, Aslan F, Sağdıç MF, Aksel B, Doğan L. Early discharge in breast cancer surgery: safety and patient satisfaction with the outpatient approach. Anatol J Gen Med Res. 2025;35(3):301-306

Abstract

Objective: Advances in minimally invasive surgery and anesthesia have enabled same-day discharge after breast cancer surgery. This study compared outcomes, complication rates, and patient satisfaction between patients discharged on the same-day of breast-conserving surgery with sentinel lymph node biopsy and patients discharged the following day.

Methods: Between December 2022 and May 2025, 349 patients with early-stage breast cancer at University of Health Sciences Türkiye, Ankara Etlik City Hospital were enrolled in the study. Same-day discharge was offered to patients who underwent general anesthesia with a laryngeal mask, had no drains or significant comorbidities, and demonstrated stable postoperative recovery. Patients accepting same-day discharge formed the outpatient group (n=215), and those who declined formed the inpatient group (n=134). Demographics, tumor characteristics, 30-day complications (Clavien-Dindo classification), and patient satisfaction (BREAST-Q) were recorded.

Results: The groups were comparable in terms of demographics and tumor features. The rates of seroma (6% vs. 6.7%), partial skin necrosis (1.4% vs. 1.5%), and surgical-site infection (1.9% vs. 1.5%) were similar (p>0.05). No major complications occurred. Overall, 90.5% of the patients experienced no complications, with minor complication rates comparable between the groups. BREAST-Q scores were slightly higher in the same-day discharge group (satisfaction: 94±4.6 vs. 92±6.3; psychosocial well-being: 91±3.8 vs. 89±4.5), although these differences were not statistically significant.

Conclusion: Outpatient breast-conserving surgery with sentinel lymph node biopsy is safe, feasible, and associated with high patient satisfaction in both groups, with no significant differences between the groups in carefully selected patients with early-stage disease who do not require axillary dissection or neoadjuvant therapy. Multicenter prospective studies are needed to evaluate long-term outcomes and optimize patient selection.

Keywords: Breast cancer, outpatient surgery, breast-conserving surgery

Öz

Amaç: Minimal invaziv cerrahi ve anestezideki gelişmeler, meme kanseri cerrahisi sonrası aynı gün taburculuğu mümkün kılmıştır. Bu çalışma, meme koruyucu cerrahi ile sentinel lenf nodu biyopsisi sonrası aynı gün taburcu edilen hastalar ile en az bir gün hastanede yatan hastalar arasında sonuçları, komplikasyon oranlarını ve hasta memnuniyetini karşılaştırmıştır.

Yöntem: Aralık 2022 ile Mayıs 2025 arasında Sağlık Bilimleri Üniversitesi, Ankara Etlik Şehir Hastanesi'nde erken evre meme kanseri tanısı alan 349 hasta çalışmaya dahil edildi. Aynı gün taburculuk; genel anestezi altında larengeal maske kullanılarak dren yerleştirilmeden, önemli ek hastalığı bulunmayan

Address for Correspondence/Yazışma Adresi: Mahmut Onur Kültüroğlu, MD, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Department of Surgical Oncology, Ankara, Türkiye
E-mail: mahmutonurkulturoglu@gmail.com
ORCID ID: orcid.org/0000-0003-2293-7139

Received/Geliş tarihi: 19.09.2025
Accepted/Kabul tarihi: 30.10.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

ve stabil postoperatif seyir gösteren hastalara önerildi. Aynı gün taburculuğu kabul edenler ayaktan grup (n=215), kabul etmeyenler ise yatan hasta grubu (n=134) olarak tanımlandı. Demografik veriler, tümör özellikleri, 30 günlük komplikasyonlar (Clavien-Dindo sınıflaması) ve hasta memnuniyeti (BREAST-Q) kaydedildi.

Bulgular: Gruplar demografik özellikler ve tümör parametreleri açısından benzerdi. Seroma (%6 vs. %6,7), kısmi cilt nekrozu (%1,4 vs. %1,5) ve cerrahi alan enfeksiyonu (%1,9 vs. %1,5) oranları benzerdi (p>0,05). Majör komplikasyon görülmedi. Genel olarak, hastaların %90,5'inde komplikasyon gelişmedi; minör komplikasyon oranları gruplar arasında karşılaştırılabilir düzeydeydi. BREAST-Q skorları ayaktan grupta biraz daha yüksek olsa da (memnuniyet: 94±4,6 vs. 92±6,3; psikososyal iyilik hali: 91±3,8 vs. 89±4,5), istatistiksel olarak anlamlı değildi.

Sonuç: Günübirlik meme koruyucu cerrahisi; aksiller diseksiyon veya neoadjuvan tedavi gerektirmeyen, erken evre hastalarda güvenli, uygulanabilir ve yüksek hasta memnuniyeti sağlayan bir yöntemdir. Uzun dönem sonuçların değerlendirilmesi ve hasta seçiminin optimize edilmesi için çok merkezli prospektif çalışmalara ihtiyaç vardır.

Anahtar Kelimeler: Meme kanseri, günübirlik cerrahi, meme koruyucu cerrahi

Introduction

Breast cancer treatment has become less invasive with advances in surgical and adjuvant therapies. Minimally invasive surgical approaches and advances in anesthetic methods have made the outpatient management of breast cancer feasible[1]. The concept of enhanced recovery, aiming to minimize pain and perioperative risk, has been adopted in various surgical fields, including breast cancer surgery. Studies have reported that outpatient surgery is associated with fewer recovery-related issues, such as postoperative pain, nausea, vomiting, and the need for psychosocial support[2].

Historically, outpatient surgery was limited to procedures performed under local or regional anesthesia, with minimal requirements for postoperative monitoring. However, it is currently possible to perform low-risk surgeries under general anesthesia in outpatient settings. In outpatient surgery, patients typically receive general anesthesia and are discharged within 12 hours of the procedure[1].

In this study, patients who underwent breast cancer surgery under general anesthesia and met specific postoperative criteria were discharged within 12 hours. Clinical outcomes were compared between those who were discharged on the same-day and those who remained hospitalized for at least one night. This study aimed to evaluate the feasibility and safety of outpatient surgery for breast cancer treatment.

Materials and Methods

Study Design and Patient Selection

This study was conducted at the Departmment of Surgical Oncology, University of Health Sciences Türkiye, Ankara Etlik City Hospital between December 2022 and May 2025 and included patients who underwent breast cancer surgery. Same-day discharge was recommended for patients who underwent breast-conserving surgery and sentinel lymph node biopsy (SLNB) in whom no drains were placed, who had no comorbidities, whose airways were managed with a laryngeal mask during surgery, and who remained hemodynamically stable postoperatively. Patients who accepted discharge within 12 hours comprised the outpatient group, while those who declined discharge for personal reasons (e.g., pain concerns or transportation difficulties) and remained hospitalized comprised the inpatient group. Demographic data [sex, age, body mass index (BMI), and smoking status, clinical and operative information (operation time and length of hospital stay), tumor characteristics (tumor size), and postoperative complications were recorded. In the same-day discharge group, patients were instructed to contact our clinic directly to communicate with the on-call physician if they experienced symptoms such as fatigue, general malaise, or deterioration in overall condition. Hospital beds were reserved for use in such situations. Our team instructed all patients and their caregivers in home wound care. For inpatients, wound care was performed by healthcare personnel during hospitalization and by patients and caregivers after discharge. In the outpatient group, all wound care was performed by the patients and caregivers. All patients were scheduled for follow-up between postoperative days 3 and 5, during which wound healing and vital signs were evaluated.

Patients who underwent mastectomy for breast cancer or axillary dissection, who received neoadjuvant therapy, who had drains placed during surgery, or who had comorbid conditions were excluded from the study. Complications that occurred within 30 days postoperatively were considered early complications. Postoperative bleeding requiring

transfusion and/or associated with hypotension was defined as "active bleeding". Wound edges that were completely separated, that required suturing, or that required prolonged dressing were defined as wound dehiscence These two were classified as major complications. Infections requiring antibiotics and/or drainage were defined as surgical-site infections (SSIs). Superficial skin necrosis or skin loss without wound dehiscence was categorized as "partial skin necrosis". Clinically stable seromas or hematomas controlled by percutaneous aspiration were considered "minor complications". The Clavien-Dindo classification was used to evaluate postoperative complications: patients without complications were classified as grade 0; complications managed with seroma aspiration or wound dressing were classified as grade 1; complications requiring antibiotics due to SSI were classified as grade 2; complications necessitating debridement under local anesthesia due to partial skin necrosis were classified as grade 3a; and patients with major complications were classified as grade 4. At the first postoperative follow-up visit (one month postoperatively), patient satisfaction and psychosocial well-being were assessed using the latest version of the BREAST-Q questionnaire. Ethical approval for this study was obtained from the Ethics Committee of University of Health Sciences Türkiye, Ankara Etlik City Hospital (approval number: AEŞH-BADEK2-2025-199, date: 08.07.2025). This study was conducted in accordance with the Declaration of Helsinki, with informed consent obtained from eligible participants.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows (version 25.0; IBM Corp., Armonk, NY, USA). Continuous variables (age, operation time, length of hospital stay, and tumor size) were reported as mean ± standard deviation and range (minimum-maximum). Categorical variables (smoking status and complications) were presented as frequencies (n) and percentages (%). The independent samples t-test was used to compare continuous variables between groups, while the chi-square test (χ^2) was used for categorical variables. The BREAST-Q satisfaction scores were compared between the outpatient and inpatient groups using the Mann-Whitney U test, as the scores did not follow a normal distribution. Statistical significance was set at $p<0.05$.

Results

Same-day discharge was offered to 349 patients who met the established criteria, of whom 215 accepted and were discharged. A total of 134 patients declined same-day discharge for personal reasons, such as concerns about pain or transportation, and spent their first postoperative day in the hospital. No patient received antibiotics at discharge, and all were prescribed oral analgesics (paracetamol 500 mg twice daily). All patients were female. Table 1 presents the age distribution, smoking status, BMI, operation time, and length of hospital stay by group.

The tumor sizes were similar between the outpatient and inpatient groups. The incidence of seroma, partial skin necrosis, SSI, and minor complications did not differ between the groups. No patient experienced major complications. None of the outpatients required unplanned readmission at night, and routine follow-up visits were completed as scheduled. During the first postoperative month, no patient required rehospitalization or reoperation because of complications (Table 2).

According to the Clavien-Dindo classification, 90.5% of patients experienced no postoperative complications. The most common complication was seroma; all cases were managed conservatively (grade 1). Six patients developed SSIs requiring antibiotic treatment (grade 2), and five patients required local debridement for partial skin necrosis

Table 1. Distribution of clinical characteristics and time-related parameters

Variable	Outpatient group (n=215)	Inpatient group (n=134)	p-value
Age, mean ± SD (minimum-maximum)	53.1±9.76 (26-70)	53.5±10.2 (28-72)	0.72
Smoking, n (%)	40 (18.6)	26 (19.4)	0.89
BMI, median (IQR)	28.2 (27.8-30.6)	28.6 (27.2-30.8)	0.14
Operation time (minutes), mean ± SD (minimum-maximum)	55.6±8.73 (45-75)	56.9±9.5 (45-75)	0.20
Length of hospital stay (hours), mean ± SD (minimum-maximum)	10.8 (9.2-12)	32.4 (18-41.5)	<0.001

SD: Standard deviation, BMI: Body mass index, IQR: Interquartile range

(grade 3). No major complications (grade 4) were observed. No statistically significant difference in the Clavien-Dindo distribution was found between the outpatient and inpatient groups (p>0.05; Table 3).

At the first month of postoperative follow-up, the BREAST-Q questionnaire, which assesses patient satisfaction and psychosocial well-being, showed higher scores in the outpatient group; however, the differences were not statistically significant (Tables 4 and 5).

Table 2. Distribution of histopathological data and complication

Variable	Outpatient group (n=215)	Inpatient group (n=134)	p-value
Tumor size (mm), mean ± SD (minimum-maximum)	16.6±6.3 (2-36)	17.2±5.8 (4-39)	0.36
Major complications, n	0	0	N/A
Seroma, n (%)	13 (6)	9 (6.7)	0.80
Partial skin necrosis, n (%)	3 (1.4)	2 (1.5)	0.95
Surgical-site infection, n (%)	4 (1.9)	2 (1.5)	0.80

SD: Standard deviation, N/A: Not applicable

Table 3. Postoperative complications according to the Clavien-Dindo classification

Clavien-Dindo grade	Outpatient group (n=215)	Inpatient group (n=134)	Total (n=349)
Grade 0	195 (90.7%)	121 (90.3%)	316 (90.5%)
Grade 1	13 (6.0%)	9 (6.7%)	22 (6.3%)
Grade 2	4 (1.9%)	2 (1.5%)	6 (1.7%)
Grade 3	3 (1.4%)	2 (1.5%)	5 (1.5%)
Grade 4	0	0	0

Table 4. BREAST-Q questionnaire results

Variable	Outpatient group (n=215)	Inpatient group (n=134)	p-value
BREAST-Q satisfaction, mean ± SD (minimum-maximum)	94±4.6 (70-100)	92±6.3 (65-100)	0.08
BREAST-Q psychosocial well-being, mean ± SD (minimum-maximum)	91±3.8 (68-100)	89±4.5 (65-100)	0.11

SD: Standard deviation

Table 5. Categorical distribution of BREAST-Q questionnaire results

Measurement/level	Outpatient group (n=215)	Inpatient group (n=134)	p-value
BREAST-Q satisfaction			
≥90 (high)	188 (87.4%)	108 (80.6%)	0.08
80-89 (moderate)	26 (12.1%)	24 (17.9%)	
<80 (low)	1 (0.5%)	2 (1.5%)	
Psychosocial well-being			
≥90 (high)	183 (85.1%)	106 (79.1%)	0.11
80-89 (moderate)	30 (14.0%)	25 (18.7%)	
<80 (low)	2 (0.9%)	3 (2.2%)	

Discussion

In our study, data from patients with similar clinical characteristics who underwent surgery under general anesthesia with a laryngeal mask and who were either discharged on the same-day or hospitalized were analyzed. Outpatient breast cancer surgery can be safely performed in patients with early-stage disease (T1-2, N0) whose airway is managed using a laryngeal mask, who do not require axillary dissection based on SLNB results, in whom no drains are placed at the surgical site, who have no comorbidities, and whose postoperative vital signs remain stable.

In recent years, outpatient approaches to breast cancer surgery have become increasingly widespread and have been implemented in many centers. This approach allows patients to be discharged on the day of surgery, thereby reducing the burden on the healthcare system and enabling a more comfortable postoperative recovery at home. Numerous studies have demonstrated that outpatient surgery provides clear advantages in terms of patient satisfaction, low complication rates, and cost-effectiveness[3]. Furthermore, modern anesthetic techniques, minimally invasive surgical approaches, and effective pain management have largely eliminated the need for prolonged hospitalization. Early mobilization, reduced risk of infection, and preservation of psychological well-being are additional key benefits of this approach[4]. Outpatient treatment allows patients with early-stage tumors who are scheduled for breast-conserving surgery to be discharged on the day of surgery.

The increase in the number of outpatient surgical procedures has raised concerns about patient safety, particularly regarding postoperative complications. However, studies have shown that outpatient breast cancer surgery, when performed with appropriate patient selection, does not increase complication rates. In particular, the incidence of complications such as SSI, hematoma, and seroma in patients discharged early has been reported to be comparable to that in traditional inpatient care[5,6]. Furthermore, outpatient surgeries conducted under structured protocols for postoperative pain management, mobilization, and patient education have been shown to reduce hospitalization without negatively affecting patient satisfaction or clinical outcomes[7]. In our study, the incidence of complications was similar between the two groups; no differences were observed in postoperative complications according to the Clavien-Dindo classification. These findings suggest that outpatient surgery can be safely implemented with appropriate infrastructure and patient selection.

Outpatient surgical procedures offer significant psychological benefits. Same-day discharge has been associated with reduced levels of anxiety, depression, and stress, owing to shorter hospital stays. Prolonged hospitalization has been reported to increase the patients' sense of dependency, which may negatively affect their psychological well-being. Early discharge allows patients to recover in their own environment, enhances comfort, and promotes better psychosocial adaptation through improved interactions with family members[8]. In a previous study, patients who underwent outpatient breast cancer surgery exhibited lower psychological distress scores on the psychological distress scale and better emotional adjustment scores on the emotional adjustment index[9]. This early recovery, facilitated by outpatient surgery, may positively influence patients' adherence to treatment, which in turn contributes to overall improvements in quality of life[10]. Patients' emotional status can be assessed using various tools, including the BREAST-Q questionnaire. The BREAST-Q is a widely used, patient-reported, breast-specific outcome measure that evaluates health-related quality of life. It is increasingly used to assess the impact of breast cancer treatment and surgical outcomes[11]. In our study, the BREAST-Q assessments of satisfaction and psychosocial well-being showed higher scores in patients who underwent outpatient breast cancer surgery; however, the differences between the groups were not statistically significant.

Outpatient surgery allows for faster turnover of hospital beds than traditional inpatient procedures, enabling healthcare institutions to serve more patients with the same bed capacity. Shorter hospitalization not only reduces costs but also allows more efficient use of hospital resources[3]. In outpatient surgeries, both the reduction in direct healthcare costs and the prevention of indirect productivity losses contribute to a more favorable cost-benefit ratio[8]. Outpatient breast surgery not only facilitates faster bed turnover and enables hospitals to serve more patients per unit time but also indirectly reduces productivity losses by decreasing the need for accompanying caregivers. Cost-effectiveness analyses have demonstrated that this approach is advantageous with respect to both direct costs (reduced hospital expenditures) and indirect costs (caregivers' time spent in the hospital and associated productivity losses)[12].

Study Limitations

The limitations of this study include the relatively small sample size and the lack of long-term follow-up data for patients who underwent outpatient surgery.

Conclusion

Outpatient breast cancer surgery is safe and effective for carefully selected patients. Particularly in early stage patients who have not received neoadjuvant therapy and who do not require axillary dissection, outpatient surgery has emerged as a sustainable option to ensure patient safety and improve healthcare system efficiency. However, prospective multicenter studies with long-term follow-up are needed to further evaluate the outcomes and refine the patient selection criteria.

Ethics

Ethics Committee Approval: Ethical approval for this study was obtained from the Ethics Committee of University of Health Sciences Türkiye, Ankara Etlik City Hospital (approval number: AEŞH-BADEK2-2025-199, date: 08.07.2025).

Informed Consent: This study was conducted in accordance with the Declaration of Helsinki, with informed consent obtained from eligible participants.

Acknowledgments

We would like to thank Editage (www.editage.com) for English language editing.

Footnotes

Authorship Contributions

Surgical and Medical Practises: M.O.K., F.A., L.D., Concept: M.O.K., F.A., B.A., L.D., Design: M.O.K., F.A., B.A., L.D., Data Collection or Processing: M.O.K., Fa.A., L.D., Analysis or Interpretation: M.O.K., Fa.A., M.F.S., L.D., Literature Search: M.O.K., M.F.S., L.D., Writing: M.O.K., M.F.S., B.A., L.D.

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

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

References

1. Kim R, Pharm AK, Wakisaka M, et al. Outcomes of outpatient breast cancer surgery at a private breast clinic. Breast J. 2018;24:628-32.
2. Duriaud HM, Kroman N, Kehlet H. Feasibility and safety of outpatient breast cancer surgery. Dan Med J. 2018;65:A5458.
3. Levine A, Khan M, Kelley JK, et al. Timing of mastectomy and the effect on the likelihood of outpatient surgery and cost savings in breast cancer patients. Surgery. 2024;175:671-6.
4. Cordeiro E, Zhong T, Jackson T, Cil T. The safety of same-day breast reconstructive surgery: an analysis of short-term outcomes. Am J Surg. 2017;214:495-500.
5. Bonnema J, van Wersch AM, van Geel AN, et al. Medical and psychosocial effects of early discharge after surgery for breast cancer: randomised trial. BMJ. 1998;316:1267-71.
6. Cordeiro E, Jackson T, Cil T. Same-day major breast cancer surgery is safe: an analysis of short-term outcomes using NSQIP data. Ann Surg Oncol. 2016;23:2480-6.
7. Purushotham AD, McLatchie E, Young D, et al. Randomized clinical trial of no wound drains and early discharge in the treatment of women with breast cancer. Br J Surg. 2002;89:286-92.
8. Susini T, Carriero C, Tani F, et al. Day surgery management of early breast cancer: feasibility and psychological outcomes. Anticancer Res. 2019;39:3141-6.
9. Marla S, Stallard S. Systematic review of day surgery for breast cancer. Int J Surg. 2009;7:318-23.
10. Dawe DE, Bennett LR, Kearney A, Westera D. Emotional and informational needs of women experiencing outpatient surgery for breast cancer. Can Oncol Nurs J. 2014;24:20-30.
11. Kuhlefelt C, Repo JP, Rasi V, et al. Preoperative reference values for breast cancer patients using the BREAST-Q. Breast. 2024;78:103832.
12. de Kok M, Dirksen CD, Kessels AG, et al. Cost-effectiveness of a short stay admission programme for breast cancer surgery. Acta Oncol. 2010;49:338-46.

DOI: 10.4274/anatoljmed.2025.79663
Anatol J Gen Med Res 2025;35(3):307-320

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Validation and Improvement of a Machine-learning-based LDL Prediction Model Using Retrospective Lipid Profile Data

Retrospektif Lipid Profili Verileri Kullanılarak Makine Öğrenmesi Tabanlı LDL Tahmin Modelinin Doğrulanması ve İyileştirilmesi

Ferhat Demirci[1,2], Murat Emeç[3], Mehmet Hilal Özcanhan[4], Özlem Gürsoy Doruk[2,5], Pınar Akan[2,5]

[1]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of Medical Biochemistry, İzmir, Türkiye
[2]Dokuz Eylül University Institute of Health Sciences, Department of Neurosciences, İzmir, Türkiye
[3]İstanbul University Faculty of Computer and Informatics, Department of Computer Engineering, İstanbul, Türkiye
[4]Dokuz Eylül University Faculty of Engineering, Department of Computer Engineering, İzmir, Türkiye
[5]Dokuz Eylül University Faculty of Medicine, Department of Medical Biochemistry, İzmir, Türkiye

Cite as: Demirci F, Emeç M, Özcanhan MH, Gürsoy Doruk Ö, Akan P. Validation and improvement of a machine-learning-based LDL prediction model using retrospective lipid profile data. Anatol J Gen Med Res. 2025;35(3):307-320

Abstract

Objective: Direct measurement of low-density lipoprotein cholesterol (LDL-C) is time-consuming and expensive when triglycerides (TG) exceed 400 mg/dL. We sought to validate and refine a machine-learning (ML) model for rapid estimation of LDL-C in hypertriglyceridemic sera.

Methods: We extracted 25.991 lipid profiles (TG: 400-800 mg/dL) collected between 2010 and 2022 from two Turkish university hospitals. After an 80/20 split, seven ML algorithms were trained; the top two (random forest and XGBoost) were stacked with a decision tree meta-learner (model-3). Performance on the external test set (n=1.279) was compared with that of direct homogeneous LDL-C assays and the Sampson's formula (NIH-Equ-2) using balanced accuracy, precision, recall, F1 score, specificity, Pearson correlation coefficient, and Bland-Altman analysis, following International Federation of Clinical Chemistry and Laboratory Medicine analytical performance specifications.

Results: Model-3 yielded balanced accuracy =99.3%, precision =98.9%, recall =98.9%, and specificity =99.8%. Predicted LDL-C correlated strongly with direct measurement (r=0.996, p<0.001) and reduced the mean absolute error by 54% compared with NIH-Equ-2. Only 0.39% of cases were underclassified relative to the European Society of Cardiology/European Atherosclerosis Society LDL-C risk categories. Bland-Altman plots demonstrated no significant proportional bias across the LDL-C range (mean bias =-0.2 mg/dL; 95% limits of agreement: -7.8 to+7.4 mg/dL).

Conclusion: A stacked ensemble ML model delivers near-assay accuracy for LDL-C prediction in high-TG samples and markedly outperforms current formula. Implementation could enable same day, low-cost LDL-C reporting without extra laboratory procedures, supporting faster dyslipidaemia management.

Keywords: LDL, lipid profile, machine-learning, artificial intelligence

Öz

Amaç: Trigliserid (TG) düzeyi 400 mg/dL'nin üzerine çıktığında düşük yoğunluklu lipoprotein kolesterol (LDL-K) ölçümü zaman alıcı ve maliyetli hale gelmektedir. Bu çalışmada, hipertrigliseridemik serumlarda hızlı LDL-K tahmini için geliştirilen makine öğrenmesi (ML) modelinin doğrulanması ve iyileştirilmesi amaçlandı.

Address for Correspondence/Yazışma Adresi: Murat Emeç, PhD, İstanbul University Faculty of Computer and Informatics, Department of Computer Engineering, İstanbul, Türkiye
E-mail: murat.emec@istanbul.edu.tr
ORCID ID: orcid.org/0000-0002-9407-1728

Received/Geliş tarihi: 20.09.2025
Accepted/Kabul tarihi: 03.11.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

Yöntem: 2010-2022 yılları arasında iki üniversite hastanesinden elde edilen 25,991 lipid profili (TG: 400-800 mg/dL) retrospektif olarak incelendi. Veriler %80/20 oranında ayrıldıktan sonra yedi ML algoritması eğitildi; en iyi iki algoritma (random forest, XGBoost), karar ağacı tabanlı bir meta-öğrenici ile birleştirilerek (model-3) istiflendi. Dış test setinde (n=1.279) modelin performansı doğrudan homojen LDL-K testleri ve Sampson formülüyle (NIH-Equ-2) karşılaştırıldı. Değerlendirmede dengeli doğruluk, kesinlik, duyarlılık, F1 skoru, özgüllük, Pearson korelasyonu ve Bland-Altman analizi kullanıldı; Uluslararası Klinik Kimya ve Laboratuvar Tıbbı Federasyonu analitik performans kriterleri dikkate alındı.

Bulgular: Model-3, dengeli doğruluk %99,3; kesinlik %98,9; duyarlılık %98,9 ve özgüllük %99,8 elde etti. Tahmin edilen LDL-K ile doğrudan ölçüm arasında güçlü korelasyon saptandı (r=0,996, p<0,001). Model, NIH-Equ-2 formülüne göre ortalama mutlak hatayı %54 azalttı. Avrupa Kardiyoloji Derneği/Avrupa Ateroskleroz Derneği LDL-K risk kategorilerine göre yanlış sınıflandırma oranı yalnızca %0,39 idi. Bland-Altman analizinde anlamlı orantısal yanlılık gözlenmedi (ortalama fark =-0,2 mg/dL; %95 güven aralığı, -7,8 ile +7,4 mg/dL).

Sonuç: Yığılmış topluluk ML modeli, yüksek TG düzeylerinde LDL-K tahmininde doğrudan testlere yakın doğruluk sağlamış ve mevcut formüllerden belirgin olarak üstün bulunmuştur. Modelin uygulanması, ek laboratuvar işlemleri olmadan aynı gün, düşük maliyetli LDL-K raporlamasına olanak tanıyabilir ve dislipidemi yönetiminde hız kazandırabilir.

Anahtar Kelimeler: LDL, lipid profili, makine öğrenmesi, yapay zeka

Introduction

Artificial intelligence (AI) applications in medicine have become increasingly widespread. Machine-learning (ML) developments have been widely adopted in medical AI (MAI) applications. The growth of MAI is due to the ever-increasing abundance of health data, the primary input for ML. Interest in MAI stems from its ability to generate diagnostic predictions from complex datasets. The MAI prediction and visualization applications have produced fast and accurate results in solving many medical problems[1].

However, dataset size is not the only driving force behind MAI. The number, variety, and accuracy of input data that are directly related to the output significantly influence the success of results produced by the designed ML models. Therefore, the data are expected to include all information related to the research output. In computer science, supervised ML (i.e., controlled ML) is currently the most widely used tool in MAI.

Computer algorithms such as computer-aided diagnosis or clinical decision support systems used for supporting diagnosis, decision-making, and prediction are classified as diagnostic devices[2]. The methods for clinical validation and development are similar to those for standard diagnostic tests. Therefore, medical devices used in MAI applications must undergo rigorous clinical and experimental validation before use in patients to ensure patient safety and the efficacy of the method. Additionally, the reproducibility of the ML prediction results is a major concern of the International Federation for Clinical Chemistry and Laboratory Medicine (IFFC)[3].

The MAI applications used for clinical validation vary according to their form, model, and function. Our work aims to rapidly predict patients' low-density lipoprotein (LDL) levels prior to the costly and delayed direct measurement. The present work aims to validate and improve our previous "LDL predictor model" (p-LDL-M {2}) designed for LDL prediction in patients with 400≤ triglyceride (TG) ≤ 800 mg/dL (abbreviated as high-TG for the rest of the article)[4]. Our ultimate goal is to recommend an improved MAI application for the research community. Although different models were tested in this work, our present and previous data were obtained from similar models of testing devices. In other words, there is no data discrepancy.

The design and validation of generalized, reproducible, and improved p-LDL-M models is a five-step iterative process. The steps involve formulating the problem, collecting and preparing the data, validating and selecting a model, and interpreting and finally implementing the model. The improved target model is obtained after identifying the model with the best performance. Finally, optimization and feature selection techniques are applied to further enhance the performance of the developed model. However, the performance results of interim models have not been included in the results section to save space and avoid repeating noncritical results.

LDL-C concentration is the principal target for lipid-lowering therapy and a key determinant of atherosclerotic cardiovascular disease (ASCVD) risk, as emphasized by recent European Society of Cardiology (ESC)/European Atherosclerosis Society (EAS) and American College of Cardiology/American Heart Association Guidelines. However,

the reliability of conventional LDL-C estimation equations is limited under hypertriglyceridemic conditions (TG >400 mg/dL). The Friedewald formula becomes invalid, the Martin-Hopkins method underestimates LDL-C in low LDL-C ranges, and even the National Institutes of Health (NIH)-Equ-2 method may introduce bias at very high-TG levels. Therefore, developing a robust ML-based estimation method is essential for precise LDL-C assessment in these patients.

The main aim of this study was to develop and validate a reliable ML-based model capable of predicting LDL-C levels in patients with hypertriglyceridemia (TG >400 mg/dL) before the costly and delayed laboratory measurements.

Materials and Methods

Study Design

Before commencing the study, the necessary approval was obtained from the Non-Interventional Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, (approval no: 2023/13-23, date: 12.04.2023) and Non-Interventional Ethics Committee of Dokuz Eylul University Faculty of Medicine (approval no: 2023/20-04, date: 14.06.2023). This study was first conducted at the University of Healt Sciences Türkiye, Dr. Suat Seren Chest Diseases and Chest Surgery Training and Research Hospital (hospital 1) and at the Dokuz Eylül University Research and Application Hospital (hospital 2) as the validation and improvement phase of the first phase. All experiments on humans were conducted according to relevant ethical guidelines and regulations. The experiments followed protocols approved by the Ethics Committees of hospital 1 and hospital 2. All experimental protocols used in this study have been reviewed and approved by the relevant institutional and/or licensing committee. The study's participants are three healthcare scientists and two engineering scientists from four institutions. The participants comply with the first recommendation to involve diverse stakeholders in developing clinically useful, practical, and ethical models. A total of 6.404 patient records with high-TG levels were presented in the hospital 1 biochemistry laboratory[4]. The hospital 2 biochemistry laboratory maintains records for 20.690 high-TG patients. During data analysis, records were omitted if they had missing results for total cholesterol (TC), TG, high-density lipoprotein (HDL), or LDL; if results exceeded the linear limits of specific analysis methods; if they contained zero or negative values; if they were from patients younger than 18 years of age; or if they lacked numerical data.

Of 27.094 patient records across the two hospitals, only 25.991 high-TG patient records (6.392 from hospital 1 and 19.599 from hospital 2) were processed using Python® software (Wilmington, Delaware, USA). As a rule of thumb in ML design and testing, the dataset was split into three training subsets (80% of all TG records, all-TG) and three testing subsets (20% of each dataset). Training and testing were conducted using nine combinations of datasets and three ML models. The study results are valid only for patients with high-TG levels. The TC, HDL, LDL, and TG were analyzed using Roche Cobas c702 (Mannheim, Germany) and Beckman Coulter AU5800 (California, USA) automated analyzers at hospital 1 and hospital 2, respectively. Our training and test sets were completely independent, meaning no test data was used in training the models. In addition, no data that would be unavailable during actual use were used; i.e., there was no data leakage in our analysis. With no data leakage and an 80:20 independent training-test split, our sample sets comply with the IFFC recommendations.

Study Population/Subjects

Our study population consisted of 6.392 lipid profile results obtained between January 2010 and December 2022 at hospital 1 and 19.559 results obtained between August 2011 and July 2022 at hospital 2. Standardized lipid profile data collected from the laboratory database included TC, TG, HDL, and LDL levels that were measured on the same day. Table 1 shows the main characteristics of the two high-TG study populations.

At hospital 1, 3.431 cases were male and 2.961 were female. The mean age of men was 49.72 years, while the mean age of women was 54.07 years. The mean directly measured LDL was 149.76±45.28 mg/dL. At hospital 2, 16.638 cases were male 2.961 were female. The mean ages were 56.81 years for men and 54.06 years for women. The mean measured direct LDL level was 151.10±46.44 mg/dL. Figure 1 displays the standard diagram used for reporting diagnostic accuracy, illustrating the progression of subjects throughout the study. Participants were divided into two datasets for statistical evaluation and ML analysis. The first dataset, typically comprising 80% of the participants, was used as the training set, while the remaining 20% formed the test set. In ML, the training set is utilized to build predictive models, and the test set is used to assess the prediction accuracy of those models.

Table 2 shows that the training set was divided into three groups. The first group (n=5113) contains direct LDL data from hospital 1 with TG levels >400 mg/dL. The second

group (n=19.599) contains hospital 2 direct LDL data with TG >400 mg/dL. The third group (n=24.712) comprises the combined direct LDL data from hospital 1 and hospital 2 for cases with TG levels >400 mg/dL. In the designed ML models, the training set of model-1 (the model most similar to our previous p-LDL-M {2} model) included only the first group of data, while model-2 used only the second data group. Model-3 was trained using the sum of the training sets[4]. It should be pointed out that the ML models also differ in their AI architectures.

To ensure unbiased comparability, the test set is the same for all three models: 20% of the hospital 1 data. The test set had to be from hospital 1 because testing newly designed models with a new training and test set from hospital 2 could have been misleading by eliminating cross-hospital prediction

The LDL level distribution of the 1279 test subjects is shown in Figure 1. The classification is based on the 2019 ESC/EAS Guidelines for managing dyslipidemia[5]. The most undesirable error in LDL level classification is assigning a patient to an LDL level below the actual classification (under-classification). Therefore, preventing under-classification was one of the primary objectives of the new model designs. The above properties of the study population indicate full compatibility with the sample size, race, gender, data diversity, and train-test set partitioning recommendations of IFFC.

Table 1. Characteristics of the study population

Characteristics	Units	Hospital 1: n=6.392 value ± SD	Hospital 2: n=19.599 value ± SD
Age		51.73±11.61	56.39±13.75
Male	years	49.72±11.20	56.81±14.05
Female	years	54.07±11.61	54.06±11.65
Sex			
Male	N/A	3431 (%53.7)	16638 (%84.9)
Female	N/A	2961 (%46.3)	2961 (%15.1)
Total cholesterol	mg/dL	243.16±52.79	248.46±57.82
	mmol/L	6.29±1.37	6.45±1.50
Triglycerides	mg/dL	510.98±96.71	509.33±97.13
	mmol/L	5.77±1.09	5.79±1.10
HDL	mg/dL	37.57±8.97	40.51±11.17
	mmol/L	0.97±0.23	0.92±0.21
Non-HDL cholesterol	mg/dL	205.62±48.12	207.96±51.30
	mmol/L	5.32±1.24	5.32±1.24
Direct LDL	mg/dL	149.76±45.28	151.10±46.45
	mmol/L	3.87±1.17	3.87±1.17

SD: Standart deviation, N/A: Not applicable, HDL: High-density lipoprotein, LDL: Low-density lipoprotein

Table 2. Description of model abbreviations according to data sets

Model	Training set	Test Set
Model-1	80% records of high-TG subjects only (group 1: 5.113)	20%
Model-2	100% records of high-TG subjects only (group 2: 19.599)	20%
Model-3	80% records of high-TG subjects in hospital 1 and 100% records of high-TG subjects in hostpital 2 (group 3: 27.712)	20%
NIH-Equ-2	The number of n in the groups for each model (group 1: 5.113, group 2: 19.599, group 3: 27.712)	20%

TG: Triglycerides, NIH-Equ-2: National Institutes of Health-Equ-2

Lipid Profile Testing

All lipid profile parameters were analyzed using automated chemistry analyzers: the Roche Cobas c702 in the biochemistry laboratory of hospital 1 and the Beckman Coulter AU5800 in the biochemistry laboratory of hospital 2. Only the initial test results of each patient were considered in the study; repeated measurements were excluded. TC and TG were determined using the enzymatic cholesterol esterase/oxidase and glycerol phosphate oxidase methods, respectively.

HDL levels were measured using a direct homogeneous assay that did not involve precipitation. LDL was quantified using a direct homogeneous assay that employs a selective protective agent to isolate LDL from chylomicrons, HDL, and very LDL, with measurement by the cholesterol esterase/oxidase method. The maximum allowable total error for LDL based on these methodologies was 11.9%. The actual total error rates recorded by the Roche c702 and Beckman AU5800 analyzers were 9.48% and 8.67%, respectively. Since both error rates were below the acceptable limit, the lipid profile data were deemed reliable and suitable for the study.

ML Analysis

Python 3.9 was used as the primary programming language. Data manipulation and analysis were performed using the Pandas Library (version 1.4.4) in Python. NumPy (version 1.21.5), which supports the handling of large, multidimensional arrays and provides advanced mathematical functions for array operations, was also used. ML models were developed using the Scikit-learn (Sklearn) library, version 1.0.2. To evaluate the contribution of individual features to model predictions, SHapley Additive exPlanations (SHAP) analysis was conducted. The SHAP library (version 0.42.1) was used to measure feature importance, and the corresponding

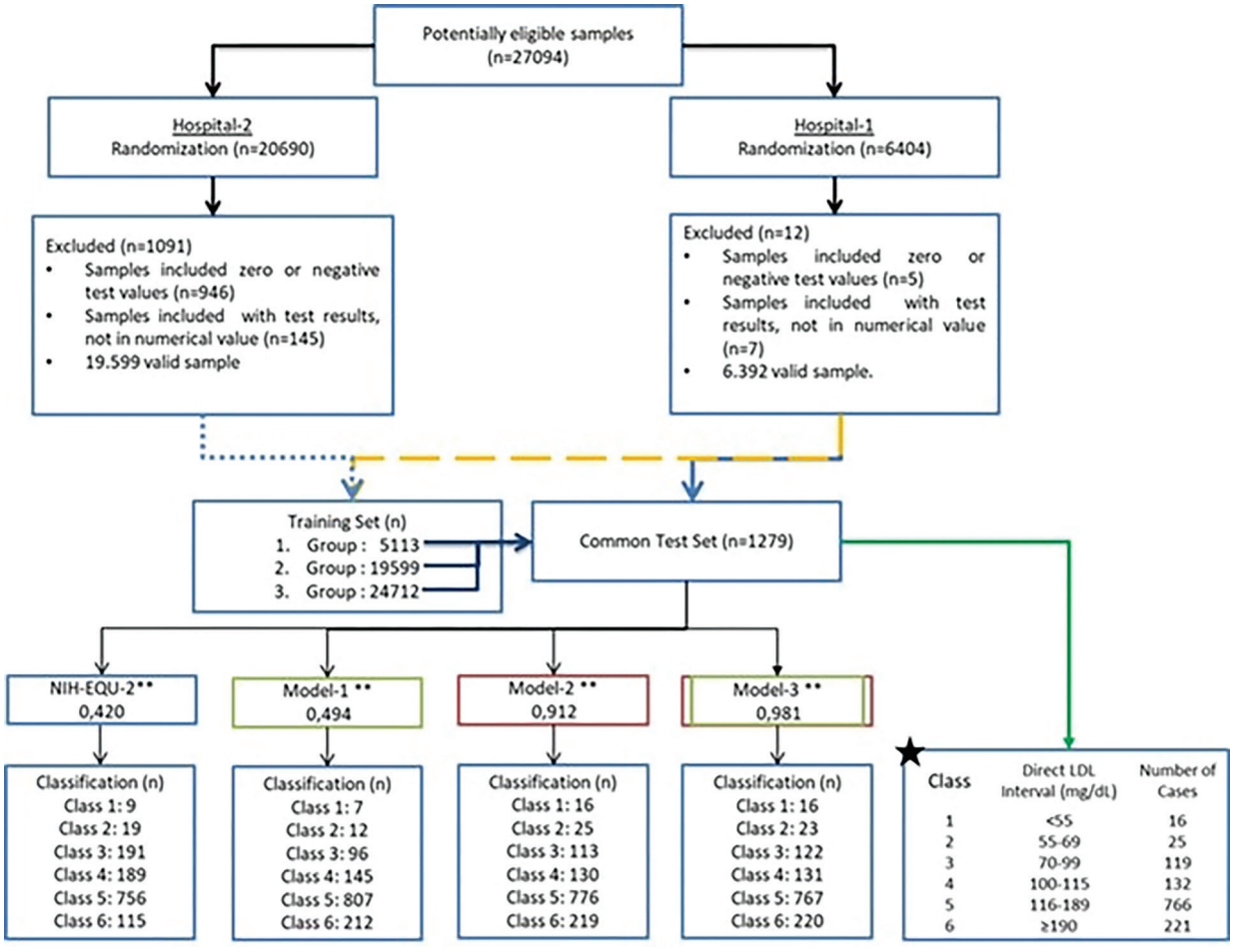


Figure 1. The flow of the subjects through the study shown in standards for reporting diagnostic accuracy diagram

NIH-Equ-2: National Institutes of Health-Equ-2, LDL: Low-density lipoprotein

SHAP graph for the LDL dataset is presented in Figure 2. Additionally, the GradientBoost library (version 1.5.0) was used to implement gradient boosting algorithms during model training and testing. A detailed literature search identified previous LDL estimation algorithms[6-10]. Our selection of analysis methods is based on various reviews of the use of the best combined ML algorithms that can detect linear or non-linear relationships between independent and dependent variables. Three ML algorithms (decision tree, random forest, and gradient boosting) were used to test linearity in the preprocessed data.

The linearity of the new dataset all-TG, obtained by combining the hospital 1 and hospital 2 subsets, was also verified. After verifying the linearity of the high-TG data set, the high-TG analysis was considered a regression analysis, and the prediction scores for LDL values from seven individual ML algorithms were determined separately for the three data sets (hospital 1, hospital 2, and all-TG). Next, LDL values for patients at both hospitals were predicted using a combination of the three algorithms described above. The new models were constructed by stacking the highest-performing algorithms: random forest, XGBoost, and decision tree. Stacking is an ensemble ML technique that combines multiple high-performing ML algorithms to produce the highest-performing predictive model. Early results indicated that the ensemble ML method using all-TG improved predictions of LDL values and LDL-level classification. Accordingly, the highest-performing random forest and XGBoost models were used as base learners, and the decision tree algorithm was stacked as the meta-learner to produce a meta-model.

Hence, a new stacked ensemble model was designed in our study. The untuned model was tested on the all-TG dataset and was later tuned. Hyperparameter tuning is a technique in ML model design used to achieve the highest final test scores across all performance parameters. After hyperparameter tuning, the highest-performing stacked ensemble ML model (model-3) shown in Figure 3 was obtained. The start-up model was model 1. Model 2 was obtained using only the hospital 2 dataset. Model-1 is our previous p-LDL-M {2} model in work[4]. The performances of all three models were tested. Model-3's LDL prediction was tested on the all-TG dataset to evaluate the effects of a larger dataset and model improvements on predictive performance. The predicted LDL values were placed into LDL-level classes in the final step, as shown in Figure 3. During the above design and selection processes, all key steps and recommendations of the IFFC for developing a medical ML application were followed. Figure 3, supported by the above-detailed explanations of our design's architecture, meets the reproducibility recommendation of IFFC.

Statistical Analysis

The measured direct LDL was accepted as the actual value. The predicted and calculated LDL values were compared with the actual LDL values. Statistical analyses were performed using IBM® SPSS® Statistics 26 for Windows®. A paired t-test was used to compare the means. Pearson's and Spearman's correlation tests were performed to assess the association between direct LDL and the predictions of the designed ML models and the Sampson-NIH equation (hereafter referred to as the NIH-Equ-2 method). Sometimes, the two correlations

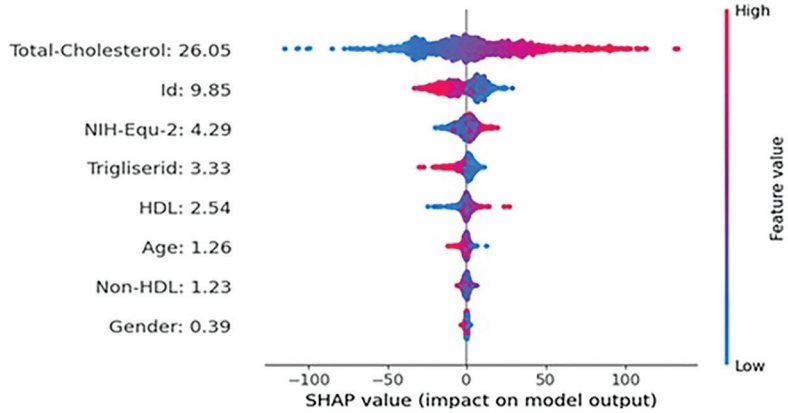


Figure 2. SHAP graph of direct LDL dataset parametersfeatures

SHAP: SHapley Additive exPlanations, LDL: Low-density lipoprotein

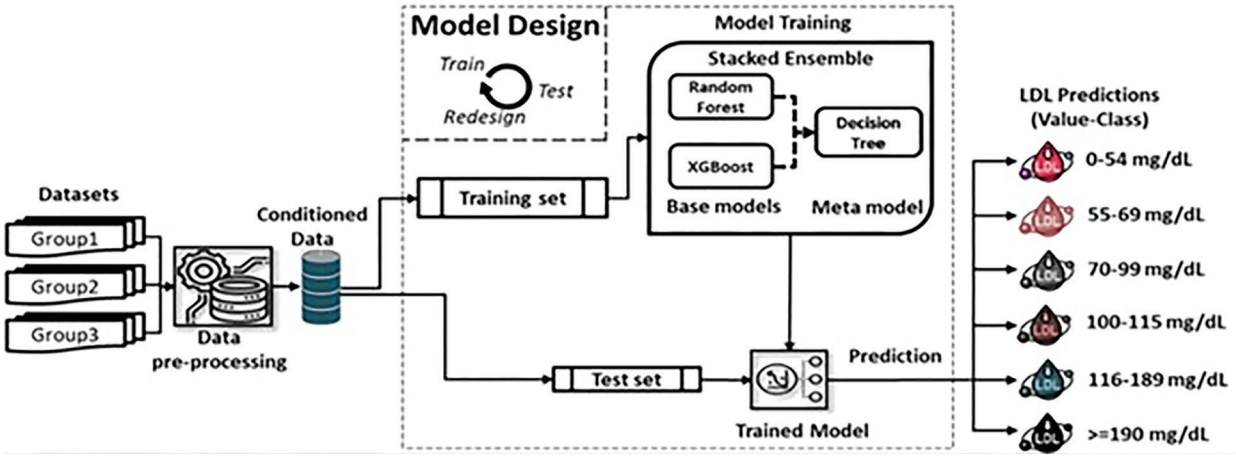


Figure 3. The proposed highest-performing LDL prediction ML model-3 architecture

LDL: Low-density lipoprotein, ML: Machine-learning

disagree on the strength of the correlation between an independent and dependent variable because of outliers[6]. Therefore, we included both in our work to determine whether such a disagreement existed. The present study found no discrepancy between the Pearson and Spearman correlation coefficient matrices; only minor differences, all less than 1.000, were observed.

Statistical significance was set at $p<0.05$, and a Passing-Bablok regression was conducted to determine the agreement between the prediction models and the current measurements. Bland-Altman plots were used to assess systematic bias across different direct LDL concentrations. In the Bland-Altman plots, differences between methods were plotted against direct LDL measurements.

The LDL levels classification performance of the designed models and the NIH-Equ-2 in classifying them was also assessed in accordance with the 2019 ESC/EAS Guidelines. Each subject's predicted or calculated LDL level class was compared with the subject's actual LDL level class. One way to measure and compare ML model performance is to report precision, recall, balanced accuracy, F1 score, and specificity for the model's predictions. The parameters used in the calculations for the equations in our study are defined as follows:

• **True positive:** The number of cases when the subject's LDL class was correctly identified.

• **False positive:** The number of cases when the subject's LDL class is incorrectly identified.

• **True negative:** The number of cases when the subjects out of an LDL class are correctly identified (not applicable in our study).

• **False negative:** The number of cases that subjects out of an LDL class is incorrectly identified Cohen's Kappa statistic was used to assess agreement between the designed models and NIH-Equ-2 classifications. The Kappa result can be interpreted as follows: values ≤0 indicating no agreement, 0.01-0.20 as none to slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1.00 as almost perfect agreement[6].

Results

Basic Statistics Results

A correlation matrix of designed models, NIH-Equ-2, and the actual direct LDL is given in Figure 4. The Figure shows that all models and NIH-Equ-2 results are strongly correlated with the exact values. However, the correlation for model-3 is exceptionally high (r=0.996). NIH-Equ-2's correlation with the actual values is the lowest, at 0.862.

The analysis of the scatter correlation plots of the compared methods (Figure 5) showed that the NIH-Equ-2 results were scattered, and the R2 value was low (R2=0.7443). Model-3 produced the best results, with R^2 close to 1 (R^2=0.9923) and low scatter. Interim model-2 exhibited a slightly high degree of scatter, with an R^2 value of 0.9494.

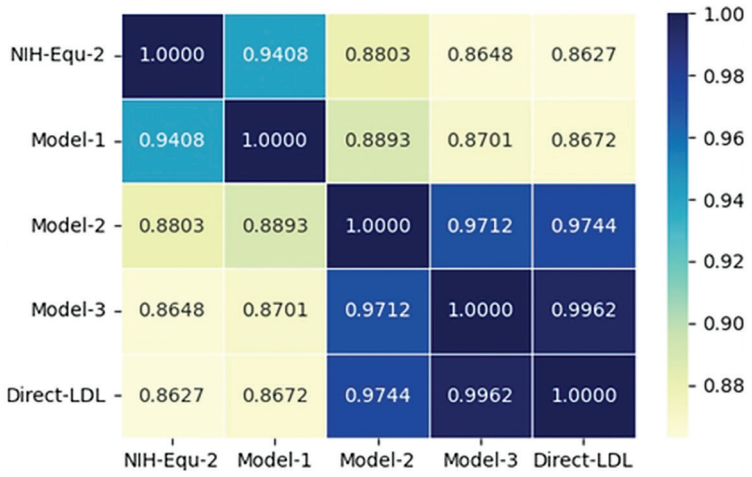


Figure 4. Correlation matrix of NIH-Equ-2, designed ML models and Direct LDL

NIH-Equ-2: National Institutes of Health-Equ-2, LDL: Low-density lipoprotein, ML: Machine-learning

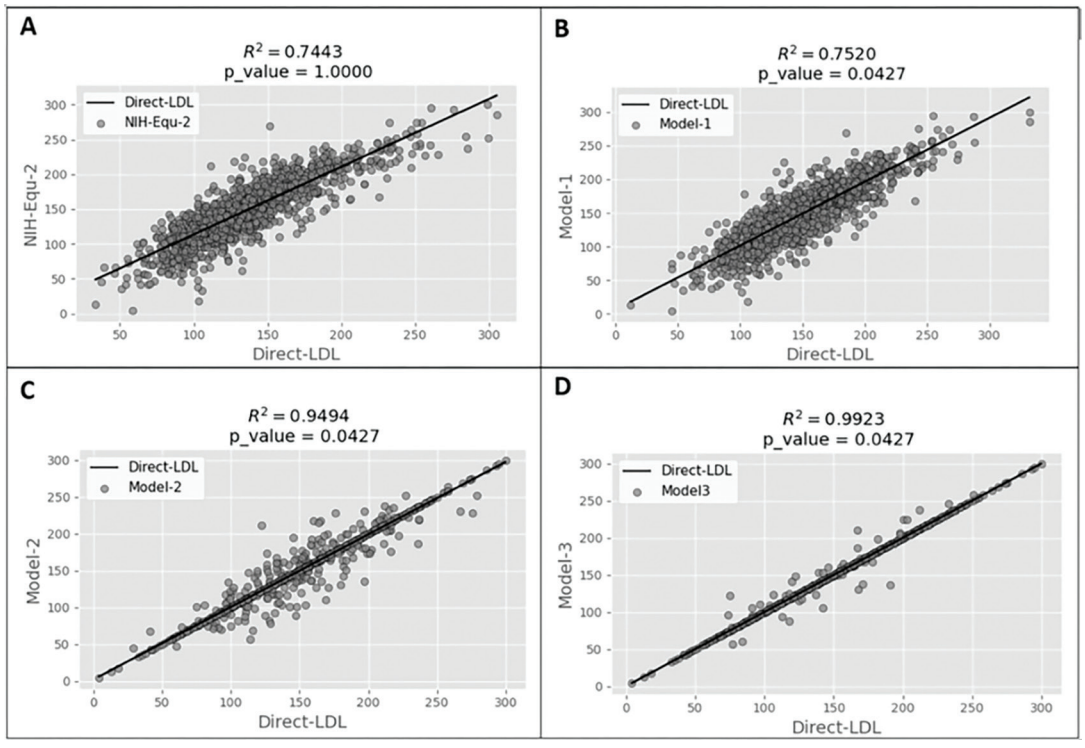


Figure 5. Scatter plots of correlations between predicted and direct LDL

A: NIH-Equ-2 vs. direct-LDL scatter graph, B: Model-1 vs. direct-LDL scatter graph, C: Model-2 vs. direct-LDL scatter graph, D: Model-3 vs. direct-LDL scatter graph, NIH-Equ-2: National Institutes of Health-Equ-2, LDL: Low-density lipoprotein

The receiver operating characteristic (ROC) curves for the six classes predicted by model-3 are shown in Figure 6. The area under the ROC curve (AUC) indicates the performance of a model across all possible classification thresholds. A value greater than 0.9 is considered outstanding. Our ROC curves showed a micro-averaged AUC of 97% across five classes.

The AUC for the non-critical class 3 was the lowest [89%; 95% confidence interval (CI), 8493%]. Therefore, the average AUC

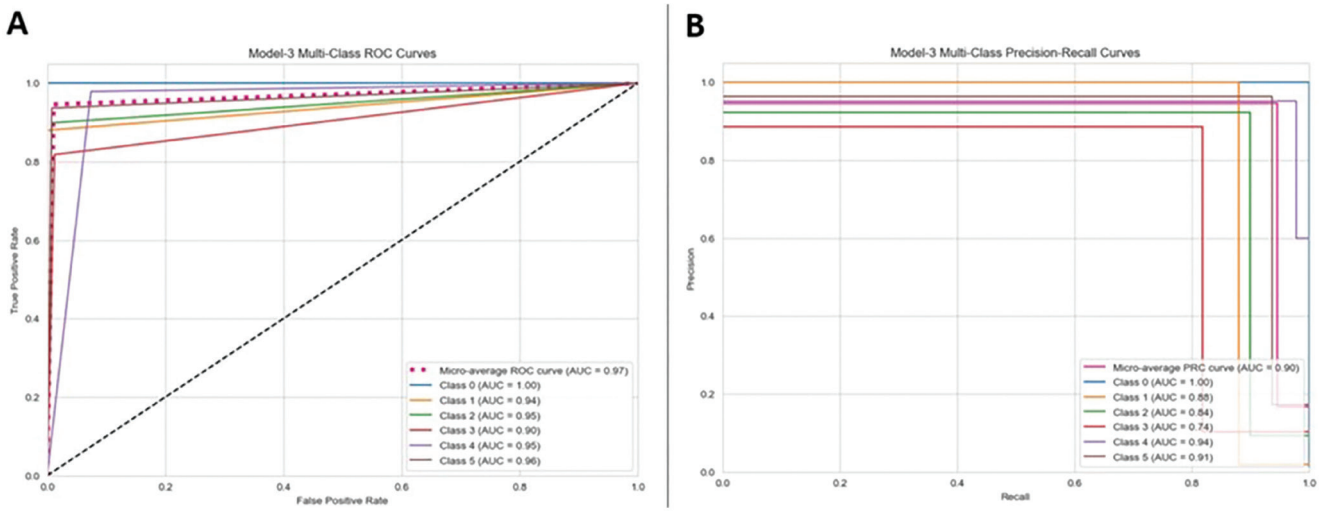


Figure 6. ROC curve and PRC of model-3 class predictions

A: Model-3 multi-class ROC curves, B: Model-3 multi-class precision-recall curves, ROC: Receiver operating characteristic, PRC: Precision-recall curve, AUC: Area under the curve

indicates that our proposed model achieves good predictive accuracy and precision across all classes. The average AUC of the precision-recall curve is also satisfactory at 0.90 (95% CI, 0.86-0.93). However, the AUC of class 3 is the lowest at 0.74 (95% CI, 0.71-0.77). Fortunately, class 3 is not the level at which LDL underestimation is critical. The AUC for the class 4 critical LDL level is high at 0.94 (95% CI, 0.90-0.98).

The Bland-Altman plot of direct LDL and model-3 is given in Figure 7. As observed, most measurements are above the mean and fall within the 95% CI. However, some values do not fall within the 95% CI, which requires an explanation. The reason for the high number of outliers in the CI is presented in the next section.

Kappa scores (Table 3) were obtained by comparing the actual LDL levels with the predicted and calculated LDL levels. The lowest Kappa score was 0.420 for NIH-Equ-2's LDL-level classification, while the highest was 0.981 for model-3. These results indicate that the performance of model 3 was best when data from hospital 1 and hospital 2 were combined.

Discussion

Our study aims to predict LDL levels in high-TG subjects. Based on our literature review, this study is the second ML-based study on high-TG subjects in Türkiye.

The categorical classification of patients' LDL levels is as important as the quantitative LDL value for guiding lipid-lowering therapy. Clinicians apply various treatments, from dietary changes and exercise to multidrug therapies, depending on the patient's LDL level. Therefore, the LDL values under study were categorized into classes according to the 2019 ESC/EAS Guideline[5]. The data preparation, model selection, design, and validation steps were completed in accordance with the IFFC recommendations. The most important findings of our research are discussed below.

As illustrated in Figure 1, class 1 (0-54 mg/dL) had the lowest number of cases, whereas class 5 (116-189 mg/dL) had the highest (766 cases). The mean values of the datasets play a crucial role in representing the characteristics of the studied population. Upon examining the lipid profiles of the individuals included in our research, it was observed that the average TC, TG, and LDL levels were elevated, whereas average HDL levels were comparatively low compared with similar ML studies[7,9,11]. These discrepancies may be attributed to the dietary patterns prevalent in our country. Nevertheless, with the exception of TG levels, the lipid values reported by the NIH in the multicenter study by Sampson et al. [12] were largely consistent with ours. In contrast, the other four centers reported lower lipid values than those in our study.

In prior studies focused on low-TG LDL prediction, random forest has been the most commonly used ML algorithm. However, alternative approaches such as XGBoost, deep neural networks, support vector machines, linear regression,

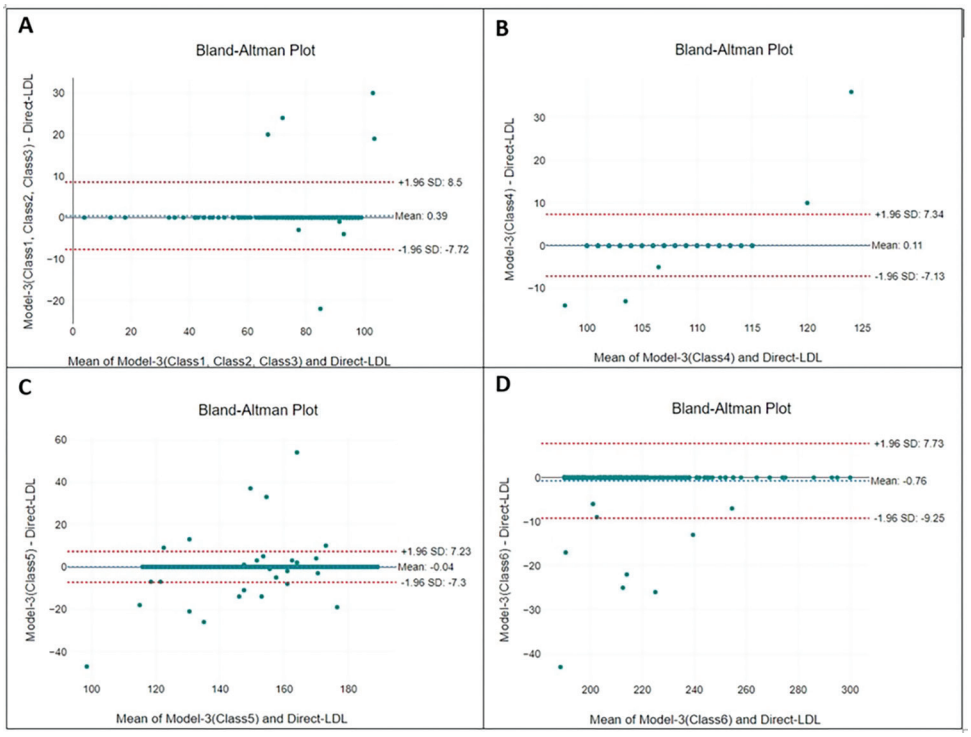


Figure 7. Bland-Altman plot between direct LDL and model-3

LDL: Low-density lipoprotein, SD: Standard deviation

Table 3. Model-1, model-2, model-3, and NIH-Equ-2 predicted Kappa scores of direct

Model/formula	Cohen's Kappa score
NIH-Equ-2	0.420[a]
Model-1	0.494[a]
Model-2	0.912[b]
Model-3	0.981[b]

[a]: Moderate aggrement, [b]: Almost perfect aggrement, NIH-Equ-2: National Institutes of Health-Equ-2

and k-nearest neighbors have also been employed[8,9,11]. In the present study, the stacked ensemble ML model-3 demonstrated superior performance and yielded the most accurate predictions.

Model 3 yielded several noteworthy findings. When its predictions were compared with direct LDL measurements, model-3 demonstrated the highest accuracy and correlation coefficients and the lowest error rates (mean absolute error, mean squared error, mean absolute percentage error). Notably, the best performance was achieved using the full all-TG dataset rather than the high-TG subset. When the entire dataset, including calculated LDL values, was utilized, model-3's prediction performance improved by 12.90% compared with predictions based solely on the high-TG group. Additionally, classification of LDL levels exhibited superior accuracy and minimal variability when using the all-TG dataset. Model-3 also outperformed the well-established NIH-Equ-2 method, showing a 13.45% improvement margin. These findings highlight the advantage of combining data from hospital 1 and hospital 2 to enhance LDL estimation. Moreover, the study confirms the following:

- The previously known strong correlations between TC, non-HDL, and direct LDL.
- The strong performance of NIH-Equ-2 in calculating LDL.
- The success of ML in estimating LDL values.
- The linear relationship between TG and LDL.

Model-3-predicted results and direct LDL measurements were significantly correlated (r=0.996). The algorithm results of Anudeep et al.[7] and Singh et al.[9] were also significantly associated with the direct-LDL measurements (0.98 and 0.982). The more robust correlation results

in independent studies indicate that ML algorithms and ensemble techniques can predict LDL values better than previously developed formulae[13].

Another interesting result was the difference between NIH-Equ-2 mean and the direct LDL mean (136.12±39.38 and 148.80±44.42 mg/dL, respectively). The relatively large difference was disappointing. In contrast, the mean value obtained in model 3 did not differ significantly. Our model-3's superior statistical performance was further supported by higher precision, recall, balanced accuracy, F1 score, and specificity.

The resulting SHAP graph is shown in Figure 2. TC was the most impactful feature in our SHAP graph (Figure 2). The impact of TC was validated by the highest Pearson correlation value, 0.844, in Figure 8. There was also a high correlation between TC and direct-LDL in the study by Chen et al.[14].

Beyond TC, the SHAP summary plot revealed that TG and non-HDL cholesterol made substantial positive contributions to the prediction of LDL, reflecting the well-known metabolic coupling between TG-rich lipoproteins and LDL particles. HDL-C exerted a mild inverse effect, consistent with its protective role in reverse cholesterol transport. Age showed a modest positive effect, whereas sex contributed minimally, likely because lipid distributions were similar between sexes in the dataset. These findings support the biological plausibility of the model outputs.

The scatter plot of model-3 results in Figure *5* shows that our model results are almost linear, in contrast to the scattered results of NIH-Equ-2. Model-3's p-value (Figure *5* 4.27%) is lower than the %5.46 of desirable biological variation database specifications for the LDL cholesterol[15]. Our scatter performance is also consistent with Anudeep et al.[7] low-scatter study. Our study also agreed with Anudeep et al.[7] find that different formulae can produce negative results, even though their correlation values (r) vary between 0.89 and 0.94.

The R^2 value of model-3 in our study was comparable to, yet slightly higher than, the R^2 reported in the study by Chen et al.[14] In the research conducted by Kim et al.[16] where the XGBoost method-also employed in our study-was applied consecutively, the R^2 value was relatively high but still moderately lower than that achieved by our model.

Extensive evidence from epidemiological studies, Mendelian randomization analyses, and randomized controlled trials has established a log-linear association between LDL levels and ASCVD. Consequently, clinical guidelines consistently emphasize lipid-lowering therapies as essential for improving ASCVD-related outcomes. The effectiveness of these interventions is supported by foundational scientific research, clinical data, genetic studies, randomized trials, and population-based analyses[17,18]. Furthermore, LDL concentrations directly inform the selection and dosage of cholesterol-lowering treatments. One study, for instance, reported that each 1 mmol/L reduction in LDL was associated with a 20% decrease in major cardiovascular events[19].

Previous studies have shown that traditional predictive models and formulae exhibit greater error rates at lower LDL concentrations (<70 mg/dL). In contrast, our model-3 achieved a classification error rate of just 4.8% (2 out of 41) in this range (Table 4), outperforming the Weill-Cornell model, which had an error rate of 7.5%[9]. Similar performance was observed in the study by Çubukçu and Topcu[11] although it is important to note that their cohort consisted of patients with TG levels between 177 and 399 mg/dL[9]. Based on the 2019 ESC/EAS Guidelines, our model exhibited a 3.4% (4/119) classification error across the first three LDL categories (LDL <100 mg/dL). For comparison, error rates were 43.75% (77/176) for the NIH-Equ-2 formula, 11.4% (143/1254) for the Weill-Cornell model, and 3.47% (53/1528) for the model by Anudeep et al.[7] In the study by Barakett-Hamade et al.[8] LDL values were categorized into three groups, with the lowest group defined as <80 mg/dL. The misclassification rate for this category was 12.5% (793/6327).

It is well established that elevated LDL levels contribute significantly to morbidity and mortality among patients with cardiovascular disease; intensive hyperlipidemia management has been shown to improve quality of life, particularly in patients who are over-classified[20,21]. A correct or slightly higher classification of LDL levels (over-classification) ensures that patients receive appropriate or more aggressive treatment regimens. In our study, model 3 occasionally overclassified LDL levels by one category (Table 4). However, we argue that this slight overestimation poses minimal clinical risk, as it would lead to intensified treatment, which is generally safer than the risk of undertreatment[22]. Notably, therapies such as ezetimibe and monoclonal antibodies, when added to statin-based treatment, effectively reduce LDL levels and improve cardiovascular outcomes and overall survival. Studies have shown that inclusion of ezetimibe lowers the risk of cardiovascular events without increasing adverse effects or toxicity, even in patients with acute coronary syndromes or with already optimal LDL

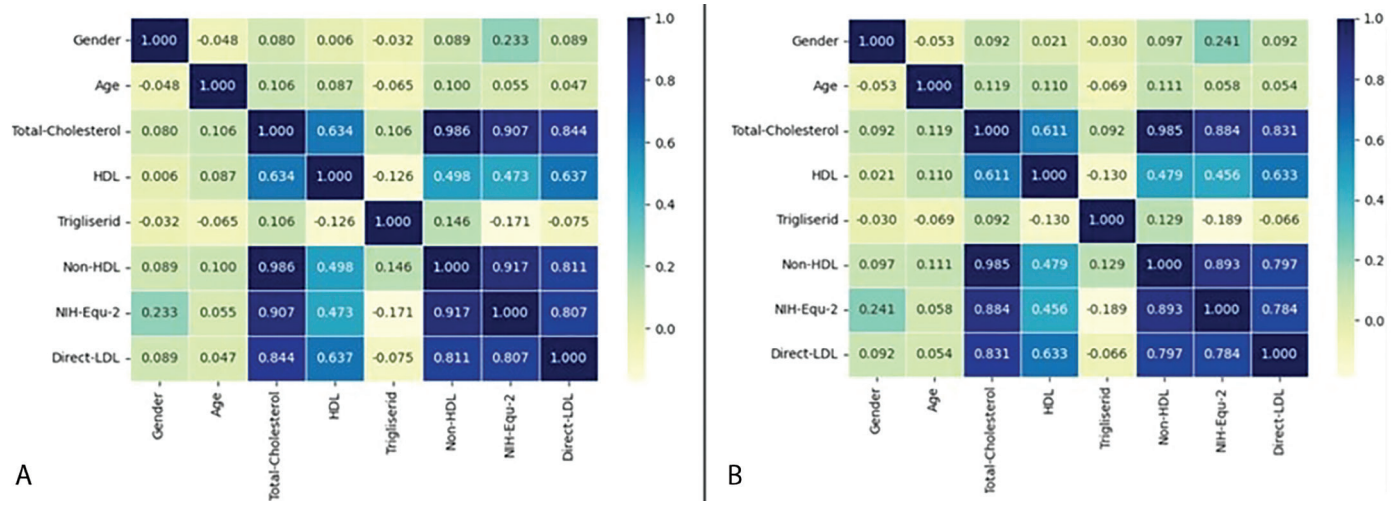


Figure 8. Correlation matrices of direct LDL parameters

A: Pearson correlation matrix, B: Spearman correlation matrix, NIH-Equ-2: National Institutes of Health-Equ-2, LDL: Low-density lipoprotein, HDL: High-density lipoprotein

Table 4. Results of classifications by model-3 (n=1279)

Model-3 class	Direct LDL class						
	1	2	3	4	5	6	Total
1	16	0	0	0	0	0	16
2	0	23	0	0	0	0	23
3	0	2	117	2	1	0	122
4	0	0	1	128	2	0	131
5	0	0	1	2	762	2	767
6	0	0	0	0	1	219	220
Total direct LDL	16	25	119	132	766	221	1279

LDL: Low-density lipoprotein

levels. Moreover, there is currently no evidence suggesting harm from adding ezetimibe in cases with moderate LDL concentrations[23,24]. In contrast, under-classification may have more serious consequences, such as misclassifying a high-risk patient from class 5 to class 4, resulting in inadequate treatment. Importantly, the under-classification rate of model 3 was remarkably low-only 0.39% (three of 766 cases).

Lipid-lowering therapies impose a considerable financial burden on healthcare systems globally[25]. Although beta-quantification is considered the reference standard for LDL measurement, it is not feasible for routine clinical use due to its high cost and labor intensity[26]. As a result, enzymatic and homogeneous immunoassay methods have largely replaced it. Following the establishment of correlations between LDL and other lipid fractions, direct LDL measurement techniques were gradually replaced by calculation-based methods such as the Friedewald, Martin-Hopkins, and NIH-Equ-2 formulae. Each of these approaches offers distinct advantages. For instance, whereas combined hospital datasets perform well when TG levels exceed 400 mg/dL, NIH-Equ-2 can extend its calculations up to 800 mg/dL. Notably, the Martin-Hopkins method utilizes variable adjustment factors tailored to TG levels, allowing for accurate estimations even in non-fasting individuals[9]. To overcome the limitations of traditional formulae, recent studies have shifted focus toward ML-based LDL prediction models. Based on our literature review, this study is the second to use ML algorithms to predict LDL levels in patients with TG >400 mg/dL, following our initial publication. This makes our work particularly relevant in clinical settings where

precise LDL classification is crucial for initiating appropriate therapeutic interventions[5]. The effectiveness of our model is further validated by its Kappa score, a statistical measure of classification agreement. Model-3 achieved a Kappa score of 0.981, significantly outperforming the NIH-Equ-2 formula, which had a Kappa score of 0.420. These results indicate that model-3 is in "almost perfect agreement" with direct LDL measurements, whereas NIH-Equ-2 falls under the category of "moderate agreement". Although numerous studies endorse the reliability of NIH-Equ-2 for LDL estimation[12,27], it is important to note that NIH-Equ-2 was calibrated against LDL values derived from the reference beta-quantification technique. In contrast to many existing studies, our research utilized results from two distinct auto-analyzer platforms. Therefore, variations in accuracy may stem from differences in analytical methods, ML models, or the analyzers themselves[26].

The number of predictions deviating from the mean in the Bland-Altman plot comparing direct LDL and model-3 was high. The highest number deviation is observed in classes 5 and 6 (Figure 7). The large numbers in both classes are due to the very wide range of class 5 (116-189) and the open-ended range of class 6 (>190). The number of deviations from the acceptable region for the critical class 4 is four. The percentage of undesired predictions is negligible, with 4 instances out of 1279 total predictions, i.e., 0.31%.

The recommendations of the IFCC Working Group[3] have been followed carefully in our present work. In the last stage, the prediction results for the same high-TG patients were used to cross-validate the final model's reliability. Thus, we can now claim the model's applicability to data from different hospitals. We believe that our work will increase acceptance of ML in MAI applications and pave the way for future real-world applications.

Study Limitations

Our study has some limitations. First, because our data were retrospective and beta-quantification is not used in routine laboratory practice, the more commonly used direct homogeneous immunoassay was employed. In this study, LDL-C values were obtained using a commercial homogeneous direct assay. However, previous studies have shown that homogeneous LDL-C methods can overestimate LDL-C levels by 7-12% compared to the beta-quantification[28]. This discrepancy is particularly pronounced in hypertriglyceridemic individuals (TG ≥ 400 mg/dL), where cholesterol from very LDL and intermediate DL fractions may be mistakenly included in LDL-C measurements. As a result, the LDL-C values used for training our model may contain a systematic bias. Future studies should validate the model using beta-quantification-based reference data or incorporate an adjustment to account for potential measurement bias in homogeneous LDL-C assays. Secondly, the effects of diseases that may influence the lipid profiles could not be assessed independently because of the study's retrospective design. Third, sub-categorization of ethnic groups could not be performed. Since the target range was TG >400 mg/dL, TG values could not be subcategorized or analyzed in detail.

Conclusion

An MAI application that fully complied with the IFFC recommendations for predicting LDL using ensemble ML methods was presented. Performance results indicate that our newly designed ML estimation model, model-3, for predicting target high-TG values outperforms the NIH-Equ-2 formula and our previous model. Our work can be included in the routine lipid profile without changing the main principles and methods if similar work is planned as a multicenter study, enriched with data from different races, and expanded to include multiple autoanalyzers.

Ethics

Ethics Committee Approval: Before commencing the study, the necessary approval was obtained from the Non-Interventional Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, (approval no: 2023/13-23, date: 12.04.2023) and Non-Interventional Ethics Committee of Dokuz Eylul University Faculty of Medicine (approval no: 2023/20-04, date: 14.06.2023).

Informed Consent: Retrospective design.

Footnotes

Authorship Contributions

Surgical and Medical Practises: F.D., M.E., Ö.G.D., P.A., Concept: F.D., M.E., M.H.Ö., P.A., Design: F.D., M.E., M.H.Ö., Ö.G.D., P.A., Data Collection or Processing: F.D., M.E., Ö.G.D., Analysis or Interpretation: F.D., M.E., M.H.Ö., Ö.G.D., Literature Search: F.D., M.E., M.H.Ö., Ö.G.D., P.A., Writing: F.D., M.E.

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

One of the authors of this article (F.D.) is a member of the Editorial Board of this journal. He was completely blinded to the peer review process of the article.

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

References

1. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019;6:94-8.
2. Park SH, Choi J, Byeon JS. Key principles of clinical validation, device approval, and insurance coverage decisions of artificial intelligence. Korean J Radiol. 2021;22:442-53.
3. Master SR, Badrick TC, Bietenbeck A, Haymond S. Machine learning in laboratory medicine: recommendations of the IFCC Working Group. Clin Chem. 2023;69:690-8.
4. Demirci F, Emec M, Gursoy Doruk O, Ormen M, Akan P, Hilal Ozcanhan M. Prediction of LDL in hypertriglyceridemic subjects using an innovative ensemble machine learning technique. Turkish Journal of Biochemistry. 2024;48:641-52.
5. Mach F, Baigent C, Catapano AL, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. Eur Heart J. 2020;41:111-88.
6. McHugh ML. Interrater reliability: the Kappa statistic. Biochem Med (Zagreb). 2012;22:276-82.
7. Anudeep PP, Kumari S, Rajasimman AS, Nayak S, Priyadarsini P. Machine learning predictive models of LDL-C in the population of eastern India and its comparison with directly measured and calculated LDL-C. Ann Clin Biochem. 2022;59:76-86.
8. Barakett-Hamade V, Ghayad JP, Mchantaf G, Sleilaty G. Is machine learning-derived low-density lipoprotein cholesterol estimation more reliable than standard closed form equations? Insights from a laboratory database by comparison with a direct homogeneous assay. Clin Chim Acta. 2021;519:220-6.
9. Singh G, Hussain Y, Xu Z, et al. Comparing a novel machine learning method to the Friedewald formula and Martin-Hopkins equation for low-density lipoprotein estimation. PLoS One. 2020;15:e0239934.
10. Hidekazu I, Nagasawa H, Yamamoto Y, et al. Dataset dependency of low-density lipoprotein-cholesterol estimation by machine learning. Ann Clin Biochem. 2023;60:396-405.
11. Çubukçu HC, Topcu Dİ. Estimation of low-density lipoprotein cholesterol concentration using machine learning. Lab Med. 2022;53:161-71.
12. Sampson M, Ling C, Sun Q, et al. A new equation for calculation of low-density lipoprotein cholesterol in patients with normolipidemia and/or hypertriglyceridemia. JAMA Cardiol. 2020;5:540-8.
13. Atabi F, Mohammadi R. Clinical validation of eleven formulas for calculating LDL-C in Iran. Iran J Pathol. 2020;15:261-7.
14. Chen L, Rong C, Ma P, Li Y, Deng X, Hua M. A new equation for estimating low-density lipoprotein cholesterol concentration based on machine learning. Medicine (Baltimore). 2024;103:e37766.
15. Westgard J. Desirable biological variation database specifications. Available from: https://www.westgard.com/clia-a-quality/quality-requirements/238-biodatabase1.html
16. Kim Y, Lee WK, Lee W. Prediction of low-density lipoprotein cholesterol levels using machine learning methods. Lab Med. 2024;55:471-84.
17. Writing Committee; Lloyd-Jones DM, Morris PB, et al. 2022 ACC expert consensus decision pathway on the role of nonstatin therapies for LDL-cholesterol lowering in the management of atherosclerotic cardiovascular disease risk: a report of the American College of Cardiology Solution Set Oversight Committee. J Am Coll Cardiol. 2022;80:1366-418. Erratum in: J Am Coll Cardiol. 2023;81:104.
18. Catapano AL, Graham I, De Backer G, et al. 2016 ESC/EAS Guidelines for the management of dyslipidaemias. Eur Heart J. 2016;37:2999-3058.
19. Cholesterol Treatment Trialists' (CTT) Collaboration; Baigent C, Blackwell L, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet. 2010;376:1670-81.
20. Emerging Risk Factors Collaboration; Di Angelantonio E, Gao P, et al. Lipid-related markers and cardiovascular disease prediction. JAMA. 2012;307:2499-506.
21. Ference BA, Yoo W, Alesh I, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. J Am Coll Cardiol. 2012;60:2631-9.
22. Sathiyakumar V, Blumenthal RS, Elshazly MB: New information on accuracy of LDL-C estimation. American College of Cardiology. March 20. 2020. (Accessed: March 27, 2024). Available from: https://www.acc.org/latest-in-cardiology/articles/2020/03/19/16/00/new-information-on-accuracy-of-ldl-c-estimation
23. Oyama K, Giugliano RP, Blazing MA, et al. Baseline low-density lipoprotein cholesterol and clinical outcomes of combining ezetimibe with statin therapy in IMPROVE-IT. J Am Coll Cardiol. 2021;78:1499-507.
24. Koskinas KC, Siontis GCM, Piccolo R, et al. Effect of statins and non-statin LDL-lowering medications on cardiovascular outcomes in secondary prevention: a meta-analysis of randomized trials. Eur Heart J. 2018;39:1172-80.
25. Michaeli DT, Michaeli JC, Boch T, Michaeli T. Cost-effectiveness of lipid-lowering therapies for cardiovascular prevention in Germany. Cardiovasc Drugs Ther. 2023;37:683-94.
26. Contois JH, Langlois MR, Cobbaert C, Sniderman AD. Standardization of apolipoprotein B, LDL-cholesterol, and Non-HDL-cholesterol. J Am Heart Assoc. 2023;12:e030405.
27. Sampson M, Wolska A, Meeusen JW, Otvos J, Remaley AT. The Sampson-NIH Equation is the preferred calculation method for LDL-C. Clin Chem. 2024;70:399-402.
28. Yano M, Matsunaga A, Harada S, et al. Comparison of two homogeneous LDL-cholesterol assays using fresh hypertriglyceridemic serum and quantitative ultracentrifugation fractions. J Atheroscler Thromb. 2019;26:979-88.

DOI: 10.4274/anatoljmed.2025.48343
Anatol J Gen Med Res 2025;35(3):321-329

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Keratocan and *Nicotinamide Mononucleotide Adenylyltransferase 1* are Prognostic Markers in Kidney Renal Papillary Cell Carcinoma

Keratokan ve *Nikotinamit Mononükleotit Adenililtransferaz 1* Renal Papiller Hücre Karsinomunda Prognostik Belirteçlerdir

Gülçin Çakan Akdoğan

Dokuz Eylül University, İzmir International Biomedicine and Genome Insitute, Department of Biomedicine and Health Technologies; İzmir Biomedicine and Genome Center, İzmir, Türkiye

Cite as: Çakan Akdoğan G. Keratocan and nicotinamide mononucleotide adenylyltransferase 1 are prognostic markers in kidney renal papillary cell carcinoma. Anatol J Gen Med Res. 2025;35(3):321-329

Abstract

Objectives: Keratan sulfate proteoglycans (KSPGs) are the least studied of the proteoglycans with emerging roles in cancer. Diverse roles of KSPGs are known; however, the KSPG interactome has not been defined to date. The aim here was to define a KSPG interactome and to determine biomarkers in cancer patients using published datasets and bioinformatic analysis tools.

Methods: The STRING database was utilized to compile a KSPG interactome centered around keratocan (KERA), lumican, fibromodulin and osteoglycin. STRING, Xena Browser, Gene Set Cancer Analysis Platform, and SmulTCan were used to perform data analyses.

Results: A 56-protein KSPG interactome network was discovered. Extracellular matrix (ECM) organization, PI3K-AKT signaling, ECM-receptor interactions, and nicotinamide metabolism pathways were enriched in the interactome. Multivariate prognostic modeling with Glmnet identified a seven-gene prognostic signature in kidney renal papillary cell carcinoma (KIRP), stratifying patients into distinct risk groups with strong predictive accuracy (area under the curve of 78.9%). Among these, *KERA* emerged as a biomarker of poor prognosis, while *nicotinamide mononucleotide adenylyltransferase 1 (NMNAT1)* emerged as a protective biomarker in KIRP.

Conclusion: This study describes a methodology to utilize available datasets for protein interactome discovery and for querying cancer prognostic genes. In this study, an extensive KSPG interactome was identified. This is the first report linking *KERA* and *NMNAT1* to prognosis in KIRP.

Keywords: KSPG, *keratocan*, *NMNAT1*, prognosis, kidney renal papillary cell carcinoma

Öz

Amaç: Keratan sülfat proteoglikanlar (KSPG'ler), kanserdeki rolleri en az çalışılmış proteoglikan grubudur. KSPG'lerin çeşitli işlevleri bilinmekle birlikte, KSPG etkileşim ağı (interaktom) bugüne kadar tanımlanmamıştır. Bu çalışmada, KSPG interaktomunun tanımlanması ve yayımlanmış veri setleri ile biyoinformatik analiz araçları kullanılarak kanser hastalarında prognostik açıdan önemli biyobelirteçlerin belirlenmesi amaçlanmıştır.

Yöntem: STRING veritabanı keratokan (KERA), lumikan, fibromodulin ve osteoglikan merkezli bir KSPG interaktomu oluşturmak için kullanılmıştır. Veri analizlerinde STRING, Xena Browser, Gene Set Cancer Analysis Platform ve SmulTCan araçları kullanılmıştır.

Address for Correspondence/Yazışma Adresi: Asst. Prof., Gülçin Çakan Akdoğan, Dokuz Eylül University, İzmir International Biomedicine and Genome Insitute, Department of Biomedicine and Health Technologies; İzmir Biomedicine and Genome Center, İzmir, Türkiye
E-mail: gulcin.cakan@ibg.edu.tr
ORCID ID: orcid.org/0000-0002-6356-5979

Received/Geliş tarihi: 02.10.2025
Accepted/Kabul tarihi: 04.11.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

Bulgular: 56 proteinden oluşan bir KSPG etkileşim ağı keşfedilmiştir. Hücre dışı matriks (HDM) organizasyonu, PI3K-AKT sinyal yolu, HDM-reseptör etkileşimleri ve nikotinamid metabolizması interaktomda zenginleştirilmiş yollar arasında bulunmuştur. Glmnet ile yapılan çok değişkenli prognostik modelleme, böbrek renal papiller karsinom (KIRP) hastalarında yedi genlik bir prognostik imza ortaya koymuş; bu imza, hastaları güçlü prediktif doğrulukla (eğri altındaki alan %78,9) farklı risk gruplarına ayırmıştır. Bu analizde *KERA* kötü prognoz biyobelirteci olarak, *nikotinamit mononükleotit adenililtransferaz 1 (NMNAT1)* ise koruyucu bir biyobelirteç olarak öne çıkmıştır.

Sonuç: Bu çalışma, mevcut veri setlerinin protein interaktom keşfi ve kanser prognostik genlerinin belirlenmesinde nasıl kullanılabileceğini gösteren bir metodoloji sunmaktadır. Çalışma sonucunda kapsamlı bir KSPG interaktomu tanımlanmıştır. Bu, *KERA* ve *NMNAT1*'ın KIRP prognozu ile ilişkilendirildiğini gösteren ilk rapordur.

Anahtar Kelimeler: KSPG, *keratokan*, *NMNAT1*, prognoz, böbrek renal papiller hücrekarsinomu

Introduction

The extracellular matrix (ECM) surrounding the tumor is the first stromal component encountered by metastatic cancer cells before they move to distant sites. Traditionally considered a physical blockade, the ECM is now recognized as a dynamic regulator of tumor invasion, metastasis, and immune modulation. Increased ECM stiffness can suppress metastatic spread, whereas activation or overexpression of matrix metalloproteinases that digest the ECM has been associated with increased cellular migratory capacity and metastasis[1,2]. The influence of the ECM on metastatic cancer cells is complex, owing to its highly intricate and dynamic composition. The main macromolecular constituents are collagens and proteoglycans, both diverse groups that contribute to the organ- and tissue-specific composition of the ECM[3]. Proteoglycans are classified into several types, named after their glycosaminoglycan side chains, such as heparan sulfate, dermatan sulfate, or keratan sulfate. The determinants of the side chains are the specific types of disaccharide repeats, such as glucuronic acid or iduronic acid paired with glucosamine for heparan sulfates, and galactose paired with N-acetylglucosamine for keratan sulfates. Each proteoglycan molecule may have different lengths of disaccharide repeats, different sulfate loads on these repeats, and different core proteins[3]. Diversity in structure is likely to be reflected in functional diversity, making it especially challenging to pinpoint the roles of a particular proteoglycan[4].

Most research on the role of proteoglycans in cancer metastasis has focused on heparan sulfate proteoglycans[5]. Keratan sulfate proteoglycans (KSPGs), on the other hand, are the least studied in this context[6]. Recent transcriptomic studies have identified that *carbohydrate sulfotransferase 6 (CHST6)*, a keratan sulfate-specific sulfotransferase, is part of a genetic signature associated with poor prognosis in cancer[7-9]. Sulfation of KSPGs is a major influence on their function and solubility, as evidenced by the formation of KSPG aggregates in patients with macular corneal dystrophy, which is linked to *CHST6* mutations[10-12]. These data indicate that KSPG regulation is important in cancer progression.

Reports on the role of lumican (LUM) in cancer progression are contradictory[13]. *LUM* overexpression has been correlated with poor prognosis in colorectal, lung, and squamous cell carcinomas[14-17]. In contrast, *LUM* overexpression was shown to decrease the malignancy of melanoma cells and to correlate negatively with osteosarcoma tumor progression[18]. These contradictory findings may be related to the location of the KSPG: in the ECM or within the cancer cell itself. Additionally, differences in chain lengths and sulfation status of LUM in different cell types influence its interactions with other proteins and its roles[19].

KSPG binding proteins are not fully defined in the literature, as exemplified by the absence of a dedicated gene ontology (GO) term (https://www.ebi.ac.uk/QuickGO/). Proteoglycan binding GO term contains child terms that include syndecan binding (GO:0045545), chondroitin sulfate proteoglycan binding (GO:0035373), heparan sulfate proteoglycan binding (GO:0043395). This study aimed to define the KSPG interactome and to explore the expression and prognostic significance of its components in cancer, using published datasets and openly available bioinformatic tools and databases.

Here, KSPG-interacting proteins were identified through manual queries in STRING[20]. The STRING database integrates both experimentally validated interactions and computational predictions based on co-expression,

literature mining, and curated databases[20]. Then, data from The Cancer Genome Atlas (TCGA) Pan-Cancer study were analyzed using the SmulTCan app to investigate gene expression patterns and correlations with survival in cancer patients[21]. This analysis revealed a candidate prognostic gene set in the tumor microenvironment of kidney renal papillary cell carcinoma (KIRP).

Materials and Methods

KSPG Interactome Discovery

Selected KSPGs were manually entered into the STRING search query. Network display options were adjusted to reveal only high-confidence interactions, defined as the "physical subnetwork". A medium-confidence level was selected, and the first and second shells of the interactome were included in the network. Functional enrichment analysis was performed using STRING's built-in functions[20].

Statistical Analysis

The KSPG interactome gene set was queried in the Gene Set Cancer Analysis Platform (https://guolab.wchscu.cn/GSCA/#/), to evaluate the correlation of the KSPG interactome expression with drug sensitivity[22, 23].

Multivariate Prognostic Modeling

Gene expression and survival data of the TCGA Pan-Cancer study was retrieved from Xena Browser and uploaded to the SmulTCan app[21,24,25]. Univariate Cox proportional hazards (CPH) regression analyses and Glmnet analyses were performed for each cancer type[21]. Genes with p-values less than 0.05 were considered statistically significant. Forest plots were generated using Python (v3.10) with the matplotlib and seaborn libraries to visualize hazard ratios, 95% confidence intervals, and corresponding p-values.

Ethical Aspects of the Research

This research was conducted in accordance with the principles of the Declaration of Helsinki. Datasets available in public repositories were used. Therefore there is no ethical consern related to the work reported here.

Results

KSPG Interactome Discovery and Functional Enrichment Analysis

To compile a list of proteins interacting with KSPGs, the STRING database was queried for the four major KSPG core proteins: LUM, keratocan (KERA), fibromodulin (FMOD), and osteoglycin (OSG/mimecan). Experimentally validated interactors, as well as first- and second-shell interaction partners, were compiled (Table 1). LUM exhibited the largest interaction network, with experimentally validated interactors: zinc finger protein 488 (ZNF488), cystathionine gamma-lyase (CTH), and matrix metalloproteinase 14 (MMP14). The KERA subnetwork included direct physical interactors bleomycin hydrolase (BLMH), nudix hydrolase 12 (NUDT12), and EP300 interacting inhibitor of differentiation 2 (EID2). FMOD showed high-confidence binding to transforming growth factor-beta (TGFB) 1, and this binding extended to include other TGFB isoforms and matrix proteins. OSG was experimentally linked to collagen type XIV alpha 1

Table 1. Summary of KSPG interactome based on STRING analysis

KSPG name	Experimentally determined interactors	STRING network-1st shell	STRING network-2nd shell
Lumican	ZNF488, CTH, MMP14	BGN, COL1A1, COL1A2, COL3A1, COL5A1, COL5A2, COL5A3, COL11A1, COL11A2, COL16A1, FN1, TGFB1, TGFB2, ITGA2, ITGB1	ALB, DCN, FMOD, FASLG, BGN, POSTN
Keratocan	BLMH, NUDT12, EID2	BLMH, NUDT12, EID2, CXCL1, LUM	ACKR1, CXCL3, DPT, COL5A3, COL11A2, NMNAT1, NMNAT2, ENPP1, ENPP3, ZNF488
Fibromodulin	TGFB1	TGFB1, TGFB2, TGFB3, DCN, LUM, BGN, FN1, HSPG2, ACAN, ELN	–
Osteoglycin	COL14A1, SRC, HLA-DRB1, CNPY3	COL14A1, SRC, HLA-DRB1, CNPY3, OMD, VCAN, LPAR3	CDCP1, ADRB3, AFAP1, TNXB, NCAN, ACAN, TPBG, MLANA, CILP, CHRM1

Each row lists a core KSPG and its experimentally validated interactors, and the proteins found in its first and second-shell STRING network
Annotations are provided in Table S1

chain (COL14A1), proto-oncogene tyrosine-protein kinase Src (SRC), major histocompatibility complex, class II, DR beta 1 (HLA-DRB1), and canopy FGF signaling regulator 3 (CNPY3).

The identified KSPG interactome contains 56 proteins and forms a highly interconnected network that is visualized and functionally annotated using the STRING database (Figure 1A, Table S1). GO biological process enrichment analysis highlighted ECM organization as the most significantly enriched term [false discovery rate (FDR) <1e-23], with more than 20 genes contributing (Figure 1B). Closely related terms included extracellular structure organization, collagen fibril organization, and connective tissue development. Additional enrichment in cell adhesion and in positive regulation of smooth muscle cell proliferation was detected. Proteoglycans in cancer was the top KEGG pathway (FDR <1e-9), followed by focal adhesion, ECM-receptor interaction, phosphoinositide 3-kinase-AKT (PI3K-AKT) signaling, and TGFB signaling pathways, Interestingly, terms related to nicotinamide adenine dinucleotide (NADH) metabolism were enriched in both GO and KEGG pathway analyses.

Prognostic Associations of the KSPG Interactome Across Cancer Types

Correlation between mRNA expression profiles and drug sensitivity was assessed using the genomics of drug sensitivity in cancer portal and the Pan-Cancer dataset[23]. The top eight drugs with significant correlations with KSPG gene expression patterns and drug response were identified (Figure 2A, Table S2). Response to BHG712, a selective EphB4 inhibitor, was found to be strongly correlated with the expression levels of most genes in the KSPG interactome. Among these genes, *NMNAT2*, *ZNF488*, *TPBG*, *ITGB1*, *ITGB2*, *CDCP1*, and *FN1* were positively correlated, while *ENPP3*, *BLMH*, *HLA-DRB1*, and *CNPY3* were negatively correlated. Response to BRAF and MEK inhibitors (Dabrafenib, PLX4720, SB590885, CI-1040, and RDEA119) were negatively correlated with expression levels of *LUM*, *FN1*, *CXCL1*, *COL16A1*, *MMP14*, *HSPG2* and *MLANA*. Expression levels of the interactome genes in skin cutaneous melanoma (SKCM), uterine endometrial cancer (UCEC), and KIRP are plotted in Figure 2B. SKCM was chosen because the KSPG interactome gene set alteration frequency is highest (data not shown); UCEC and KIRP were chosen to complement CPH analyses explained below. Most KSPG genes are overexpressed in cancer patients. Among the core KSPGs, *LUM*, *FMOD*, and *biglycan* were highly overexpressed; *decorin*, *osteomodulin*, and *aggrecan* were moderately overexpressed; *KERA* was not overexpressed in the analyzed cancer types.

Next, CPH regression and Glmnet analyses were using the SmulTCan app across TCGA cancer types to define prognostic gene sets. The CPH analysis revealed specific multi-gene signatures in various cancer types, with uterine corpus endometrial carcinoma (UCEC) and KIRP exhibiting the most extensive lists of genes (Table S3). CPH predicted a highly significant multigene prognostic signature comprising 18 genes associated with survival in KIRP ($p<0.05$). These genes included the proteoglycans *LUM*, *FMOD*, and *KERA*, which are significantly associated with poor prognosis (Figure 3A). The metabolic genes *NMNAT1*, *NUDT12*, and *ENPP1* were significantly associated with a better prognosis. Other genes with prognostic value were *CXCL1*, *MLANA*, *CHRM1*, *ZNF488*, *BLMH*, and *CTH*. The model effectively stratified patients into high- and low-risk groups, which exhibited a significant difference in overall survival on the Kaplan-Meier plot (Figure 3B).

Next, the Glmnet method was applied to select the best gene subset; however, a fitted model was obtained only for KIRP. The KIRP prognostic signature included *KERA*, *ZNF488* and *CTH* with positive coefficients, suggesting their expression is associated with increased risk and poorer survival (Figure 3C). In contrast, *nicotinamide mononucleotide adenylyltransferase 1 (NMNAT1)* was assigned a negative coefficient, consistent with a protective effect. Additional low-magnitude positive coefficients were retained for *COL11A2*, *COL1A1*, and *COL5A2*. Receiver operating characteristic (ROC) curve analysis yielded an area under the curve (AUC) of 78.9%, indicating strong predictive accuracy (Figure 3D).

Discussion

KSPGs are found both extracellularly and intracellularly, with capacity to interact with a large number and variety of proteins[4]. Nevertheless, the experimental data on KSPG-interacting proteins are fragmented, and a definition of the KSPG interactome is needed. To address this gap, we used the STRING database to construct a physical interaction network centered on the KSPG core proteins KERA, LUM, FMOD, and OGN and identified a 56-protein KSPG interactome. The intereactome not only included ECM related proteins but also major cancer related pathways including PI3K-AKT signaling. Enrichment of focal adhesion, integrin binding, and TGFB binding may be linked to the structural remodeling required for cancer metastasis.

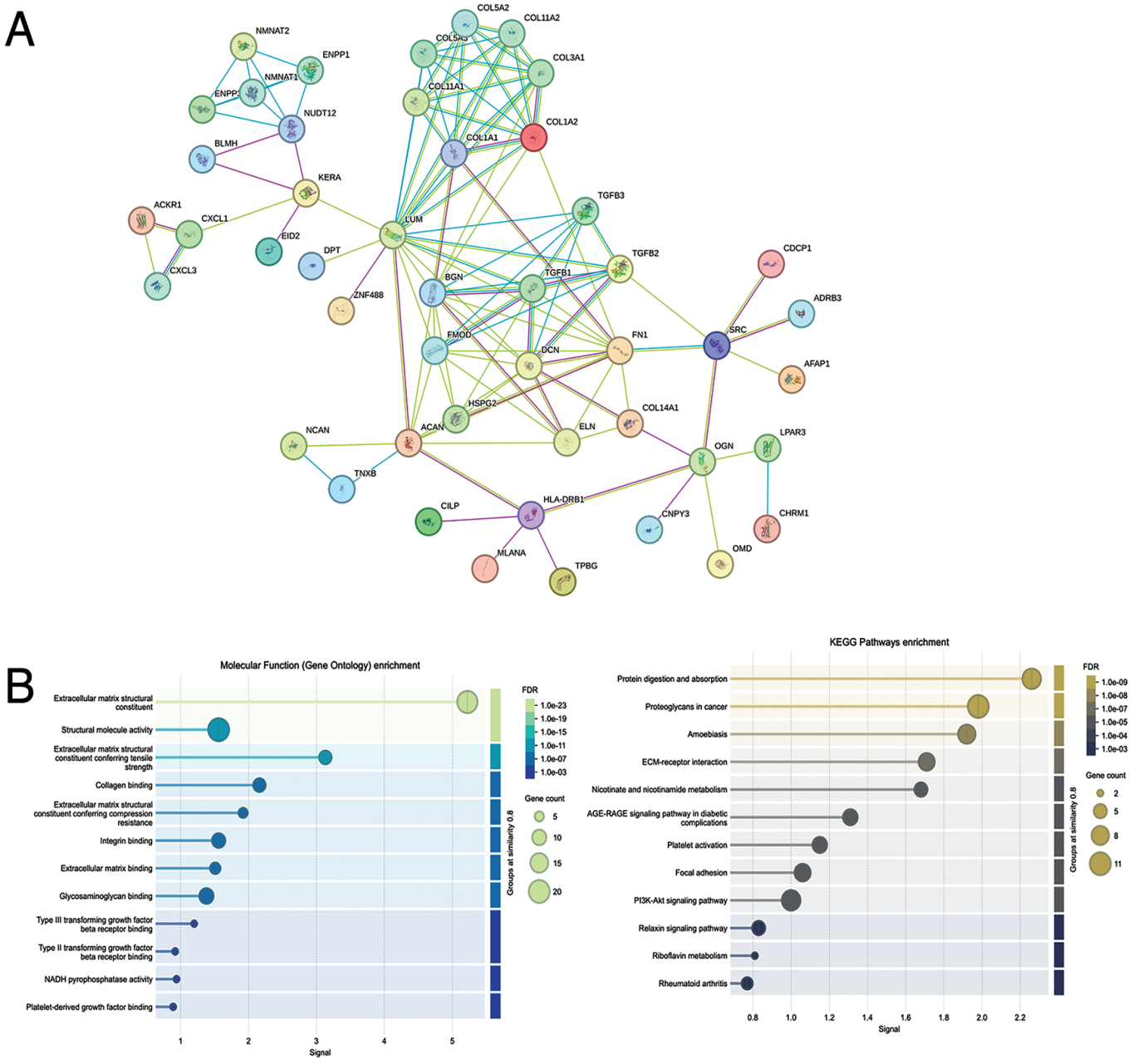


Figure 1. STRING-derived interaction network of the KSPG interactome

A) The network includes 56 proteins identified through STRING as part of the keratan sulfate proteoglycan (KSPG) interactome. Nodes represent proteins; edge colors indicate the type of supporting evidence: purple (experimentally determined interactions); blue (curated database associations); and green (text mining). **B)** GO biological process enrichment of the KSPG interactome. Bubble plot showing the top enriched GO biological processes. The x-axis represents gene ratio, while the y-axis lists the GO terms. Bubble size corresponds to the number of genes involved in each term; color intensity reflects statistical significance (FDR)

KSPG: Keratan sulfate proteoglycan, GO: Gene ontology, KEGG: Kyoto Encyclopedia of Genes and Genomes, FDR: False discovery rate

The composition of the ECM is known to modulate drug sensitivity in cancer[26]. Correlations between gene expression and BRAF-MEK inhibitors suggest a potential modulation of drug sensitivity via the KSPG interactome. The prognostic potential of KSPG interactome genes was analyzed via univariate CPH, and best subsets were selected using Glmnet multivariate penalized Cox regression. This approach enables simultaneous variable selection and regularization, identifying a minimal gene set with optimal predictive performance while excluding redundant or non-informative features[21]. The CPH approach identified numerous genes with significant associations with survival across tumor types, including uterine corpus endometrial carcinoma UCEC and KIRP. While this method is straightforward and sensitive to any gene-outcome association, it does not account for correlations among genes, potentially overestimating the

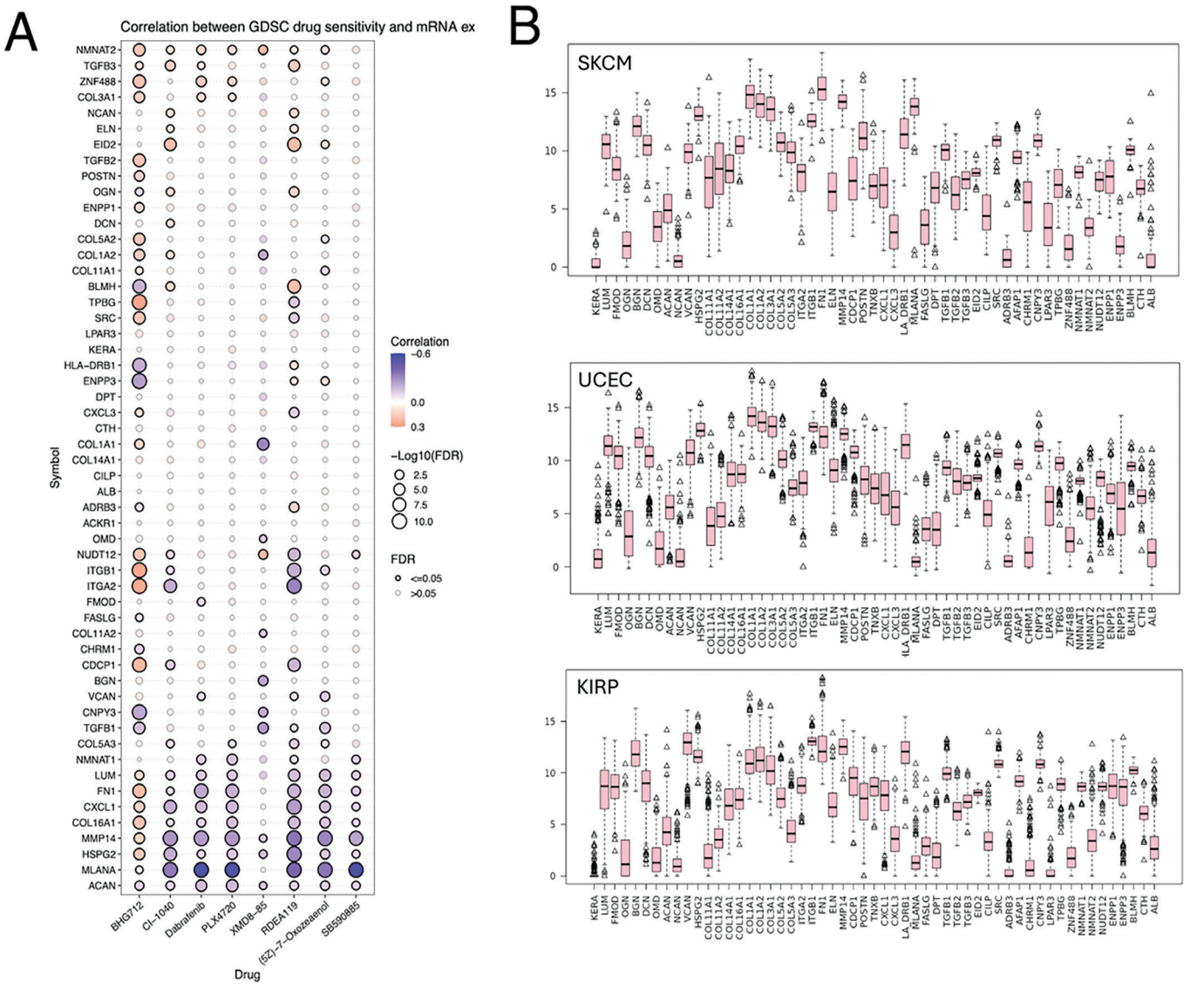


Figure 2. KSPG interactome expression in cancer

A) Drug sensitivity correlation figure summarizes the correlation between gene expression and the sensitivity of GDSC drugs (top 30) in Pan-Cancer (full list Table S2). B) Gene expression distribution of KSPG interactome genes in SKCM, UCEC and KIRP

GDSC: Genomics of drug sensitivity in cancer, KSPG: Keratan sulfate proteoglycan, SKCM: Skin cutenous melanoma, UCEC: Uterine endometrial cancer, KIRP: Kidney renal papillary cell carcinoma

number of relevant variables. By contrast, Glmnet applies lasso regularization to select a parsimonious multigene model with independent predictive contributions and improved generalizability[21]. However, this increased stringency can lead to failure to retain variables in the final model when predictors are highly correlated or when the incremental predictive value is modest relative to noise. This scenario occurred in UCEC, where no multigene signature was selected despite numerous univariate associations. However, in KIRP, Glmnet successfully identified a multigene signature consisting of *KERA*, *NMNAT1*, *ZNF488*, *COL5A2*, *COL1A1*, *COL11A2*, and *CTH*. The transcription factor *ZNF488* is functionally linked to *LUM* and *COL5A3* in the STRING network. Recently, the roles of *ZNF488* in cancer progression

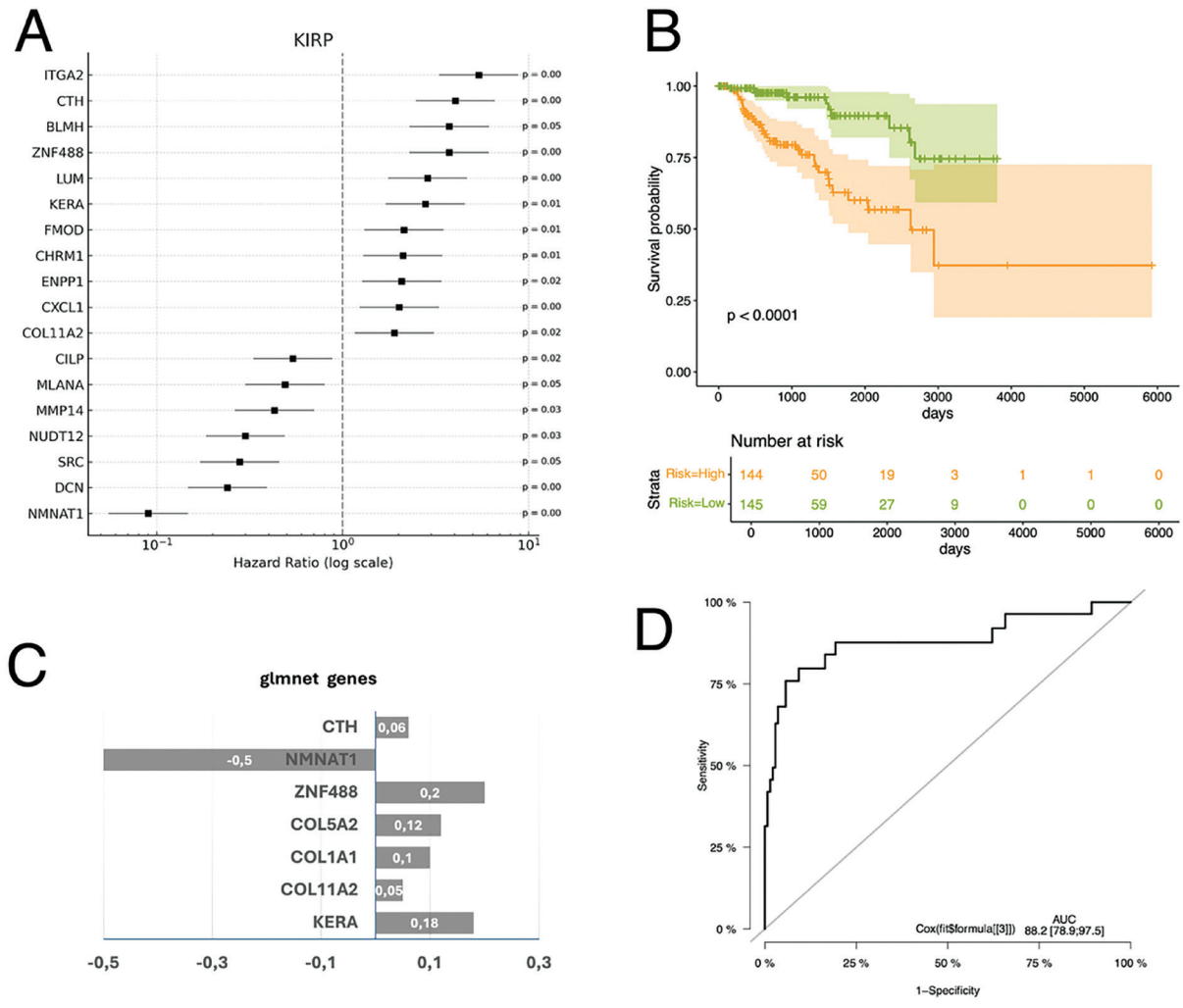


Figure 3. A prognostic gene signature is discovered in KIRP

A) Gene signature recovered with Cox modeling in KIRP. HRs and their 95% confidence intervals are displayed on a logarithmic scale. Genes with HR >1 are associated with poorer prognosis, while those with HR <1 are associated with improved survival. P-values are shown to the right of each gene. The gene order reflects increasing HR for visual clarity. **B)** Kaplan-Meier survival curves comparing overall survival between high-risk and low-risk groups. The high-risk group (orange) exhibited significantly poorer survival than the low-risk group (green) (log-rank p<0.0001). Shaded areas indicate 95% confidence intervals. Tick marks on the curves denote censored observations. The table below the plot shows the number of subjects at risk at each time point (in days). **C)** Prognostic gene signature discovered with Glmnet best subset selection. Coeficients are displayed on bar graph. **D)** ROC curve illustrating the predictive performance of the Glmnet model in distinguishing between the specified outcome groups. The AUC was 78.9%, indicating strong predictive accuracy. The diagonal dashed line represents the performance of a random classifier (AUC: 0.5)

KIRP: Kidney renal papillary cell carcinoma, HR: Hazard ratio, ROC: Receiver operating characteristic, AUC: Area under the curve

and resistance have been indicated, although the mechanism and its link to KSPGs have not been studied [27,28]. *KERA* had a positive coefficient, indicating that higher expression is associated with worse survival. Interestingly, *KERA* expression in KIRP is markedly lower than in other cancers and than that of other genes in the KSPG interactome. However, high variability in gene expression among patients was observed, suggesting that *KERA* expression must be kept very low for better patient prognosis.

KERA encodes a small leucine-rich proteoglycan involved in collagen organization and ECM architecture, but its role in cancer has been described in only a few studies[29,30]. *NMNAT1* had the most negative coefficient in the penalized model, suggesting a protective association in KIRP. Enrichment of the nicotinate and nicotinamide metabolism pathways suggested involvement of the KSPG network in regulating energy metabolism, which is further supported by the prognostic value of *NMNAT1* identified in this study. The NMNAT1 enzyme catalyzes a critical step in NAD^+ biosynthesis, potentially supporting metabolic homeostasis and resilience to oxidative stress. Notably, NMNAT1 is part of the KERA-centered STRING subnetwork together with NMNAT2, ENPP1, ENPP3, and NUDT12. These proteins regulate nicotinate and nicotinamide metabolism. Human NMNAT enzymes catalyze the formation of NAD^+ from adenosine triphosphate and nicotinamide mononucleotide (NMN) and the cleavage of NADH in the reverse reaction[31]. NUDT12 is a peroxisomal NADH pyrophosphatase and an mRNA-decapping enzyme that specifically removes the NAD cap from a subset of transcripts, thereby generating NMN and promoting mRNA decay, particularly under nutrient stress conditions[32,33]. It also hydrolyzes free NAD(P)H and diadenosine diphosphates to regulate peroxisomal nicotinamide nucleotide pools required for oxidative metabolism. Together, these findings suggest a potentially relevant ECM-metabolism regulatory axis, linking KERA and its interactors to energy metabolism and nucleotide turnover within the tumor microenvironment.

This analysis leveraged the SmulTCan web application, which integrates both univariate and multivariate survival modeling in an accessible interface[21]. SmulTCan proved particularly useful for rapidly evaluating prognostic associations without requiring advanced bioinformatics expertise or command-line tools, making it well suited for translational researchers and clinicians who wish to explore publicly available cancer datasets reproducibly. The methodology followed here illustrates that integrating curated protein interaction networks with public genomics data can reveal cancer-type-specific gene modules associated with prognosis and drug response.

Importantly, a KSPG interactome was identified and proposed in this study. *KERA* and *NMNAT1* were identified as having high prognostic value in KIRP, a finding not previously reported. Future studies are needed to functionally validate the contributions of the candidate genes identified here and to determine whether they can inform stratified therapeutic strategies or combination treatments that target ECM remodeling and tumor metabolism.

Ethics

Ethics Committee Approval: Datasets available in public repositories were used. Therefore there is no ethical consern related to the work reported here.

Informed Consent: This study does not involve any experiments conducted with patients, or patient materials, animals or animal materials. The researh was conducted with publicly available datasets.

Acknowledgements

Author thanks Prof. MD, Özlen Konu for critical reading of the manuscript and help with the SmulTCan App analyses.

Footnotes

Authorship Contributions

Concept: G.Ç.A., Design: G.Ç.A., Data Collection or Processing: G.Ç.A., Analysis or Interpretation: G.Ç.A., Literature Search: G.Ç.A., Writing: G.Ç.A.

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

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

References

1. Wang M, Yang Y, Han L, et al. Effect of three-dimensional ECM stiffness on cancer cell migration through regulating cell volume homeostasis. Biochem Biophys Res Commun. 2020;528:459-65.
2. Deng B, Zhao Z, Kong W, Han C, Shen X, Zhou C. Biological role of matrix stiffness in tumor growth and treatment. J Transl Med. 2022;20:540.
3. Hardingham T. Proteoglycans and glycosaminoglycans. In: Seibel MJ, Bilezikian JP, Robins SP, (editors). Dynamics of bone and cartilage metabolism. 2nd ed. Elseiver; 2006:85-98.
4. Caterson B, Melrose J. Keratan sulfate, a complex glycosaminoglycan with unique functional capability. Glycobiology. 2018;28:182-206.

5. Wei J, Hu M, Huang K, Lin S, Du H. Roles of proteoglycans and glycosaminoglycans in cancer development and progression. Int J Mol Sci. 2020;21:5983.
6. Appunni S, Anand V, Khandelwal M, Gupta N, Rubens M, Sharma A. Small leucine rich proteoglycans (decorin, biglycan and lumican) in cancer. Clin Chim Acta. 2019;491:1-7.
7. Grant RC, Denroche RE, Borgida A, et al. Exome-wide association study of pancreatic cancer risk. Gastroenterology. 2018;154:719-22.
8. Xu F, Xu H, Li Z, et al. Glycolysis-based genes are potential biomarkers in thyroid cancer. Front Oncol. 2021;11:534838.
9. Liu G, Lu Y, Gao D, Huang Z, Ma L. Identification of an energy metabolism-related six-gene signature for distinguishing and forecasting the prognosis of low-grade gliomas. Ann Transl Med. 2023;11:146.
10. Melrose J. Keratan sulfate (KS)-proteoglycans and neuronal regulation in health and disease: the importance of KS-glycodynamics and interactive capability with neuroregulatory ligands. J Neurochem. 2019;149:170-94.
11. Basol M, Ersoz-Gulseven E, Ozaktas H, Kalyoncu S, Utine CA, Cakan-Akdogan G. Loss of carbohydrate sulfotransferase 6 function leads to macular corneal dystrophy phenotypes and skeletal defects in zebrafish. FEBS J. 2025;292:373-90.
12. Abbruzzese C, Kuhn U, Molina F, Rama P, De Luca M. Novel mutations in the *CHST6* gene causing macular corneal dystrophy. Clin Genet. 2004;65:120-5.
13. Appunni S, Rubens M, Ramamoorthy V, et al. Lumican, pro-tumorigenic or anti-tumorigenic: a conundrum. Clin Chim Acta. 2021;514:1-7.
14. Seya T, Tanaka N, Shinji S, et al. Lumican expression in advanced colorectal cancer with nodal metastasis correlates with poor prognosis. Oncol Rep. 2006;16:1225-30.
15. Matsuda Y, Yamamoto T, Kudo M, et al. Expression and roles of lumican in lung adenocarcinoma and squamous cell carcinoma. Int J Oncol. 2008;33:1177-85.
16. Xu W, Chen S, Jiang Q, et al. LUM as a novel prognostic marker and its correlation with immune infiltration in gastric cancer: a study based on immunohistochemical analysis and bioinformatics. BMC Gastroenterol. 2023;23:455.
17. Chen X, Li X, Hu X, et al. LUM expression and its prognostic significance in gastric cancer. Front Oncol. 2020;10:605.
18. Karamanou K, Franchi M, Proult I, Rivet R, Vynios D, Brezillon S. Lumican inhibits *in vivo* melanoma metastasis by altering matrix-effectors and invadopodia markers. Cells. 2021;10:841.
19. Chasan S, Hesse E, Atallah P, et al. Sulfation of glycosaminoglycan hydrogels instructs cell fate and chondral versus endochondral lineage decision of skeletal stem cells *in vivo*. Adv Funct Mater. 2021;32:2109176.
20. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43:D447-52.
21. Ozhan A, Tombaz M, Konu O. SmulTCan: a shiny application for multivariable survival analysis of TCGA data with gene sets. Comput Biol Med. 2021;137:104793.
22. Liu CJ, Hu FF, Xia MX, Han L, Zhang Q, Guo AY. GSCALite: a web server for gene set cancer analysis. Bioinformatics. 2018;34:3771-2.
23. Yang W, Soares J, Greninger P, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41:D955-61.
24. Goldman MJ, Craft B, Hastie M, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38:675-8.
25. Hutter C, Zenklusen JC. The Cancer Genome Atlas: creating lasting value beyond its data. Cell. 2018;173:283-5.
26. Huang J, Zhang L, Wan D, et al. Extracellular matrix and its therapeutic potential for cancer treatment. Signal Transduct Target Ther. 2021;6:153.
27. Weng K, Li L, Zhou H. Transcription factor ZNF488 accelerates cervical cancer progression through regulating the MEK/ERK signaling pathway. Histol Histopathol. 2023;38:1381-90.
28. Chen D, Le SB, Manektalia H, et al. The EP3-ZNF488 axis promotes self-renewal of glioma stem-like cells to induce resistance to tumor treating fields. Cancer Res. 2025;85:360-77.
29. Chakravarti S. Focus on molecules: keratocan (KERA). Exp Eye Res. 2006;82:183-4.
30. Gao H, Qian R, Ren Q, et al. The upregulation of keratocan promotes the progression of human pancreatic cancer. Mol Cell Toxicol. 2024;20:271-80.
31. Sorci L, Scotti S, Peterlli R, et al. Initial-rate kinetics of human NMN-adenylyltransferases: substrate and metal ion specificity, inhibition by products and multisubstrate analogues, and isozyme contributions to NAD+ biosynthesis. Biochemistry. 2007;46:4912-22.
32. Grudzien-Nogalska E, Wu Y, Jiao X, et al. Structural and mechanistic basis of mammalian Nudt12 RNA deNADding. Nat Chem Biol. 2019;15:575-82.
33. Carreras-Puigvert J, Zitnik M, Jemth AS, et al. A comprehensive structural, biochemical and biological profiling of the human NUDIX hydrolase family. Nat Commun. 2017;8:1541.

Supplementary Table 1. https://d2v96fxpocvxx.cloudfront.net/cf9d60d6-523c-458a-a2e6-78728d3ffbb0/documents/Table%20S1_KSPGinteractome_annotation.xlsx

Supplementary Table 2. https://d2v96fxpocvxx.cloudfront.net/cf9d60d6-523c-458a-a2e6-78728d3ffbb0/documents/Table%20S3_CPH%20coefficients.xlsx

Supplementary Table 3. https://d2v96fxpocvxx.cloudfront.net/cf9d60d6-523c-458a-a2e6-78728d3ffbb0/documents/Table%20S2_GdscIC50_drug%20sensitivity_AndExprTable.xlsx

DOI: 10.4274/anatoljmed.2025.33154
Anatol J Gen Med Res 2025;35(3):330-338

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

The Relationship Between Thyroid Function and Menstrual-related Migraine in Women: A Cross-sectional Analysis

Kadınlarda Tiroid Fonksiyonu ve Menstrüasyon Dönemiyle İlişkili Migren Arasındaki İlişki: Kesitsel Bir Analiz

Emiş Cansu Yaka[1], Dinçer Atila[2]

[1]İzmir City Hospital, Clinic of Neurology, İzmir, Türkiye
[2]Menemen No. 1 Family Health Center, Clinic of Family Medicine, İzmir, Türkiye

Cite as: Yaka EC, Atila D. The relationship between thyroid function and menstrual-related migraine in women: a cross-sectional analysis. Anatol J Gen Med Res. 2025;35(3):330-338

Abstract

Objective: Migraine is a common and disabling neurological disorder, particularly in women, and its menstrual subtypes-menstrual migraine (MM) and menstruation-related migraine (MRM)-highlight the influence of hormonal fluctuations. Estrogen withdrawal is a recognized trigger, but the influence of thyroid function remains unclear. The aim of this research was to investigate the association between thyroid hormone levels and migraine phenotypes in women.

Methods: This cross-sectional study included 175 women categorized by migraine type: chronic migraine (n=54), MRM [n=56; including 12 with pure MM (PMM)], and a control group (n=65). Participants with known thyroid disease were excluded. Serum thyroid-stimulating hormone (TSH), triiodothyronine (T3), and thyroxine (T4) were measured, and hematologic indices were examined. Migraine severity and related disability were measured using the visual analog scale and the migraine disability assessment (MIDAS). Group comparisons were performed using one-way analysis of variance, Kruskal-Wallis, and chi-square tests.

Results: TSH, T3, T4, and T3/T4 ratio did not differ significantly between groups (all $p>0.05$). The PMM and MRM subgroups also showed similar thyroid values. However, patients with chronic migraine had significantly higher MIDAS scores and a higher frequency of prophylactic drug use (both $p<0.001$). Hypertension was more common among patients with migraine, but the difference was not statistically significant.

Conclusion: There were no differences in thyroid hormone levels between migraine types and controls, indicating that thyroid changes are not a key factor in migraine among women with normal thyroid function. Further longitudinal research is needed to explore potential subclinical or autoimmune thyroid contributions to hormonally mediated migraine.

Keywords: Migraine, menstruation-related migraine, thyroid function tests, women's health

Öz

Amaç: Migren, özellikle kadınlarda yaygın ve sakatlayıcı bir nörolojik hastalıktır ve menstrüel alt tipleri-menstrüel migren (MM) ve menstrüasyonla ilişkili migren (MİM)-hormonal dalgalanmaların etkisini vurgular. Östrojen yoksunluğu bilinen bir tetikleyicidir, ancak tiroid fonksiyonunun etkisi henüz net değildir. Bu araştırmanın amacı, kadınlarda tiroid hormon seviyeleri ile migren fenotipleri arasındaki ilişkiyi araştırmaktır.

Address for Correspondence/Yazışma Adresi: Emiş Cansu Yaka, MD, İzmir City Hospital, Clinic of Neurology, İzmir, Türkiye
E-mail: emiscansu@gmail.com
ORCID ID: orcid.org/emiscansu@gmail.com

Received/Geliş tarihi: 03.11.2025
Accepted/Kabul tarihi: 10.11.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

Yöntem: Bu kesitsel çalışmaya, migren tipine göre kategorize edilmiş 175 kadın dahil edilmiştir: kronik migren (n=54), MİM [n=56; 12 saf MM (SMM) dahil] ve kontrol (n=65). Bilinen tiroid hastalığı olan katılımcılar çalışma dışı bırakılmıştır. Serum tiroid uyarıcı hormon (TSH), triiyodotironin (T3), tiroksin (T4) ve hematolojik indeksler incelenmiştir. Migrenle ilişkili şiddet ve engellilik, görsel analog skala ve migren engellilik değerlendirmesi (MIDAS) kullanılarak ölçüldü. Grup karşılaştırmaları tek yönlü varyans analizi, Kruskal-Wallis ve ki-kare testleri ile yapıldı.

Bulgular: TSH, T3, T4 ve T3/T4 oranları gruplar arasında anlamlı farklılık göstermedi (tümü p>0,05). SMM ve MİM alt grupları da benzer tiroid değerleri gösterdi. Ancak, kronik migrenli hastalarda anlamlı olarak daha yüksek MIDAS skorları ve daha sık profilaktik ilaç kullanımı vardı (her ikisi de p<0,001). Hipertansiyon migren hastalarında daha yaygındı, ancak istatistiksel olarak anlamlı değildi.

Sonuç: Migren tipleri veya kontroller arasında tiroid hormonu seviyelerinde farklılık bulunmaması, tiroid değişiminin normal tiroid fonksiyonuna sahip kadınlar için önemli bir migren faktörü olmadığını göstermektedir. Hormonal aracılı migrene olası subklinik veya otoimmün tiroid katkılarını araştırmak için daha fazla uzunlamasına araştırmaya ihtiyaç vardır.

Anahtar Kelimeler: Migren, menstrüasyon ile ilişkili migren, tiroid fonksiyon testleri, kadın sağlığı

Introduction

Migraine is a prevalent and disabling primary headache disorder characterized by recurrent episodes of moderate-to-severe headaches, often accompanied by nausea, photophobia, and phonophobia[1]. Approximately 12% of people worldwide are affected, and it occurs about three times more frequently in women than in men, particularly during women's reproductive years[2]. Among its subtypes, menstrual migraine (MM) and menstruation-related migraine (MRM) are especially challenging, as attacks occur in close temporal relation to the menstrual cycle-typically from two days before to three days after onset of menstruation[3]. Estrogen withdrawal is thought to be the key hormonal trigger; however, the underlying mechanisms remain incompletely understood[4].

Other than sex hormones, there is growing interest in how the endocrine system, especially the thyroid, affects migraine pathophysiology. Thyroid hormones-triiodothyronine (T3) and thyroxine (T4)-are key regulators of metabolism, vascular tone, and neurochemical balance. Altered levels of these hormones can change neuronal excitability, cerebral metabolism, and pain processing, potentially increasing migraine risk[5]. Epidemiological studies indicate that migraine and thyroid dysfunction co-occur more often than expected by chance, suggesting shared pathophysiological pathways or potential causal interactions[5-7]. Importantly, both conditions appear more frequently in women, indicating that estrogen might modulate thyroid activity, which could potentially affect migraine severity[8].

Several studies have explored this association. A systematic review by Michalik et al.[5] found that individuals with hypothyroidism frequently report headaches, supporting the notion of a hormonal contribution to migraine genesis. A longitudinal study suggested that headache disorders may precede new-onset hypothyroidism, implying a possible bidirectional relationship[9]. Likewise, Starikova et al.[10] discovered that lower thyroid-stimulating hormone (TSH) levels correlated with more severe migraine symptoms, although other studies showed no significant differences in TSH or free T4 between migraine and control groups[11]. The inconsistency emphasizes the need for greater clarity, particularly within hormonally sensitive populations.

The association between thyroid disorders and migraines in women with MRM or pure MM (PMM) remains poorly understood. Fluctuations in estrogen levels during the menstrual cycle may affect thyroid hormone activity, potentially altering headache timing and intensity[7,12]. However, it is difficult to draw definitive conclusions from the currently incomplete and contradictory data. The research indicated that women of reproductive age with migraines may particularly benefit from thyroid screening, given the correlation between increased headache frequency and subclinical hypothyroidism characterized by elevated TSH with normal T4[8,9].

Given these uncertainties, the present study aimed to comprehensively assess thyroid hormone profiles (T3, T4, and TSH) and related hematologic indices among women with chronic migraine, MRM, or PMM, and healthy controls. This research used clinical, endocrine, and hematologic parameters to investigate whether variations in thyroid function contribute to migraine burden, particularly in hormonally influenced subtypes.

Materials and Methods

Study Design and Participants

This cross-sectional study involved 190 premenopausal women who visited the neurology clinic between April and June 2025. The participants were divided into three groups: chronic migraine (n=54), MRM (n=56), and healthy controls (n=65). The MRM group included 12 patients diagnosed with PMM. Definitive diagnoses of migraine were made according to the International Classification of Headache Disorders, 3rd edition[1], based on a clinical interview conducted by a neurologist specializing in headache disorders. Participants with a history of thyroid disease (n=15; 7.9%) were excluded to minimize potential endocrine confounding factors. Consequently, the final analyses included 175 participants.

Data Collection and Variables

Structured forms were used to collect clinical and demographic data. Age, menstrual status, presence of hypertension (HT), and migraine characteristics were documented. The migraine disability assessment (MIDAS) questionnaire[13] and visual analog scale (VAS) were used to assess headache disability and pain severity, respectively. The number of prophylactic drugs in current use was also recorded.

Blood samples were collected from the veins after an 8-hour fast. An automated hematology analyzer (Sysmex Corporation, Kobe, Japan) was used to measure hematological parameters, including hemoglobin (Hb), hematocrit (Hct), mean corpuscular volume (MCV), red cell distribution width (RDW), platelet distribution width (PDW), platelet count (PLT), and neutrophil and lymphocyte counts. The neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, Hb/RDW ratio, PDW/PLT ratio, and RDW/MCV ratio were calculated. Thyroid hormone measurements included TSH, T3, and T4. All assays were performed using standardized chemiluminescent immunoassay techniques at the same laboratory.

Ethical Approval and Informed Consent

This research adhered to the Declaration of Helsinki and was ethically approved by the University of Health Sciences Türkiye, İzmir City Hospital Ethics Committee (approval no: 2025/52, date: 19.03.2025). Before joining the study, all participants provided written informed consent.

Data Availability Statement

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

Declaration Regarding the Use of Artificial Intelligence and Artificial Intelligence-assisted Technologies

During the preparation of this work, the author used Grammarly to assist with linguistic editing and improve phrasing. The final content was reviewed and approved by the author, who takes full responsibility for the manuscript's scientific integrity.

Statistical Analysis

All statistical analyses were conducted using R version 4.4.2. Hypothesis testing was performed using the rstatix package; flextable was used for tables; and R6 was used for modular data and workflow automation. The Shapiro-Wilk test was used to assess the normality of continuous variables. Normally distributed variables were expressed as mean±standard deviation (SD), whereas non-normally distributed variables were summarized as median (minimum-maximum).

The independent samples t-test or Mann-Whitney U test was used to compare two independent groups, as appropriate. When comparing three or more groups, one-way analysis of variance or the Kruskal-Wallis test was applied, followed by Bonferroni-corrected post-hoc tests. Categorical variables were compared using the chi-square test or Fisher's exact test when necessary, and were presented as counts and percentages. Statistical significance was determined by a two-tailed p-value below 0.05.

Results

A total of 190 participants were initially enrolled in the study. Fifteen participants were excluded owing to known thyroid disease (9.1% in the chronic migraine group, 8.9% in the MRM group, 16.7% in the PMM subgroup, and 5.8% in the control group). Hence, the final analysis comprised 175 participants: 56 in the MRM group, 54 in the chronic migraine group, and 65 in the control group (Table 1). With a mean age of 36.4±8.8 years, 113 participants (64.6%) were aged 40 years or younger. HT was present in 11 individuals (6.3%). The overall mean Hb was 12.6±1.4 g/dL, and the mean Hct was 37.6±3.7%. The median values for TSH, T3, and T4 were 1.7 mIU/L (0.1-18.9), 3.4 ng/dL (2.4-4.6), and 0.8 ng/dL (0.5-2.0), respectively.

As shown in Table 2, demographic and thyroid-related variables did not differ significantly between the chronic migraine (including MRM) group and the control group. Mean age was similar between groups (mean±SD: 36.5±8.1 vs. 36.2±9.8 years; p=0.874). In the chronic migraine group, 69.1% of participants were ≤40 years old, compared with 56.9% of controls (p=0.144). No significant group differences were observed in T3, T4, TSH, or the T3/T4 ratio (all p>0.05).

Table 1. Baseline demographic, hematological, and thyroid parameters of the participants without thyroid disease population

	Overall, n (%) n=175		**Overall, n (%) n=175**
Age (years)*	36.4±8.8	**RDW (%)***	14.4±1.9
Age		**Neutrophil (x10³/µL)***	4.4 (1.6-16.2)
≤40 years old	113 (64.6)	**Lymphocyte (x10³/µL)***	2.1 (0.3-4.3)
>40 years old	62 (35)	**Platelets (x10³/µL)***	271 (140-516)
Study group		**PDW (%)***	16.8±0.5
MRM	56 (32.0)	**T3 (ng/dL)***	3.4 (2.4-4.6)
Chronic migraine	54 (30.9)	**T4 (ng/dL)***	0.8 (0.5-2.0)
Control	65 (37.1)	**TSH (mIU/L)***	1.7 (0.1-18.9)
MIDAS score*	18 (5-46)	**T3/T4 ratio***	4.1 (0.0-7.3)
VAS score*	8 (7-9)	**RDW/MCV ratio***	0.2 (0.1-0.5)
Number of prophylactic drugs*	0 (0-6)	**NLR***	2.1 (0.6-38.7)
HT	11 (6.3)	**PLR***	130.7 (68.6-1076.7)
Hb (g/dL)*	12.6±1.4	**PDW/PLT ratio***	0.1 (0.0-0.1)
Hematocrit (%)*	37.6±3.7	**Hb/RDW ratio***	0.9 (0.4-1.2)
MCV (fL)*	83.1±9.5		

*: Numeric variables were presented as medians (minimum-maximum) or mean±standard deviation, Hb: Hemoglobin, HT: Hypertension, MCV: Mean corpuscular volume, MIDAS: Migraine disability assessment, MRM: Mensturation-related migraine, NLR: Neutrophil-to-lymphocyte ratio, RDW: Red cell distribution width, PDW: Platelet distribution width, PLT: Platelet count, PLR: Platelet-to-lymphocyte ratio, T3: Triiodothyronine, T4: Thyroxine, TSH: Thyroid-stimulating hormone, VAS: Visual analog scale

Table 2. Comparison of parameters between the MRM+chronic migraine and control groups

	MRM€+chronic migraine, n (%) n=110	**Control, n (%) n=65**	**p-value**
Age (years)*	36.5±8.1	36.2±9.8	0.874
Age			0.144
≤40 years old	76 (69.1)	37 (56.9)	
>40 years old	34 (30.9)	28 (43.1)	
MIDAS score*	18 (5-46)	NA	**NA**
VAS score*	8 (7-9)	NA	**NA**
Number of prophylactic drugs*	2 (0-6)	NA	**NA**
HT	9 (8.2)	2 (3.1)	0.215
PDW (%)*	16.8±0.5	16.8±0.6	0.861
T3 (ng/dL)*	3.4 (2.6-4.5)	3.5 (2.4-4.6)	0.199
T4 (ng/dL)*	0.8 (0.5-1.2)	0.9 (0.5-2.0)	0.152
TSH (mIU/L)*	1.7 (0.1-19.0)	1.7 (0.1-6.5)	0.914
T3/T4 ratio*	4.2 (0-7.3)	4.0 (1.6-6.1)	0.425

*: Numeric variables were presented as median (minimum-maximum) or mean±standard deviation, €: PMM and MRM patients were evaluated in the same group within MRM, PMM: Pure menstrual migraine, MRM: Menstruation-related migraine, MIDAS: Migraine disability assessment, NA: Not applicable, VAS: Visual analog scale, HT: Hypertension, PDW: Platelet distribution width, T3: Triiodothyronine, T4: Thyroxine, TSH: Thyroid-stimulating hormone

Table 3 provides a comparison of the MRM, chronic migraine, and control groups. Age distribution (p=0.231) and thyroid function parameters (all p>0.05) did not differ significantly. Nevertheless, the chronic migraine group had considerably higher MIDAS scores than the MRM group (31 vs. 13, p<0.001). Use of preventive drugs was higher in the chronic migraine group than in the MRM group (0 vs. 4; p<0.001). HT was more common in chronic migraine (13.0%) than in MRM (3.6%) or in controls (3.1%); however, this difference did not reach statistical significance (p=0.052).

A subgroup analysis comparing PMM and MRM participants (Table 4) revealed that the PMM group was significantly younger than the MRM group (mean age 29.6 vs. 36.7 years; p=0.002). The MIDAS score was significantly lower in PMM than in MRM (8.6 vs. 12.6; p=0.002). VAS scores, prophylactic medication use, and thyroid-related values (T3, T4, TSH, and T3/T4 ratio) did not differ significantly between PMM and MRM (all p>0.50).

Age-stratified comparisons in Table 5 show that participants aged >40 years had a higher prevalence of HT (14.5%) compared with younger individuals (1.8; p=0.002). The mean VAS scores were slightly but significantly lower in the older group (7.5 vs. 8.0; p<0.001). However, other blood and thyroid measures showed no significant differences across age groups (all p>0.05). Age group showed no significant differences in study group distribution or menstrual status (p=0.061 and p=0.095, respectively).

Discussion

In this cross-sectional study of 175 women-comprising MRM, chronic migraine, and healthy controls-we observed no statistically significant differences in serum thyroid hormone parameters (T3, T4, TSH, and the T3/T4 ratio) among groups. Likewise, comparison of PMM and MRM subtypes found no significant difference in thyroid hormone levels. Stratification by age showed a higher prevalence of HT among older participants, consistent with epidemiological expectations, yet thyroid and hematologic measures remained comparable between age groups. These results indicate that thyroid hormone levels do not distinguish migraine types or severity in an otherwise euthyroid population.

Our findings align with previous research that has not demonstrated a robust relationship between thyroid hormones and migraine features. Starikova et al.[10] reported no significant association between TSH/free T4 and headache frequency, and another study[11] found no intergroup differences in thyroid parameters. Comparable results reinforce the notion that routine biochemical assays of thyroid hormones may not fully detect subtle endocrine-neurological interactions in patients with migraine.

Nevertheless, other evidence points toward a more complex and possibly indirect thyroid-migraine connection. According to Michalik et al.[5] systematic review, there was a higher-than-expected co-occurrence of migraine and

Table 3. Comparison of parameters among the MRM, chronic migraine, and control groups

	MRM€, n (%) n=56	Chronic migraine, n (%) n=54	Control, n (%) n=65	p-value
Age (years)*	35.4±7.0	37.7±9.0	36.2±9.8	0.231
Age				0.061
≤40 years old	43 (76.8)	33 (61.1)	37 (56.9)	
>40 years old	13 (23.2)	21 (38.9)	28 (43.1)	
MIDAS score*	13 (5-19)	31 (18-46)	NA	**<0.001**
VAS score*	8 (7-9)	8 (7-9)	NA	0.913
Number of prophylactic drugs*	0 (0-6)	4 (1-6)	NA	**<0.001**
HT	2 (3.6)	7 (13.0)	2 (3.1)	0.052
T3 (ng/dL)*	3.3 (2.6-4.3)	3.5 (2.6-4.5)	3.5 (2.4-4.6)	0.439
T4 (ng/dL)*	0.8 (0.5-1.1)	0.8 (0.5-1.2)	0.9 (0.5-2.0)	0.271
TSH (mIU/L)*	1.7 (0.4-5.2)	1.7 (0.1-18.9)	1.7 (0.1-6.5)	0.942
T3/T4 ratio*	4.0 (0.0- 7.3)	4.3 (2.5-6.1)	4.0 (1.6-6.1)	0.131

*: Numeric variables were presented as median (minimum-maximum) or mean±standard deviation, €: PMM and MRM patients were evaluated in the same group within MRM, PMM: Pure menstrual migraine, MRM: Menstruation-related migraine, MIDAS: Migraine disability assessment, VAS: Visual analog scale, HT: Hypertension, NA: not applicable, T3: Triiodothyronine, T4: Thyroxine, TSH: Thyroid-stimulating hormone

Table 4. MRM, migraine, and control group evaluation of clinical, hematological, and thyroid parameters

	PMM, n (%) n=10	MRM, n (%) n=46	p-value
Age (years)*	29.6±6.2	36.7±6.5	**0.002**
Age			
≤40 years old	10 (100.0)	33 (71.7)	0.095
>40 years old	0 (0.0)	13 (28.3)	
MIDAS score*	8.6±1.8	12.6±3.6	**0.002**
VAS score*	8 (7.5-8.5)	8 (7-9)	0.350
Number of prophylactic drugs*	0 (0-2)	0 (0-6)	0.738
HT	0 (0.0)	2 (4.4)	>0.999
Hb (g/dL)*	12.2±1.3	12.5±1.3	0.335
Hematocrit (%)*	36.5±3.0	37.3±3.0	0.653
MCV (fL)*	71.0±24.6	83.7±7.9	0.164
RDW (%)*	14.8±2.2	14.5±2.0	0.716
Neutrophil (x10^3/µL)*	4.2 (3.1-8.6)	4.2 (1.6-10.9)	0.991
Lymphocyte (x10^3/µL)*	2.8 (1.5-4.3)	2.1 (1.1-4.3)	0.125
Platelets (x10^3/µL)*	271 (194-383)	281 (175-437)	0.881
PDW (%)*	17.0±0.4	16.8±0.5	0.262
T3 (ng/dL)*	3.7 (3.0-4.0)	3.3 (2.6-4.3)	0.131
T4 (ng/dL)*	0.9 (0.8-1.0)	0.8 (0.5-1.1)	0.218
TSH (mIU/L)*	1.1 (0.4-5.2)	1.8 (0.7-4.6)	0.063
T3/T4 ratio*	4.0 (3.4-4.7)	4.0 (0.0-7.3)	0.623
RDW/MCV ratio*	0.2 (0.1-0.5)	0.2 (0.1-0.4)	0.392
NLR*	1.7 (1.2-3.1)	2.4 (0.6-5.7)	0.185
PLR*	130.7 (68.6-152.0)	134.9 (68.6-264.6)	0.325
PDW/PLT ratio*	0.1 (0.1-0.1)	0.1 (0.0-0.1)	0.115
Hb/RDW ratio*	0.9 (0.5-1.0)	0.9 (0.4-1.1)	0.454

*: Numeric variables were presented as median (minimum-maximum) or mean±standard deviation, PMM: Pure menstrual migraine, MRM: Menstruation-related migraine, MIDAS: Migraine disability assessment, VAS: Visual analog scale, HT: Hypertension, Hb: Hemoglobin, MCV: Mean corpuscular volume, RDW: Red cell distribution width, PDW: Platelet distribution width, T3: Triiodothyronine, T4: Thyroxine, TSH: Thyroid-stimulating hormone, NLR: Neutrophil-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, PLT: Platelet count

hypothyroidism, especially Hashimoto's thyroiditis, although a causal relationship was not established. Another study[6] found that overlapping genetic loci associated with both migraine and thyroid traits imply shared neuroendocrine regulatory pathways. Moreover, a Mendelian randomization analysis found an association between autoimmune thyroid disease and migraine risk[14]. These data collectively suggest that thyroid involvement in migraine may be subclinical, genetic, or immune-mediated-mechanisms that would not necessarily manifest as abnormal TSH or T4 levels in biochemically euthyroid individuals.

In contrast to prior smaller studies[7,12], the MRM and PMM subgroups showed no differences in thyroid parameters. Albay and Tütüncü[12] observed a lower mean T4 in the MM group compared with the chronic migraine and control groups, suggesting that thyroid fluctuations could affect hormone-related migraines. Nonetheless, our inability to replicate these findings might be due to several interrelated factors. First, our cohort was deliberately restricted to euthyroid participants, excluding those with overt dysfunction; thus, any subtle endocrine influence would likely fall within reference ranges and below the detection threshold. Additionally, blood sample timing was not aligned with the menstrual cycle phase. Because thyroid hormones exhibit circadian and menstrual-phase variations and interact with estrogen and progesterone[3], unscheduled timing could have masked transient shifts. Thirdly, tests for anti-thyroid

Table 5. Baseline characteristics by age group			
	≤40 years old, n (%) n=113	**>40 years old, n (%) n=62**	**p-value**
Age (years)*	31.5±6.9	45.3±2.9	**<0.001**
Study group			
MRM	43 (38.1)	13 (21.0)	0.061
Chronic migraine	33 (29.2)	21 (33.9)	
Control	37 (32.7)	28 (45.2)	
Menstrual status			
PMM	10 (23.3)	0 (0.0)	0.095
MRM	33 (76.7)	13 (100.0)	
MIDAS score*	16 (5-46)	22 (8-42)	0.675
VAS score*	8 (7-9)	7.5 (7-8.5)	**<0.001**
Number of prophylactic drugs*	0 (0-6)	0 (0-6)	0.824
HT	2 (1.8)	9 (14.5)	**0.002**
Hb (g/dL)*	12.7±1.3	12.5±1.4	0.253
Hematocrit (%)*	37.8±3.4	37.1±4.1	0.109
MCV (fL)*	83.4±10.2	82.4±8.3	0.337
RDW (%)*	14.3±1.8	14.6±2.0	0.178
Neutrophil (x10³/µL)*	4.3 (1.6-16.2)	5.0 (2.3-11.6)	0.054
Lymphocyte (x10³/µL)*	2.1 (0.8-4.3)	2.1 (0.3-4.3)	0.930
Platelets (x10³/µL)*	265 (143-516)	287 (140-458)	0.589
PDW (%)*	16.8±0.5	16.8±0.6	0.779
T3 (ng/dL)*	3.4 (2.6-4.5)	3.4 (2.4-4.6)	0.942
T4 (ng/dL)*	0.8 (0.5-1.3)	0.9 (0.5-2.0)	0.154
TSH (mIU/L)*	1.7 (0.1-6.8)	1.8 (0.3-18.9)	0.365
T3/T4 ratio*	4.2 (0.0-7.3)	3.8 (1.6-6.2)	0.196
RDW/MCV ratio*	0.2 (0.1-0.5)	0.2 (0.1-0.4)	0.229
NLR*	2.1 (0.6-13.3)	2.5 (0.9-38.7)	0.121
PLR*	133.1 (68.6-339.2)	129.8 (71.9-1076.7)	0.975
PDW/PLT ratio*	0.1 (0.1-0.1)	0.1 (0.1-0.1)	**<0.001**
Hb/RDW ratio*	0.9 (0.4-1.2)	0.9 (0.4-1.2)	0.112

*: Numeric variables were presented as median (minimum-maximum) or mean±standard deviation, MRM: Menstruation-related migraine, PMM: Pure menstrual migraine, MIDAS: Migraine disability assessment, VAS: Visual analog scale, HT: Hypertension, Hb: Hemoglobin, MCV: Mean corpuscular volume, RDW: Red cell distribution width, PDW: Platelet distribution width, T3: Triiodothyronine, T4: Thyroxine, TSH: Thyroid-stimulating hormone, NLR: Neutrophil-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, PLT: Platelet count

peroxidase antibodies or anti-thyroglobulin antibodies were not used to evaluate autoimmune thyroiditis in the current study, which has been identified as a key mediator in earlier reports. Migraine has been associated with systemic immune activation[14]; thus, unrecognized early autoimmunity could represent a missing mechanistic link. The homogeneous thyroid profiles among our euthyroid group, despite variable migraine symptoms, could be attributed to these combined methodological and biological factors.

Migraine is a heterogeneous condition that involves vascular, neuronal, inflammatory, and hormonal dimensions. The thyroid axis may influence only selected pathways-for instance, modulating cerebral metabolism or vascular tone-rather than serving as a universal driver. This heterogeneity might also explain the inconsistent results in different studies. Thyroid issues may worsen neurovascular sensitivity through inflammation or changes in nitric oxide in those with genetic risk factors or autoimmune comorbidities[6]. In

contrast, in physiologically stable individuals, compensatory mechanisms of the hypothalamic-pituitary-thyroid axis might neutralize small hormonal changes, thereby precluding measurable biochemical differences. Hence, the thyroid's contribution is likely to be circumstantial, appearing only under specific biological conditions.

Patients with migraines, particularly those with MRM, exhibited considerably higher MIDAS scores and more frequent use of prophylactic medication than the control group, which is consistent with prior research indicating increased disability in hormonally influenced migraine[15]. The persistence of clinical burden despite normal thyroid status underscores that migraine severity is determined primarily by migraine-specific neurovascular mechanisms rather than systemic thyroid factors.

A modest, though non-significant, increase in HT prevalence was also noted in the migraine group (13% vs. 3% in the controls), echoing the recognized comorbidity between migraine and vascular dysfunction[16]. Vascular aging and hormonal shifts contribute to an increased prevalence of HT after age 40. Importantly, neither HT nor hematologic indices correlated with thyroid measures, suggesting that vascular alterations observed in migraine are not mediated by thyroid activity in euthyroid women.

In a clinical context, these results discourage routine thyroid screening in euthyroid migraine patients, even those with MMs, unless other symptoms indicate endocrine problems. However, the thyroid-migraine hypothesis should not be dismissed altogether. Thyroid autoantibody testing may be useful in patients presenting with fatigue, weight changes, menstrual irregularities, or a personal history of autoimmune disease. Furthermore, therapeutic adjustment of thyroid function may warrant investigation in patients with subclinical or autoimmune thyroid disease; a recent pilot trial found that low-dose T4 supplementation reduced migraine frequency in patients with subclinical hypothyroidism[17].

Study Limitations

The present study's strengths include well-defined diagnostic criteria, exclusion of confounding thyroid disease, inclusion of multiple migraine phenotypes, and analysis of hematologic indices alongside hormonal data. Nonetheless, several limitations may be acknowledged. The cross-sectional design prevents causal conclusions, and early thyroid autoimmunity could be missed due to a lack of autoantibody testing. The timing of blood samples was not aligned with circadian and menstrual cycles, which potentially masked subtle phase-dependent shifts. The sample size was sufficient to detect moderate effects, but may have been too small to identify small differences between subgroups, such as PMM.

Future investigations should employ longitudinal, cycle-synchronized sampling, incorporating both hormonal and immune parameters. A more complete picture of endocrine-immune-neurovascular interactions could be achieved by measuring thyroid autoantibodies and neuropeptide markers. Genomic studies may clarify shared predispositions by linking migraine susceptibility loci with thyroid trait variants[6]. Ultimately, interventional research on adjusting thyroid hormones, with a focus on subclinical or autoimmune subgroups, might determine whether endocrine correction confers therapeutic benefit.

Conclusion

In this study, Thyroid hormone levels were comparable among women with MRM, chronic migraine, and healthy controls. These results imply that thyroid hormone changes are unlikely to be the main factor in the migraine presentation in people with normal thyroid function. Nonetheless, given mounting evidence for genetic, autoimmune, and hormonal interplay, further studies integrating immunologic, genetic, and longitudinal data are warranted. Understanding these subtle links between the endocrine-neurological systems might lead to more tailored approaches for preventing and managing migraines.

Ethics

Ethics Committee Approval: This research adhered to the Declaration of Helsinki and was ethically approved by the University of Health Sciences Türkiye, İzmir City Hospital Ethics Committee (approval no: 2025/52, date: 19.03.2025).

Informed Consent: Before joining the study, all participants provided written informed consent.

Footnotes

Authorship Contributions

Surgical and Medical Practises: E.C.Y., Concept: E.C.Y., Design: E.C.Y., D.A., Data Collection or Processing: E.C.Y., D.A., Analysis or Interpretation: E.C.Y., D.A., Literature Search: E.C.Y., D.A., Writing: E.C.Y.

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

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

References

1. Headache classification committee of the International Headache Society (IHS). The International Classification of headache disorders, 3rd edition. Cephalalgia. 2018;38:1-211.
2. Steiner TJ, Stovner LJ, Jensen R, Uluduz D, Katsarava Z; Lifting The Burden: the global campaign against headache. Migraine remains second among the world's causes of disability, and first among young women: findings from GBD2019. J Headache Pain. 2020;21:137.
3. MacGregor EA. Menstrual and perimenopausal migraine: a narrative review. Maturitas. 2020;142:24-30.
4. Parashar R, Bhalla P, Rai NK, Pakhare A, Babbar R. Migraine: is it related to hormonal disturbances or stress? Int J Womens Health. 2014;6:921-5.
5. Michalik M, Łapicka J, Sota M, Zawieska J, Grodzka O, Kępczyńska K. What is the link between migraine and hypothyroidism? A systematic literature review. J Clin Med. 2025;14:4645.
6. Tasnim S, Nyholt DR. Migraine and thyroid dysfunction: co-occurrence, shared genes and biological mechanisms. Eur J Neurol. 2023;30:1815-27.
7. Seo JG. Menstrual migraine: a review of current research and clinical challenges. Headache Pain Res. 2024;25:16-23.
8. Tasnim S, Wilson SG, Walsh JP, Nyholt DR; International Headache Genetics Consortium (IHGC). Shared genetics and causal relationships between migraine and thyroid function traits. Cephalalgia. 2023;43:3331024221139253.
9. Martin AT, Pinney SM, Xie C, et al. Headache disorders may be a risk factor for the development of new-onset hypothyroidism. Headache. 2017;57:21-30.
10. Starikova NL, Baidina TV, Kalashnikova TP. Thyrotropin levels and severity of symptoms in migraine patients of tertiary headache center. Cephalalgia. 2019;39:148-52.
11. Altaş M, Burgucu HÇ, Yarar Z. Evaluation of serum TSH and free T4 levels in migraine patients. Namik Kemal Med J. 2023;11:22-6.
12. Albay VB, Tütüncü M. Evaluation of the relationship between thyroid dysfunction and menstrual migraine in adult females. Med Bull Haseki. 2020;58:110-4.
13. Stewart WF, Lipton RB, Kolodner K, Liberman J, Sawyer J. Reliability of the migraine disability assessment score in a population-based sample of headache sufferers. Cephalalgia. 1999;19:107-14.
14. Biscetti L, De Vanna G, Cresta E, et al. Headache and immunological/autoimmune disorders: a comprehensive review of available epidemiological evidence with insights on potential underlying mechanisms. J Neuroinflammation. 2021;18:259.
15. Chao S, Na L, Libin J, Guiran Y. Analysis of thyroid autoantibodies and thyroid stimulating hormone expression in patients with thyroid diseases in high iodine areas of Cangzhou. Am J Biomed Life Sci. 2024;12:12-5.
16. Sayiner ZA, Eraydın A, Metin T, Özkaya M. Interferon alpha-induced non-immune thyrotoxicosis treated by plasmapheresis. BMJ Case Rep. 2017;2017:bcr2017221347.
17. Alokley A, ALNasser MN, Alabdulqader RA, et al. Effectiveness of low dose thyroxine in patients with subclinical hypothyroidism and migraine: a systematic review and meta-analysis. BMC Neurol. 2025;25:198.

DOI: 10.4274/anatoljmed.2025.26818
Anatol J Gen Med Res 2025;35(3):339-344

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Susceptibility of Bile Cultures to Empiric Antibiotic Therapy and Impact on Clinical Outcomes in Patients with Acute Cholecystitis Undergoing Percutaneous Cholecystostomy

Perkütan Kolesistostomi Uygulanan Akut Kolesistitli Hastalarda Safra Kültürlerinin Ampirik Antibiyotik Tedavisine Duyarlılığı ve Klinik Sonuçlar Üzerindeki Etkisi

Ozan Barış Namdaroğlu[1], Fatma Dikişer[1], Selen Öztürk[1], Erdinç Kamer[2], Savaş Yakan[1]

[1]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of General Surgery, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir City Hospital, Department of General Surgery, İzmir, Türkiye

Cite as: Namdaroğlu OB, Dikişer F, Öztürk S, Kamer E, Yakan S. Susceptibility of bile cultures to empiric antibiotic therapy and impact on clinical outcomes in patients with acute cholecystitis undergoing percutaneous cholecystostomy. Anatol J Gen Med Res. 2025;35(3):339-344

Abstract

Objective: To characterize biliary microbiology, quantify empiric antibiotic-culture susceptibility concordance, and identify mortality-associated factors among acute cholecystitis (AC) patients undergoing percutaneous cholecystostomy (PC).

Methods: We conducted a retrospective single-center cohort (January 2019-June 2024) of adults with AC treated by PC (n=86). Diagnosis relied on clinical, laboratory, and imaging criteria; severity was graded per Tokyo Guidelines 2018 (grade II-III). Collected variables included demographics, comorbidities [Charlson comorbidity index (CCI), C-reactive protein (CRP)/procalcitonin (PCT)], empiric regimen, bile culture/susceptibility, and clinical outcomes. Concordance was defined as full coverage of all cultured organisms by the initial empiric regimen. Multivariable logistic regression assessed predictors of mortality.

Results: Mean age was 72.1±14.0 years; 48.8% were female. Tokyo grade II 85%, grade III 15%; CCI: 1.10±1.08. The predominant empiric regimen was 3rd-generation cephalosporin plus metronidazole (89.5%). Overall bile culture positivity was 68.6%; leading organisms were *Escherichia coli* (35%), *Enterococcus* spp. (22%), and *Klebsiella* spp. (18%). Empiric-culture concordance was 57.1%, with the most prominent discordance for *Enterococcus* (coverage 31.0%). Length of stay, drain removal time, and CCI did not differ between concordant versus discordant therapy. Overall mortality was 10.5%. In multivariable analysis, age independently predicted mortality [odds ratio (OR)=1.12, p=0.03]; higher Tokyo grade showed a non-significant upward trend (p=0.16). Concordance showed a protective trend for mortality (OR=0.23, p=0.12).

Conclusion: In AC managed with PC and timely source-control, empiric-culture concordance is moderate and appears to have limited impact on short-term mortality, which is primarily driven by age and disease severity. The frequency of *Enterococcus* underlies most discordance and should inform empiric choices. These findings support rapid de-escalation and short-course, targeted antibiotics following adequate drainage.

Keywords: Acute cholecystitis, percutaneous cholecystostomy, bile culture, empiric antibiotics, *Enterococcus*, mortality

Address for Correspondence/Yazışma Adresi: Fatma Dikişer, MD, University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of General Surgery, İzmir, Türkiye
E-mail: dikiserfatma@gmail.com
ORCID ID: orcid.org/0000-0001-6256-0124

Received/Geliş tarihi: 14.10.2025
Accepted/Kabul tarihi: 11.11.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

Amaç: Akut kolesistit (AK) nedeniyle perkütan kolesistostomi (PK) uygulanan hastalarda mikrobiyolojik dağılımı, ampirik antibiyotik-kültür duyarlılık uyumunu ve mortalite ile ilişkili faktörleri değerlendirmek.

Yöntem: Ocak 2019-Haziran 2024 arasında tek merkezde yürütülen retrospektif kohortta PK yapılan erişkin AK hastaları incelendi (n=86). Tanı klinik, laboratuvar ve görüntüleme bulgularına; şiddet sınıflaması Tokyo Guidelines 2018 (grade II-III) ölçütlerine göre yapıldı. Kayıt edilen değişkenler: demografi, komorbiditeler [Charlson komordibite indeksi (CCI), C-reaktif protein (CRP)/prokalsitonin (PCT)], ampirik rejim, safra kültürü/antibiyogram ve klinik sonuçlar. Ampirik-kültür uyumu, kültürde üreyen tüm mikroorganizmaların başlanılan ampirik rejimle kapsanması olarak tanımlandı. Çok değişkenli lojistik regresyonla mortalite belirleyicileri analiz edildi.

Bulgular: Ortalama yaş 72,1±14,0 yıl; %48,8 kadın. Tokyo grade II %85, grade III %15; CCI: 1,10±1,08. Başlangıç ampirik tedavi çoğunlukla 3. kuşak sefalosporin+metronidazol (%89,5) idi. Safra kültürü pozitifliği %68,6; en sık etkenler *Escherichia coli* (%35), *Enterococcus* spp. (%22) ve *Klebsiella* spp. (%18). Ampirik-kültür uyumu %57,1 olup uyumsuzluk en belirgin *Enterococcus*'ta izlendi (kapsama %31,0). Uyumlu-uyumsuz gruplar arasında yatış süresi, dren çekilme süresi ve CCI açısından fark yoktu. Toplam mortalite %10,5 idi. Çok değişkenli analizde yaş mortalite ile bağımsız ilişkiliydi [olasılık oranı (OO)=1,12, p=0,03]; daha yüksek Tokyo derecesi anlamlı olmayan artış eğilimi gösterdi (p=0,16). Uyum, mortalite için koruyucu eğilimdeydi (OO=0,23, p=0,12).

Sonuç: PK ile etkin kaynak kontrolü sağlanan AK olgularında ampirik-kültür uyumu orta düzeyde olup kısa dönem mortalite üzerinde sınırlı etkiye sahiptir; mortalite daha çok yaş ve hastalık şiddeti ile ilişkilidir. *Enterococcus* sıklığı uyumsuzluğun başlıca nedenidir ve ampirik rejim seçimini etkiler. Sonuçlar, kültür sonrası hızlı daraltma ve kısa süreli, hedefe yönelik antibiyotik stratejilerini desteklemektedir.

Anahtar Kelimeler: Akut kolesistit, perkütan kolesistostomi, safra kültürü, ampirik antibiyotik, *Enterococcus*, mortalite

Introduction

Acute cholecystitis (AC) is a common cause of acute abdomen, most commonly resulting from cystic duct obstruction and biliary stasis, and is associated with substantial morbidity. Contemporary guidelines emphasize severity grading [Tokyo Guidelines 2018, (TG18)] alongside clinical and laboratory findings for diagnosis and recommend rapid source-control and rational, patient-tailored use of antibiotics[1,2]. Percutaneous cholecystostomy (PC) is an effective bridging or definitive treatment option for patients for whom early cholecystectomy is not preferred or who are at high surgical risk[1].

The duration and spectrum of antibiotic therapy are among the most debated issues during the post-PC period. The TG18 and World Society of Emergency Surgery (WSES) Guidelines generally recommend short-courses when adequate source-control has been achieved (most cases, 4-7 days), and suggest prolongation in the presence of Grampositive coccal bacteremia (especially *Enterococcus* or *Streptococcus*)[2-5]. This approach aligns with the STOP-IT paradigm, which supports short-course therapy after adequate source-control in intra-abdominal infections[6]. Nevertheless, recent reviews highlight the limited evidence supporting these recommendations and the need for personalization using realworld data[7,8].

Microbiologically, while *Enterobacterales*-particularly *Escherichia coli* (*E. coli*)-remain the most frequent pathogens in AC, a relative increase in *Enterococcus* spp. within PC populations has been reported and may thereby heighten the risk of discordance between empiric therapy and culture-based susceptibility results[9-11]. Reported empiric-culture concordance rates vary across PC cohorts. In a 10 year experience, Nitzan et al.[12] reported ≈67% concordance, whereas more recent cohorts suggest that discordance may have limited impact on outcomes when source-control is decisive.

Among biomarkers, procalcitonin (PCT) and C-reactive protein (CRP) may aid in discriminating the severity of AC. Some studies suggest that PCT outperforms CRP for diagnosis and severity, yet evidence regarding independent associations with mortality is inconsistent[13,14]. Therefore, PCT and CRP should be interpreted alongside severity grading and comorbidity burden.

This study aimed to evaluate, among AC patients undergoing PC at our center between 2019 and 2024, the microbiological profile, concordance between empiric antibiotic therapy and culture susceptibility, factors associated with mortality, and clinical outcomes.

Materials and Methods

This study was a retrospective, single-center cohort analysis of adult patients who were admitted to the University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Department of General Surgery, between January

2019 and June 2024, had a diagnosis of AC, and were treated with PC. Diagnosis was based on clinical findings (right upper quadrant pain, fever, Murphy's sign), laboratory findings (leukocytosis, elevated liver function tests), and imaging findings (ultrasound or computed tomography showing wall thickening, pericholecystic fluid, and gallstones). Disease severity was graded I-III according to TG18. PC was preferred as a source-control method in patients at high surgical risk (American Society of Anesthesiologists ≥III) or with poor general condition.

From patient records, the following variables were collected: demographics (age, sex); clinical data (Tokyo Grade, comorbidities); laboratory (white blood cell count, CRP, PCT); treatment (initial empiric antibiotic regimen); microbiology (bile culture results and antibiotic susceptibility); and outcomes (length of stay, time to drain removal, mortality). Comorbidity burden was calculated using the Charlson comorbidity index (CCI), which included the presence of cerebrovascular disease, diabetes, chronic obstructive pulmonary disease (COPD), coronary artery disease, renal failure, dementia, and malignancy. Because hypertension is not included in the CCI, it was reported descriptively but excluded from CCI calculations.

During PC, bile samples were obtained using sterile techniques and cultured using standard laboratory methods. Empiric-culture concordance was defined as complete coverage of all cultured microorganisms by the initial empiric regimen, as determined by susceptibility testing. Patients with "no growth" were excluded from the concordance analysis. Empirical antibiotics were selected according to the TG18 and WSES 2020 recommendations (group 1: third-generation cephalosporin plus metronidazole; group 2: carbapenem; group 3: piperacillin/tazobactam). Antibiotic duration was generally ≤7 days; when resistant isolates, bacteremia, or a persistent systemic response were present, the duration was extended.

Exclusion criteria were: Age <18 years; acute cholangitis or pancreatitis not meeting TG18 criteria for AC; primary perforated cholecystitis with diffuse peritonitis requiring emergent cholecystectomy at presentation; lack of bile sampling at PC; bile culture sampled >48-72 h after initiation of antibiotics, or growth deemed a contaminant; missing culture/susceptibility results; missing essential variables (age, sex, Tokyo Grade, CCI, empiric regimen, CRP/PCT, length of stay, and drain times); loss to follow-up within 30 days; and early failure of source-control after PC necessitating urgent revision/surgery within 24 h.

Ethics Approval

This study was approved by the University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital Non-Interventional Clinical Research Ethics Committee (approval no: 2024/10-13, date: 07.11.2024). The study was conducted in accordance with the principles of the Declaration of Helsinki.

Statistical Analysis

Continuous variables were summarized as mean ± standard deviation or median [interquartile range (IQR)]; categorical variables were summarized as number (%). Appropriate parametric or non-parametric tests were used for group comparisons. Factors predicting mortality were assessed using multivariable logistic regression. Model 1 included age, sex, Tokyo Grade, logCRP, logPCT, and CCI; model 2 (subgroup) included age, Tokyo Grade, and empiric-culture concordance. Two-sided p-values were considered statistically significant. Analyses were performed using SPSS v26.0 (IBM Corp., USA).

Results

Among 86 AC patients who underwent PC, the mean age was 72.1±14.0 years; 48.8% were female. The severity distribution was Tokyo Grade II (85%) and Grade III (15%). The mean CCI was 1.10±1.08. The most frequent comorbidities were hypertension (47.7%), diabetes mellitus (32.6%), and COPD (17.4%) (Table 1). The mean length of stay was 10.1±6.8 days, and the median time to drain removal was 38 days (IQR: 22-54); these did not differ significantly between concordant and discordant antibiotic groups (Table 1).

Initial empirical therapy consisted predominantly of a third-generation cephalosporin plus metronidazole (89.5%), with less frequent use of carbapenems (6.0%) and piperacillin/tazobactam (4.5%) (Table 1). This distribution shaped the coverage rates calculated against culture results (Table 2).

The most commonly isolated organisms in bile cultures were *E. coli* (35%), *Enterococcus* spp. (22%), and *Klebsiella* spp. (18%); lower frequencies included *Enterobacter* spp. (10%), *Streptococcus* spp. (5%), and others (mainly *Pseudomonas* and *Candida*), accounting for 8%. The overall culture positivity rate was 68.6% (Table 2).

Empiric-culture concordance was 57.1%. By organism, coverage was 90.4% for *E. coli*, 72.7% for *Klebsiella* spp., 50.0% for *Enterobacter* spp., 100.0% for *Streptococcus* spp., and 40.0% for the "other" group. Discordance was most pronounced for *Enterococcus* spp. (coverage: 31.0%), reflecting the predominance of cephalosporin-based empiric regimens (Table 2).

There were no significant differences between the concordant and discordant groups in length of stay, time to drain removal, or CCI (Table 1). Overall mortality was 10.5%. In multivariable analysis, age was independently associated with mortality [odds ratio (OR)=1.12, p=0.03], while a higher Tokyo Grade showed a nonsignificant trend (p=0.16). Empiric-culture concordance showed a protective trend for mortality (OR=0.23, p=0.12; Table 3).

Discussion

In this singlecenter retrospective cohort study (2019-2024), we analyzed microbiology, empiric-culture concordance,

Table 1. Demographic and clinical characteristics (n=86)

Variable	Value
Age (years), mean ± SD	72.1±14.0
Female sex, n (%)	42 (48.8)
Tokyo grade II/III, n (%)	73 (84.9)/13 (15.1)
CCI, mean ± SD	1.10±1.08
Comorbidities, n (%)	Hypertension 41 (47.7), diabetes mellitus 28 (32.6), COPD 15 (17.4), cerebrovascular disease 9 (10.4), malignancy 7 (8.1)
Empiric antibiotic regimen	3rd-Generation cephalosporin+metronidazole 77 (89.5), carbapenem 5 (6.0), piperacillin/tazobactam 4 (4.5)
Length of stay (days), mean ± SD	10.1±6.8
Drain removal time (days), median (IQR)	38 (22-54)

SD: Standard deviation, CCI: Charlson comorbidity index, COPD: Chronic obstructive pulmonary disease, IQR: Interquartile range

Table 2. Bile culture results and empiric therapy-culture concordance

Microorganism	n (%)	Covered by empiric regimen (%)
Escherichia coli	21 (35.0)	90.4
Enterococcus spp.	13 (22.0)	31.0
Klebsiella spp.	11 (18.0)	72.7
Enterobacter spp.	6 (10.0)	50.0
Streptococcus spp.	3 (5.0)	100.0
Other (e.g., *Pseudomonas*, *Candida*)	5 (8.0)	40.0
Total positive cultures	59 (68.6)	-
Empiric-culture concordance	-	57.1

Table 3. Multivariable logistic regression analysis (mortality)

Variable	OR	95% CI	p
Age (per 1 year)	1.12	1.01-1.25	0.03*
Sex (female)	0.81	0.19-3.48	0.78
Tokyo grade (II–III)	3.96	0.59-26.7	0.16
log-CRP	1.09	0.73-1.63	0.67
log-PCT	1.15	0.82-1.76	0.48
CCI	1.34	0.89-2.14	0.21
Subgroup: culture concordance	0.23	0.04-1.44	0.12

OR: Odds ratio, CI: Confidence interval, CRP: C-reactive protein, PCT: Procalcitonin, CCI: Charlson comorbidity index

and mortality among AC patients treated with PC. Culture positivity was 68.6%, empiric-culture concordance was 57.1%, and mortality was 10.5%. In multivariable analysis, age was the only independent predictor of mortality; higher Tokyo Grade showed a nonsignificant upward trend (OR=3.96, 95% confidence interval: 0.59-26.7, p=0.16). These findings are broadly consistent with larger series in the literature[2,4,7-9,12].

E. coli (35%), *Enterococcus* spp. (22%), and *Klebsiella* spp. (18%) were the leading organisms, consistent with the flora anticipated by TG18 and WSES[1,5]. The higher proportion of *Enterococcus* warrants attention. Although *E. coli* remains the main pathogen-reflecting the classical calculus-stasis-colonization pathway-aging, comorbidities, prior antibiotic exposure, and nosocomial contact may shift flora toward Gram-positive and resistant organisms[7,10]. Lahav et al.[4] reported *E. coli* in 33% and *Enterococcus* in 24%, while Nitzan et al.[12] reported 60% culture positivity and 67% concordance. Our 57% concordance rate likely reflects increased isolation of *Enterococcus* and widespread use of cephalosporin-based empiric therapy.

Enterobacterales constitute the classical etiologic group in AC. In our cohort, coverage was high for *E. coli* (90%) and *Klebsiella* spp. (73%), consistent with reports by Suh et al.[2] and Fico et al.[3]. The 2024 European Centre for Disease Prevention and Control Report notes extended-spectrum beta-lactamase (ESBL) rates of ≈31% in *E. coli* and ≈38% in *Klebsiella pneumoniae*[7]. Although such trends can drive carbapenem use, both TG18 and WSES advocate short, targeted therapy. Notably, high concordance within *Enterobacterales* did not translate into mortality differences after effective drainage, underscoring the primacy of source-control.

Historically regarded as secondary colonizers, enterococci may act as true pathogens in elderly PC patients, patients with comorbidities, and antibioticexposed PC patients[10,11]. In our cohort, *Enterococcus* spp. were isolated in 22% of cases, similar to Nitzan(14%)[12] and Lahav (24%)[3,4]. Given that most empiric regimens were cephalosporin-based, coverage was only 31%. Although we found no direct association between *Enterococcus* isolation and mortality, most discordant cases were linked to this organism, potentially prolonging clinical recovery. In highrisk scenarios (healthcareassociated AC, immunosuppression, high CCI), initial coverage with ampicillinbased agents or piperacillin/tazobactam can be considered; for vancomycin-resistant enterococci (VRE), linezolid or daptomycin are options[15-18].

Streptococcus spp. (5%) were within the expected range (3-10%)[18]. *Pseudomonas* spp. are typically encountered in nosocomial or postendoscopic retrograde cholangiopancreatography settings, where empiric coverage is often inadequate; piperacillin/tazobactam or a carbapenem may be preferred[1,5]. Candida species were rare (approximately 2%), but may require antifungal therapy in patients with malignancy or immunosuppression[19].

Empiric-culture discordance showed a protective trend compared with concordance (OR≈0.23) but did not significantly affect shortterm mortality, echoing Lahav et al.[4] who reported higher 90 day readmission among patients receiving discordant therapy (64% vs. 47%) without differences in mortality. These observations suggest that antibiotic concordance may impact longer-term morbidity rather than short-term mortality when source-control is prompt and effective[15]. Accordingly, source-control remains a prognostic factor that can outweigh the effects of empiric mismatch.

CRP and PCT were not independently associated with mortality in our cohort, aligning with prior work indicating utility for diagnosis and severity assessment but limited prognostic value for hard outcomes[13,14]. Biomarkers should therefore inform, rather than dictate, clinical decision-making alongside TG18 Grade and CCI.

Study Limitations

This retrospective, singlecenter design limits statistical power for mortality analyses; molecular resistance genotyping (e.g., for ESBL, AmpC, and VRE) was not available. Strengths include a contemporary (post2019) cohort that reflects current antimicrobial stewardship principles and the applications of TG18/WSES. Prospective multicenter studies across diverse geographies are needed to refine local adaptations of guideline-based therapy.

Conclusion

This single-center retrospective study demonstrates that among patients undergoing PC for AC empiric antibiotic-culture concordance was achieved in only 57% of cases. Despite this moderate concordance, mortality was primarily driven by age and disease severity (Tokyo Grade) rather than by microbiological mismatch. The findings emphasize that timely and effective source-control remains the cornerstone of management in this fragile, high-risk population. While discordant empiric therapy did not significantly affect short-term mortality, it may contribute to prolonged

hospitalization, recurrent infection, or readmission, highlighting the need for ongoing antimicrobial stewardship. From a microbiological perspective, *Enterobacterales* remain dominant and generally respond to cephalosporin-based regimens; however, the rising proportion of *Enterococcus* spp., largely uncovered by these agents, signals an evolving shift toward resistant or mixed biliary flora. Empiric regimens should therefore be individualized, considering patient age, comorbidities, prior healthcare exposure, and local resistance data. Routine broadening of coverage is unnecessary, but selective inclusion of Enterococcus-active agents may be justified in high-risk or healthcare-associated cases. In accordance with the TG18 and WSES recommendations, short-course targeted antibiotic therapy (≤7 days) following successful drainage should be prioritized. Future multicenter prospective studies integrating molecular resistance profiling and long-term outcomes (reinfection, reintervention, post-PC cholecystectomy) are warranted.

Ethics

Ethics Committee Approval: This study was approved by the University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital Non-Interventional Clinical Research Ethics Committee (approval no: 2024/10-13, date: 07.11.2024).

Informed Consent: This study was a retrospective, single-center cohort study.

Footnotes

Authorship Contributions

Surgical and Medical Practises: O.B.N., F.D., S.Ö., E.K., S.Y., Concept: O.B.N., E.K., S.Y., Design: O.B.N., S.Ö., S.Y., Data Collection or Processing: O.B.N., F.D., S.Ö., Analysis or Interpretation: O.B.N., E.K., S.Y., Literature Search: O.B.N., F.D., S.Ö., Writing: O.B.N., F.D., E.K.

Conflict of Interest: No conflict of interest was declared by the authors. One of the authors of this article (S.Y.) is a member of the Editorial Board of this journal. He was completely blinded to the peer review process of the article.

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

References

1. Gomi H, Solomkin JS, Schlossberg D, et al. Tokyo Guidelines 2018: antimicrobial therapy for acute cholangitis and cholecystitis. J Hepatobiliary Pancreat Sci. 2018;25:3-16.
2. Suh SW, Choi YS, Choi SH, et al. Antibiotic selection based on microbiology and resistance profiles of bile from gallbladder of patients with acute cholecystitis. Sci Rep. 2021;11:2969.
3. Fico V, La Greca A, Tropeano G, et al. Updates on antibiotic regimens in acute cholecystitis. Medicina (Kaunas). 2004;60:1040.
4. Lahav L, Goldberg N, Jiryis T, et al. Impact of discordant antibiotics on outcomes after percutaneous cholecystostomy for acute cholecystitis: a retrospective analysis of 184 PCC Patients. J Clin Med. 2025;14:6589.
5. Pisano M, Allievi N, Gurusamy K, et al. 2020 World Society of Emergency Surgery updated guidelines for the diagnosis and treatment of acute calculus cholecystitis. World J Emerg Surg. 2020;15:61.
6. Sawyer RG, Claridge JA, Nathens AB, et al. Trial of short-course antimicrobial therapy for intraabdominal infection. N Engl J Med. 2015;372:1996-2005. Erratum in: N Engl J Med. 2018;378:686.
7. European Centre for Disease Prevention and Control. Antimicrobial resistance in the EU/EEA (EARS-Net) - Annual Epidemiological Report 2023. Stockholm: ECDC; 2024. Available from: www.ecdc.europa.eu/en/publications-data/antimicrobial-resistance-eueea-ears-net-annual-epidemiological-report-2023
8. Masuda S, Imamura Y, Ichita C, et al. Efficacy of short-course antibiotic therapy for acute cholangitis with positive blood cultures: a retrospective study. Cureus. 2024;16:e58883.
9. Brook I. Aerobic and anaerobic microbiology of biliary tract disease. J Clin Microbiol. 1989;27:2373-5.
10. Sung YK, Lee JK, Lee KH, Lee KT, Kang CI. The clinical epidemiology and outcomes of bacteremic biliary tract infections caused by antimicrobial-resistant pathogens. Am J Gastroenterol. 2012;107:473-83.
11. Mussa M, Martínez Pérez-Crespo PM, Lopez-Cortes LE, et al. Risk factors and predictive score for bacteremic biliary tract infections due to Enterococcus faecalis and Enterococcus faecium: a multicenter cohort study from the PROBAC project. Microbiol Spectr. 2022;10:e0005122.
12. Nitzan O, Brodsky Y, Edelstein H, et al. Microbiologic data in acute cholecystitis: ten years' experience from bile cultures obtained during percutaneous cholecystostomy. Surg Infect (Larchmt). 2017;18:345-9.
13. Hou Q, Li H, Liu C, et al. Procalcitonin and C-reactive protein as biomarkers for diagnosing and assessing the severity of acute cholecystitis. Open Med (Wars). 2025;20:20251258.
14. Spoto S, Valeriani E, Caputo D, et al. The role of procalcitonin in the diagnosis of bacterial infection after major abdominal surgery: advantage from daily measurement. Medicine (Baltimore). 2018;97:e9496.
15. Zarour S, Imam A, Kouniavsky G, Lin G, Zbar A, Mavor E. Percutaneous cholecystostomy in the management of high-risk patients presenting with acute cholecystitis: timing and outcome at a single institution. Am J Surg. 2017;214:456-61.
16. World Health Organization (WHO). Global antimicrobial resistance and use surveillance system (GLASS) report 2022. Geneva: World Health Organization; 2022.
17. Lee KJ, Park SW, Park DH, et al. Gallbladder perforation in acute acalculous vs. calculous cholecystitis: a retrospective comparative cohort study with 10-year single-center experience. Int J Surg. 2024;110:1383-91.
18. Lee JM, Kang JS, Choi YJ, et al. Suggested use of empirical antibiotics in acute cholecystitis based on bile microbiology and antibiotic susceptibility. HPB (Oxford). 2023;25:568-76.
19. Musial CE, Cockerill FR 3rd, Roberts GD. Fungal infections of the immunocompromised host: clinical and laboratory aspects. Clin Microbiol Rev. 1988;1:349-64.

DOI: 10.4274/anatoljmed. 2025.28290
Anatol J Gen Med Res 2025;35(3):345-352

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Investigation of Risk Factors for Exocrine Pancreatic Insufficiency and Type 3c Diabetes After Pancreatic Surgery

Pankreas Cerrahisi Sonrası Egzokrin Pankreas Yetmezliği ve Tip 3c Diyabet için Risk Faktörlerinin Araştırılması

Serhan Akalın[1], Ogün Aydoğan[1], Erdem Barış Cartı[1], Akay Edizsoy[1], Muhammed Çağrı Coşkun[2], Volkan Taşçı[3]

[1]Aydın Adnan Menderes University Faculty of Medicine, Department of General Surgery, Aydın, Türkiye
[2]Aydın Atatürk State Hospital, Clinic of General Surgery, Aydın, Türkiye
[3]Sakarya Training and Research Hospital, Clinic of Radiology, Sakarya, Türkiye

Cite as: Akalın S, Aydoğan O, Cartı EB, Edizsoy A, Coşkun MÇ, Taşçı V. Investigation of risk factors for exocrine pancreatic insufficiency and type 3c diabetes after pancreatic surgery. Anatol J Gen Med Res. 2025;35(3):345-352

Abstract

Objective: This study aimed to evaluate the risk factors for exocrine pancreatic insufficiency and type 3c diabetes following pancreatic surgery.

Methods: A total of 48 patients who underwent pancreatic surgery were retrospectively analyzed. Demographic data, preoperative diabetes history, type of surgery (Whipple, total or distal pancreatectomy), tumor stage, and biochemical parameters [CA19-9, CA125, carcinoembryonic antigen, body mass index (BMI), H-index] were recorded. Postoperative glycated hemoglobin (HbA1c), C-peptide, elastase levels, and incidence of hypoglycemic episodes were evaluated. Regression analyses were performed to identify related predictors.

Results: The mean age of the patients was 59.6 years, with an average BMI of 24.1 kg/m^2. Whipple procedure was the most commonly performed surgery (64.6%). Preoperative diabetes prevalence was 22.9%. CA19-9 levels showed a significant positive correlation with postoperative C-peptide levels (p=0.014). Significant predictors of HbA1c levels were preoperative diabetes (p<0.001), CA19-9 (p=0.045), and BMI (p=0.030). The incidence of hypoglycemia was 4.2%, with no significant difference in HbA1c levels (p=0.333). Postoperative elastase level was the only significant determinant of C-peptide (p=0.048).

Conclusion: Preoperative CA19-9, history of diabetes, and BMI may serve as important predictors for postoperative pancreatic dysfunction. Elastase level is a potential independent indicator of pancreatic endocrine reserve.

Keywords: CA19-9, C-peptide, elastase, exocrine insufficiency, HbA1c, type 3c diabetes, pancreatic surgery

Öz

Amaç: Bu çalışmada, pankreas cerrahisi sonrası gelişebilecek ekzokrin pankreas yetmezliği ve tip 3c diyabetin risk faktörlerinin değerlendirilmesi amaçlanmıştır.

Yöntem: Pankreas cerrahisi uygulanan 48 hastanın demografik özellikleri, preoperatif diyabet öyküsü, cerrahi yöntem (Whipple, total veya distal pankreatektomi), tümör evresi ve biyokimyasal parametreleri [CA19-9, CA125, karsinoembriyonik antijen, vücut kitle indeksi (VKİ), H-indeks] retrospektif olarak incelenmiştir. Postoperatif dönemde glikozillenmiş hemoglobin (HbA1c), C-peptit, elastaz düzeyleri ve hipoglisemik atak insidansı değerlendirilmiş; regresyon analizleri ile ilişkili belirleyiciler araştırılmıştır.

Address for Correspondence/Yazışma Adresi: Assoc. Prof., Erdem Barış Cartı, Aydın Adnan Menderes University Faculty of Medicine, Department of General Surgery, Aydın, Türkiye
E-mail: erdemcarti@yahoo.com
ORCID ID: orcid.org/0000-0002-7139-0339

Received/Geliş tarihi: 05.10.2025
Accepted/Kabul tarihi: 12.11.2025
Published date/Yayınlanma tarihi: 30.12.2025

Bulgular: Hastaların ortalama yaşı 59,6 yıl, ortalama VKİ değeri 24,1 kg/m2 olup, en sık uygulanan cerrahi prosedür Whipple'dır (%64,6). Preoperatif diyabet oranı %22,9'dur. CA19-9 düzeyi ile postoperatif C-peptit düzeyi arasında anlamlı pozitif korelasyon saptanmıştır (p=0,014). HbA1c düzeyini öngören başlıca faktörler; preoperatif diyabet varlığı (p<0,001), CA19-9 (p=0,045) ve VKİ (p=0,030) olmuştur. Hipoglisemik atak insidansı %4,2 olup HbA1c düzeyleri ile anlamlı fark göstermemiştir (p=0,333). Elastaz düzeyi, C-peptit düzeyinin bağımsız belirleyicisi olarak bulunmuştur (p=0,048).

Sonuç: CA19-9 düzeyi, preoperatif diyabet öyküsü ve VKİ, pankreas cerrahisi sonrası gelişebilecek pankreatik disfonksiyonun öngörülmesinde önemli biyobelirteçlerdir. Elastaz düzeyi ise endokrin rezervin bağımsız göstergesi olarak değerlendirilebilir.

Anahtar Kelimeler: CA19-9, C-peptit, elastaz, ekzokrin yetmezlik, HbA1c, tip 3c diyabet, pankreas cerrahisi

Introduction

The pancreas is a complex organ with both exocrine and endocrine functions. Its exocrine component secretes digestive enzymes, including trypsin, chymotrypsin, lipase, and amylase, which are essential for gastrointestinal function[1]. The endocrine pancreas regulates glucose homeostasis through hormones secreted by islet cells, including insulin, glucagon, somatostatin, and pancreatic polypeptide[2]. These systems are closely integrated, and any disruption may impair both.

Pancreatic surgery is performed for malignant and benign tumors, chronic pancreatitis, trauma, or cystic lesions. Common operations include pancreaticoduodenectomy (PD), distal pancreatectomy, and total pancreatectomy (TP)[3]. Such procedures alter pancreatic anatomy and function, predisposing patients to complications.

One of the most significant is exocrine pancreatic insufficiency (EPI), which is caused by inadequate enzyme secretion. EPI results in malnutrition, weight loss, steatorrhea, and reduced quality of life[4,5]. Management includes pancreatic enzyme-replacement therapy, dietary adjustments, and lifestyle modifications; these interventions help restore digestive function and reduce morbidity[5].

Endocrine dysfunction is also a concern, often leading to type 3c diabetes (pancreatogenic diabetes) secondary to pancreatic disease[6]. Unlike type 1 or type 2 diabetes, this condition involves a deficiency of both insulin and glucagon, making glycemic control difficult. This condition is associated with chronic pancreatitis, cystic fibrosis, pancreatic neoplasms, and post-pancreatectomy states[7-9]. Although its prevalence is uncertain, studies suggest it accounts for 5-10% of diabetes cases[10].

Type 3c diabetes is characterized by unstable glycemic patterns, frequent hypoglycemic episodes, and impaired counterregulatory responses due to hormonal deficiencies[7,11,12]. These issues are especially pronounced after surgery, as pancreatectomy contributes to glycemic instability[13].

Reported risk factors for EPI and type 3c diabetes include the type and extent of resection, remaining pancreatic tissue, histological features, and comorbidities[14]. Yet many remain unclear, limiting the effective prediction and management.

The aim of this study is to determine the incidence of EPI and type 3c diabetes following pancreatic surgery and to identify related clinical, surgical, and pathological risk factors. The findings are intended to support improved postoperative monitoring and tailored therapeutic strategies.

Materials and Methods

Participants and Inclusion Criteria

The study included patients who underwent pancreatic surgery, specifically the Whipple procedure, distal pancreatectomy, or TP, at Aydın Adnan Menderes University Hospital between 2015 and 2024. This study approved by Aydın Adnan Menderes Unıversity Ethic Committe (approval no: 2024/140, date: 25.07.2024).

Data Collection

Demographic data, surgical details, laboratory results, and follow-up records of the participants were obtained from the hospital information management system and archived medical records. All collected data were standardized and entered into an electronic data form. Preoperative and postoperative parameters were compared in relation to both endocrine and exocrine pancreatic functions.

Evaluated Parameters

Demographic variables included age, sex, and body mass index (BMI). Surgical variables recorded included the type of procedure (Whipple, distal, or TP) and the duration of

postoperative follow-up. Tumor staging was based on the tumor nodes metastasis (TNM) classification system, and the presence of neuroendocrine tumors was also evaluated. Preoperative biochemical parameters included CA19-9, CA125, and carcinoembryonic antigen levels; H-index; pancreatic duct diameter; and a history of preoperative diabetes. In the postoperative period, HbA1c (%), C-peptide (ng/mL), and serum elastase (ng/dL) levels were measured and analyzed as indicators of metabolic function. The occurrence of hypoglycemic episodes was also assessed as a postoperative complication. In the present study, serum elastase levels were measured. Although fecal elastase is regarded as the gold standard, it was not assessed in our patients because, as documented in the literature, its levels are influenced by oral pancreatic enzyme-replacement therapy, which our patients were receiving.

Postoperative EPI was defined as a fecal elastase level below 200 ng/dL. Type 3c diabetes was defined as newly diagnosed diabetes after pancreatic surgery in individuals with no prior history of diabetes, based on an HbA1c level ≥6.5% or a fasting plasma glucose level ≥126 mg/dL. A hypoglycemic episode was considered present in cases where plasma glucose levels were recorded as <70 mg/dL in conjunction with clinical symptoms.

Statistical Analysis

All data obtained in the study were analyzed using IBM SPSS Statistics for Windows, version 27.0 (IBM Corp., Armonk, NY, USA). To assess the distribution characteristics of continuous variables, normality tests were performed, including the Kolmogorov-Smirnov and Shapiro-Wilk tests. In addition, histogram plots, Q-Q plots, and Skewness and Kurtosis values were reviewed to interpret distribution patterns. Continuous variables that were normally distributed were analyzed parametrically and are presented as mean ± standard deviation. Non-normally distributed variables were expressed as the median (minimum-maximum). Categorical variables were reported as counts and percentages. Comparisons between two groups were performed using the independent samples t-test for normally distributed data and the Mann-Whitney U test for nonparametric data. For categorical comparisons, the Pearson chi-square test was used; however, Fisher's exact test was applied when more than 20% of the expected cell counts were below 5. Relationships between continuous variables were assessed using Pearson correlation analysis (when normality was confirmed). Correlations between postoperative HbA1c, C-peptide, and serum elastase levels were reported with correlation coefficients (r) and p-values. To determine the effects of variables on dependent outcomes, regression analyses were performed. Initially, univariate linear regression was used to assess the effect of each independent variable. Variables found to be significant or borderline significant ($p<0.20$) were included in a multiple linear regression model. These models were applied to continuous outcome variables, such as postoperative levels of HbA1c, C-peptide, and serum elastase. Model fit was assessed using the F-test, and regression coefficients (B), standard errors, 95% confidence intervals (CI), and p-values were reported. Multicollinearity was assessed using the variance inflation factor. A p-value of <0.05 was considered statistically significant in all analyses.

Results

Table 1 presents the basic demographic characteristics and diabetes status of 48 patients who underwent pancreatic surgery. The mean age of participants was 59.6±11.0 years, and the mean BMI was 24.1±4.5 kg/m^2. Of the total cohort, 43.8% (n=21) were female and 56.3% (n=27) were male. Preoperatively, 22.9% of patients (n=11) had a diagnosis of diabetes, whereas 77.1% (n=37) did not. The mean duration of postoperative follow-up was 21.9±27.2 months. Among the procedures performed, 13 patients (27.1%) underwent distal pancreatectomy, 4 patients (8.3%) underwent TP, and 31 patients (64.6%) underwent the Whipple procedure.

Table 1. Baseline demographic characteristics and diabetes status of patients

Characteristic	n (%) / Mean ± SD
Age (years)	59.6±11.0
Sex	Female: 21 (43.8%) Male: 27 (56.3%)
BMI (kg/m^2)	24.1±4.5
Preoperative diabetes	Yes: 11 (22.9%) No: 37 (77.1%)
Postoperative follow-up (months)	21.9±27.2
Type of surgery	Distal: 13 (27.1%) Total: 4 (8.3%) Whipple: 31 (64.6%)
Stage	1a: 1 (2.5%) 1b: 8 (20.0%) 2a: 3 (7.5%) 2b: 15 (37.5%) 3: 11 (27.5%) Neuroendocrine tumor: 2 (5.0%)

SD: Standard deviation, BMI: Body mass index, a,b: TNM classification

Table 2 presents the results of a multiple linear regression analysis evaluating the association between preoperative variables and postoperative HbA1c levels. The analysis revealed that both preoperative diabetes and CA19-9 levels significantly influenced postoperative HbA1c levels. Preoperative diabetes was found to have a strong, statistically significant negative effect on postoperative HbA1c levels (B=-2.858, standard coefficient B=-0.774, t=-7.267, p<0.001, 95% CI: -3.671 to -2.044), indicating that patients with preoperative diabetes had significantly lower postoperative HbA1c levels. Similarly, CA19-9 levels demonstrated a significant, albeit weak, negative effect on postoperative HbA1c levels (B=0.000, standard coefficient B=-0.297, t=-2.123, p=0.045, 95% CI: 0.000 to 0.000).

Table 3 displays the results of a multivariate linear regression analysis investigating preoperative variables that affect postoperative serum C-peptide levels. The analysis identified postoperative elastase level as the only variable with a

Table 2. Multiple linear regression analysis of preoperative variables predicting postoperative HbA1c level

Variable	Unstandard coefficient B	Standard coefficient B	t	p-value	95% CI lower	95% CI upper
(Constant)	11.066	–	4.557	0.000	6.043	16.088
Age (years)	-0.006	-0.036	-0.275	0.786	-0.048	0.037
Sex	0.168	0.052	0.461	0.649	-0.587	0.923
Type of surgery	-0.085	-0.044	-0.325	0.748	-0.623	0.454
CA19-9 (U/mL)	0.000	-0.297	-2.123	0.045	0.000	0.000
CA125 (U/mL)	0.002	0.065	0.335	0.741	-0.009	0.012
CEA (ng/mL)	0.076	0.314	1.470	0.155	-0.031	0.183
H-index	0.003	0.036	0.301	0.766	-0.017	0.022
Preoperative pancreatic diameter	-0.077	-0.107	-0.913	0.371	-0.250	0.097
TNM classification	-0.114	-0.086	-0.782	0.442	-0.417	0.188
BMI (kg/m²)	0.053	0.112	0.874	0.391	-0.073	0.179
Preop diabetes	-2.858	-0.774	-7.267	<0.001	-3.671	-2.044

CI: Confidence interval, CEA: Carcinoembryonic antigen, TNM: Tumor nodes metastasis, BMI: Body mass index

Table 3. Multivariate linear regression analysis of preoperative variables affecting postoperative serum C-peptide level

Variable	Unstandard B	p-value	95% CI lower	95% CI upper
(Constant)	-5.633	0.520	-23.614	12.349
Age (years)	-0.059	0.181	-0.148	0.030
Sex	-0.280	0.707	-1.814	1.254
Type of surgery	0.425	0.401	-0.611	1.460
CA19-9 (U/mL)	0.000	0.103	0.000	0.001
CA125 (U/mL)	-0.004	0.724	-0.026	0.018
CEA (ng/mL)	0.018	0.857	-0.191	0.228
H-index	-0.026	0.264	-0.073	0.021
Preoperative pancreatic diameter	-0.238	0.166	-0.583	0.108
TNM classification	0.237	0.411	-0.353	0.827
BMI (kg/m²)	-0.171	0.288	-0.498	0.156
Preoperative diabetes	2.000	0.179	-1.003	5.003
Hypoglycemic attack	4.851	0.149	-1.896	11.598
Postoperative HbA1c (%)	0.390	0.402	-0.561	1.340
Postoperative elastase (ng/dL)	1.271	0.048	-0.048	2.590

CI: Confidence interval, CEA: Carcinoembryonic antigen, TNM: Tumor nodes metastasis, BMI: Body mass index

statistically significant association with C-peptide levels (B=1.271, p=0.048, 95% CI: -0.048 to 2.590). An increase in postoperative elastase was significantly associated with higher serum C-peptide levels. No other preoperative or clinical variable showed a statistically significant correlation with postoperative C-peptide levels (all p>0.05).

Table 4 summarizes the results of a multivariate linear regression analysis evaluating preoperative variables influencing postoperative serum elastase levels. The analysis indicated that apart from age and postoperative C-peptide levels, no other variables had statistically significant effects on serum elastase levels (all p>0.05).

Table 5 presents the results of the correlation analysis between postoperative levels of HbA1c, C-peptide, and elastase. The analysis revealed a negative correlation between postoperative HbA1c and C-peptide levels (correlation coefficient=-0.263), although this relationship did not reach statistical significance (p=0.071). A positive correlation was observed between postoperative HbA1c and elastase levels (correlation coefficient=0.237); however, this finding was not statistically significant (p=0.108). Likewise, although a positive correlation was noted between postoperative C-peptide and elastase levels, the correlation did not achieve statistical significance (correlation coefficient=0.163, p=0.274).

Table 6 compares postoperative HbA1c levels between patients who experienced hypoglycemic episodes and those who did not after TP. The mean postoperative HbA1c level was 9.4±0.8 in patients who experienced hypoglycemia, compared to 7.3±1.0 in those without such episodes. However, the difference between the two groups was not statistically significant (p=0.333).

Table 4. Multivariate linear regression analysis of preoperative variables affecting postoperative serum elastase level

Variable	Unstandard B	p-value	95% CI lower	95% CI upper
(Constant)	-1.467	0.613	-7.434	4.500
Age (years)	0.030	0.035	0.002	0.057
Sex	0.162	0.508	-0.341	0.665
Type of surgery	-0.142	0.396	-0.484	0.200
CA19-9 (U/mL)	-5.57	0.945	0.000	0.000
CA125 (U/mL)	0.003	0.365	-0.004	0.010
CEA (ng/mL)	-0.023	0.495	-0.091	0.046
H-index	0.005	0.544	-0.011	0.021
Preoperative pancreatic diameter	0.063	0.274	-0.054	0.179
TNM classification	-0.036	0.706	-0.234	0.162
BMI (kg/m²)	0.028	0.597	-0.082	0.139
Preoperative diabetes	-0.242	0.630	-1.277	0.793
Hypoglycemic attack	-0.312	0.785	-2.665	2.042
Postoperative HbA1c (%)	0.028	0.856	-0.292	0.348
Postoperative C-peptide (ng/mL)	0.139	0.048	-0.005	0.283

CI: Confidence interval, CEA: Carcinoembryonic antigen, TNM: Tumor nodes metastasis, BMI: Body mass index

Table 5. Correlation analysis among postoperative HbA1c, C-peptide, and elastase levels

	Postoperative HbA1c (%)	Postoperative C-peptide (ng/mL)	Postoperative elastase (ng/dL)
Postoperative HbA1c (%)-correlation	1	-0.263	0.237
p-value		0.071	0.108
Postoperative C-peptide (ng/mL)-correlation	-0.263	1	0.163
p-value	0.071		0.274
Postoperative elastase (ng/dL)-correlation	0.237	0.163	1
p-value	0.108	0.274	

Table 6. Comparison of postoperative HbA1c levels according to hypoglycemic attack status in patients undergoing total pancreatectomy

Variable	Hypoglycemic attack present (n=2) Mean ± SD	Hypoglycemic attack absent (n=2) Mean ± SD	p-value
Postoperative HbA1c (%)	9.4±0.8	7.3±1.0	0.333

SD: Standard deviation

Discussion

In our study, the presence of preoperative diabetes was shown to have a significant impact on postoperative glycemic control. Multivariate analysis identified preoperative diabetes as the most influential factor affecting postoperative HbA1c levels; with diabetic patients demonstrating significantly lower HbA1c values after surgery. Interestingly, this finding contrasts with expectations in the literature, where patients with pre-existing diabetes undergoing pancreatic surgery are generally considered at higher risk of persistent postoperative hyperglycemia[15]. Indeed, previous studies have reported that a history of preoperative metabolic disorders (particularly diabetes) is a strong predictor of postoperative insulin requirement[16]. This phenomenon may be explained by the fact that patients with a prior diagnosis of diabetes tend to possess greater awareness of their condition, leading to better adherence to dietary recommendations and prescribed treatments.

In the general diabetic population, female patients may also be at a disadvantage compared with male patients in achieving optimal glycemic control, with a lower likelihood of reaching target HbA1c levels[17]. This disparity is thought to be influenced by sex-based differences in insulin sensitivity, visceral fat distribution, and hormonal profiles[17]. In our study, a significant interaction was observed between sex and preoperative diabetes status. The glycemic impact of preoperative diabetes was less pronounced in male patients than in female patients, suggesting that women with diabetes may face greater challenges in achieving postoperative glycemic targets. In addition to serving as a tumor marker for pancreatic cancer, elevated CA19-9 levels are indicative of tumor burden[18]. A higher tumor burden often necessitates more extensive resections, which can lead to secondary loss of pancreatic function-particularly in the remnant tissue following a Whipple procedure, a situation often compounded by underlying chronic pancreatitis. In this context, elevated preoperative CA19-9 levels may be a predictor of postoperative endocrine and exocrine insufficiency. In our analysis, a weak but statistically significant negative correlation was observed between CA19-9 levels and postoperative HbA1c values. On the other hand, in our study, increased BMI was significantly associated with elevated postoperative HbA1c levels. This finding aligns with existing evidence that obesity impairs glycemic control by exacerbating insulin resistance[19]. In one study, patients who underwent pancreatectomy and were classified as obese had a significantly higher incidence of postoperative diabetes than their non-obese counterparts (p=0.029)[15]. Therefore, a high preoperative BMI is associated with poorer glycemic profiles, both to the underlying insulin resistance that predisposes individuals to type 2 diabetes and to the increased physiological burden placed on the remaining pancreatic reserve following surgery.

Our findings demonstrated a negative correlation between postoperative HbA1c and C-peptide levels (correlation coefficient=-0.263). C-peptide serves as a marker of endogenous insulin secretion and reflects the functional integrity of pancreatic β-cells. In our study, the interaction between female sex and the Whipple procedure was significantly associated with higher postoperative C-peptide levels. It is well established that surgical resection of different regions of the pancreas can have varying impacts on the development of diabetes. According to the literature, the incidence of new-onset diabetes following PD (Whipple procedure) is generally lower than the incidence observed after distal pancreatectomy[15]. For instance, one study reported that among patients with normal pancreatic tissue, the incidence of diabetes after the Whipple procedure ranged from 10% and 24%, whereas after distal pancreatectomy it was more variable and often higher, ranging from 8% to as much as 60%[20,21]. Particularly in patients with underlying chronic pancreatitis, the incidence of diabetes has been shown to reach approximately 40% after the Whipple procedure and up to 85% after distal pancreatic resection[20]. However, a study by Lee et al.[16] suggested that the primary determinant of diabetes development after pancreatectomy is the patient's pre-resection insulin secretory capacity, while factors such as BMI or the extent of pancreatic tissue resected play a secondary role. In this context, the observation that female patients exhibited higher postoperative C-peptide levels may be attributed to a more robust β-cell reserve prior to surgery.

In a study by Kato et al.[22], the incidence of endocrine insufficiency was reported as 14%, while exocrine insufficiency

was found to be 24%. In a comprehensive review, Pathanki et al.[23] reported that the frequency of pancreatic exocrine insufficiency (PEI) following PD ranged from 38% to 93%. In our study, the incidence of postoperative diabetes was 24.4%, consistent with findings reported in the literature. We specifically evaluated postoperative HbA1c elevation and the rate of exocrine insufficiency, and found significant associations with parameters such as preoperative diabetes, CA19-9 levels, and BMI. Similarly, Kato et al.[22] highlighted that high preoperative HbA1c, elevated BMI, and reduced residual pancreatic volume are major risk factors for both endocrine and exocrine pancreatic insufficiencies. However, in contrast to their findings, our study revealed a negative association of both preoperative diabetes and CA19-9 levels with postoperative HbA1c elevation.

In our study, the postoperative diabetes incidence was 24.4%, which is consistent with findings reported in the literature. We specifically evaluated postoperative HbA1c elevation and the rate of exocrine insufficiency, demonstrating significant associations with parameters such as preoperative diabetes, CA19-9 levels, and BMI. Similarly, Kato et al.[22] highlighted that high preoperative HbA1c, elevated BMI, and reduced residual pancreatic volume are major risk factors for both endocrine and exocrine insufficiency. However, in contrast to their findings, our study revealed a negative association between preoperative diabetes and CA19-9 levels with postoperative HbA1c elevation.

In a study conducted by Iwase et al.[24], significant differences in the postoperative trajectory of serum elastase-1 levels were observed depending on the type of surgical procedure performed. In patients who underwent TP, elastase levels were reported to decline rapidly, dropping below the lower limit of the normal range (120 ng/dL) by postoperative day 14. In contrast, in patients who underwent PD (Whipple) or distal pancreatectomy, elastase levels initially increased during the first 7 days-likely due to the presence of residual pancreatic tissue-but were reported to normalize within two weeks postoperatively. In their 2020 study evaluating the development of PEI following PD, Pathanki et al.[23] identified preoperative serum elastase levels as an important predictive marker. They reported that a significant postoperative decline in elastase levels was associated with the onset of PEI. Additionally, they noted that low postoperative elastase levels were linked to diabetes, lower BMI, and advanced age. Our study yielded similar findings regarding postoperative elastase levels; specifically, the presence of preoperative diabetes had a significant negative impact on elastase concentrations. Furthermore, a near-significant positive correlation was observed between patient age and elastase levels. Moreover, we chose to use serum elastase levels to assess exocrine insufficiency, as this parameter is not influenced by the administration of pancreatic enzyme supplements[25]. However, in our study, no significant differences in serum elastase levels were observed based on the type of surgical procedure performed, including TP. Therefore, we suggest that serum elastase may not be a fully reliable indicator of postoperative exocrine pancreatic function.

Study Limitations

Despite the strengths of this study, several limitations should be acknowledged. First, the retrospective design and reliance on hospital records may have introduced bias due to missing data, recording errors, or variability in follow-up documentation, which could affect data reliability. Additionally, the study's relatively small sample size and single-center setting limit the generalizability of the findings. Future multicenter, prospective studies with larger patient cohorts are warranted to more clearly elucidate the predictive value of the investigated parameters.

Conclusion

In conclusion, preoperative parameters such as CA19-9 levels, diabetes, and BMI should be carefully considered during postoperative follow-up, as they are associated with an increased risk of subsequent endocrine and exocrine insufficiency after pancreatic surgery. Based on these findings, patients with elevated CA19-9 and obesity may benefit from closer monitoring and early intervention targeting pancreatic function in the postoperative period. Furthermore, routine follow-up of HbA1c and elastase levels after surgery is crucial for the early detection of emerging diabetes and malabsorption.

Ethics

Ethics Committee Approval: This study approved by Aydın Adnan Menderes Unıversity Ethic Committee (approval no: 8, date: 25.07.2024).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practises: S.A., O.A., E.B.C., A.E., M.Ç.C., Concept: S.A., O.A., E.B.C., Design: S.A., A.E., M.Ç.C., V.T., Data Collection or Processing: S.A., O.A., E.B.C., Analysis

or Interpretation: S.A., O.A., E.B.C., Literature Search: S.A., M.Ç.C., V.T., Writing: S.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. Atkinson MA, Campbell-Thompson M, Kusmartseva I, Kaestner KH. Organisation of the human pancreas in health and in diabetes. Diabetologia. 2020;63:1966-73.
2. Pignatelli C, Campo F, Neroni A, Piemonti L, Citro A. Bioengineering the vascularized endocrine pancreas: a fine-tuned ınterplay between vascularization, extracellular-matrix-based scaffold architecture, and ınsulin-producing cells. Transpl Int. 2022;35:10555.
3. Adsay NV, Basturk O, Saka B, et al. Whipple made simple for surgical pathologists: orientation, dissection, and sampling of pancreaticoduodenectomy specimens for a more practical and accurate evaluation of pancreatic, distal common bile duct, and ampullary tumors. Am J Surg Pathol. 2014;38:480-93.
4. Whitcomb DC, Buchner AM, Forsmark CE. AGA clinical practice update on the epidemiology, evaluation, and management of exocrine pancreatic ınsufficiency: expert review. Gastroenterology. 2023;165:1292-301.
5. Lindkvist B. Diagnosis and treatment of pancreatic exocrine insufficiency. World J Gastroenterol. 2013;19:7258-66.
6. Vonderau JS, Desai CS. Type 3c: understanding pancreatogenic diabetes. JAAPA. 2022;35:20-4.
7. Ewald N, Bretzel RG. Diabetes mellitus secondary to pancreatic diseases (Type 3c) -- are we neglecting an important disease? Eur J Intern Med. 2013;24:203-6.
8. Enríquez-Navascués JM, Borda N, Lizerazu A, et al. Patterns of local recurrence in rectal cancer after a multidisciplinary approach. World J Gastroenterol. 2011;17:1674-84.
9. Katsenos S, Becker HD. Recurrent respiratory papillomatosis: a rare chronic disease, difficult to treat, with potential to lung cancer transformation: apropos of two cases and a brief literature review. Case Rep Oncol. 2011;4:162-71.
10. Satman I, Omer B, Tutuncu Y, et al. Twelve-year trends in the prevalence and risk factors of diabetes and prediabetes in Turkish adults. Eur J Epidemiol. 2013;28:169-80.
11. Knop FK, Vilsbøll T, Holst JJ. Incretin-based therapy of type 2 diabetes mellitus. Curr Protein Pept Sci. 2009;10:46-55.
12. Cúrdia Gonçalves T, Capela TL, Cotter J. Nutrition in pancreatic diseases: a roadmap for the gastroenterologist. GE Port J Gastroenterol. 2023;31:1-13.
13. Keskin Ö, Balcı B. Diabetes mellitus and cardiovasculary complications. Kafkas J Med Sci. 2011;2:81-5.
14. Hart PA, Bellin MD, Andersen DK, et al. Type 3c (pancreatogenic) diabetes mellitus secondary to chronic pancreatitis and pancreatic cancer. Lancet Gastroenterol Hepatol. 2016;1:226-37.
15. Kim KJ, Jeong CY, Jeong SH, et al. Pancreatic diabetes after distal pancreatectomy: incidence rate and risk factors. Korean J Hepatobiliary Pancreat Surg. 2011;15:123-7.
16. Lee BW, Kang HW, Heo JS, et al. Insulin secretory defect plays a major role in the development of diabetes in patients with distal pancreatectomy. Metabolism. 2006;55:135-41.
17. Choe SA, Kim JY, Ro YS, Cho SI. Women are less likely than men to achieve optimal glycemic control after 1 year of treatment: a multi-level analysis of a Korean primary care cohort. PLoS One. 2018;13:e0196719.
18. Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: an evidence-based appraisal. J Gastrointest Oncol. 2012;3:105-19.
19. Hutchins RR, Hart RS, Pacifico M, Bradley NJ, Williamson RC. Long-term results of distal pancreatectomy for chronic pancreatitis in 90 patients. Ann Surg. 2002;236:612-8.
20. Huang JJ, Yeo CJ, Sohn TA, et al. Quality of life and outcomes after pancreaticoduodenectomy. Ann Surg. 2000;231:890-8.
21. Lemaire E, O'Toole D, Sauvanet A, Hammel P, Belghiti J, Ruszniewski P. Functional and morphological changes in the pancreatic remnant following pancreaticoduodenectomy with pancreaticogastric anastomosis. Br J Surg. 2000;87:434-8.
22. Kato T, Watanabe Y, Oshima Y, et al. Long-term outcomes and risk factors of pancreatic insufficiency after a pancreatoduodenectomy: a retrospective study. Surgery. 2024;176:880-9.
23. Pathanki AM, Attard JA, Bradley E, et al. Pancreatic exocrine insufficiency after pancreaticoduodenectomy: current evidence and management. World J Gastrointest Pathophysiol. 2020;11:20-31.
24. Iwase K, Miyata M, Tanaka Y, Izukura M, Nakaba H, Matsuda H. Serial changes in plasma levels of pancreatic elastase 1 after pancreatic surgeries. Res Exp Med (Berl). 1995;195:93-100.
25. Vanga RR, Tansel A, Sidiq S, El-Serag HB, Othman MO. Diagnostic performance of measurement of fecal elastase-1 in detection of exocrine pancreatic ınsufficiency: systematic review and meta-analysis. Clin Gastroenterol Hepatol. 2018;16:1220-8.

DOI: 10.4274/anatoljmed.2025.20082
Anatol J Gen Med Res 2025;35(3):353-367

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Horizoning the Convergence of Artificial Intelligence and Healthcare: An Exploratory Analysis Using Latent Semantic Analysis

Yapay Zeka ve Sağlık Hizmetlerinin Yakınsamasını Ufukta Görmek: Gizli Anlamsal Analiz Kullanarak Yapılan Keşifsel Bir Analiz

Hüseyin Demir[1], Süleyman Mertoğlu[2]

[1]İzmir Katip Çelebi University Faculty of Economics and Administrative Sciences, Department of Health Management, İzmir, Türkiye
[2]İzmir Provincial Health Directorate Personnel Services Directorate, İzmir, Türkiye

Cite as: Demir H, Mertoğlu S. Horizoning the convergence of AI and healthcare: an exploratory analysis using latent semantic analysis. Anatol J Gen Med Res. 2025;35(3):353-367

Abstract

Objective: This study aims to provide insights for healthcare stakeholders by applying latent semantic analysis to the most-cited scientific publications in artificial intelligence and healthcare published over the past decade.

Methods: Publications were retrieved from the Web of Science database, focusing on the 1.000 most-cited papers published between 2015 and 2025. Latent semantic analysis was employed for text analysis, encompassing corpus creation, text preprocessing, tokenization, lowercasing, stopword removal, stemming, lemmatization, term-document matrix construction, weighting, singular value decomposition, dimensionality reduction, semantic construction, and interpretation. Model performance was assessed using singular values, explained variance, and cumulative variance, with the optimal number of dimensions determined to be 160.

Results: The latent semantic analysis model effectively uncovered the underlying semantic relationships within the dataset. The gradual decline in decomposition values supported the appropriateness of the model's structure and dimensionality. The analysis revealed that artificial intelligence and healthcare are converging on two primary clinical themes: deep learning and predictive applications. The deep learning theme reflects the training of artificial intelligence systems with patient data, whereas the predictive theme emphasizes the use of artificial intelligence in diagnostic and therapeutic decision-making. Additionally, the distinct semantic positioning of the coronavirus disease-2019 theme highlighted the model's ability to differentiate thematic clusters accurately.

Conclusion: Findings indicate a clear convergence between artificial intelligence and healthcare, demonstrating increasing interconnectivity. Advances in artificial intelligence are increasingly influencing clinical decision-support systems and patient-centered applications. Policymakers should develop strategic frameworks to ensure the safe, ethical, and effective integration of AI into healthcare, particularly in clinical decision-support and patient care. Such frameworks must emphasize regulatory standards, professional training, data privacy, and patient safety to maximize the benefits of AI while mitigating potential risks.

Keywords: Healthcare, artificial intelligence, text analytics, natural language processing, latent semantic analysis

Öz

Amaç: Bu çalışma, son on yılda yapay zeka ve sağlık alanında yayımlanan ve en çok atıf alan bilimsel yayınlara gizli anlamsal analiz uygulayarak sağlık hizmeti paydaşlarına yönelik içgörüler sağlamayı amaçlamaktadır.

Address for Correspondence/Yazışma Adresi: Hüseyin Demir, İzmir Katip Çelebi University Faculty of Economics and Administrative Sciences, Department of Health Management, İzmir, Türkiye
E-mail: huseyin.demir@ikc.edu.tr
ORCID ID: orcid.org/0000-0002-8990-7228

Received/Geliş tarihi: 10.11.2025
Accepted/Kabul tarihi: 18.11.2025
Published date/Yayınlanma tarihi: 30.12.2025

Öz

Yöntem: Yayınlar, Web of Science veri tabanından taranmış olup 2015-2025 yılları arasında yayınlanan ve en çok atıf alan 1,000 makale belirlenmiştir. Metin analitiği için gizli anlamsal analiz kullanılmıştır. Bu süreç, veri kümesi oluşturma, metin ön işleme, tokenizasyon, küçük harfe dönüştürme, durak kelimelerinin çıkarılması, kök bulma, terim-belge matrisi oluşturma, ağırlıklandırma, tekil değer ayrışımı, boyut indirgeme, anlamsal yapı oluşturma ve yorumlamayı kapsamaktadır. Model performansı, tekil değerler, açıklanan varyans ve kümülatif varyans kullanılarak değerlendirilmiş ve model için optimal boyut sayısı 160 olarak elde edilmiştir.

Bulgular: Anlamsal model, veri kümesindeki temel anlamsal ilişkileri etkili bir şekilde ortaya çıkarmıştır. Tekil değerlerdeki sistematik düşüş, modelin yapısının ve boyut sayısının uygunluğunu desteklemiştir. Analizler, yapay zeka ve sağlık alanlarının derin öğrenme ve öngörüsel uygulamalar olmak üzere iki ana klinik tema etrafında yakınsadığını göstermiştir. Derin öğrenme teması, yapay zeka sistemlerinin hasta verileriyle eğitilmesini ifade ederken, öngörüsel uygulamalar teması ise yapay zekanın klinik karar verme süreçlerinde tanısal ve tedavi amaçlı tahminlerde kullanımını vurgulamaktadır. Ayrıca, koronavirüs hastalığı-2019 temasının anlamsal uzayda farklı konumlanması, modelin tematik kümeleri doğru şekilde ayırt edebilme yeteneğini göstermiştir.

Sonuç: Bulgular, yapay zeka ve sağlık alanları arasında güçlü bir yakınsama eğilimi olduğunu ve bu alanların giderek daha fazla iç içe geçtiğini göstermektedir. Yapay zeka alanındaki teknolojik ilerlemeler, klinik karar destek sistemlerini ve hasta odaklı uygulamaları giderek daha fazla etkilemektedir. Bu nedenle, politika yapıcıların, yapay zekanın sağlık hizmetlerine güvenli, etik ve etkin entegrasyonunu sağlamak amacıyla stratejik bir çerçeve geliştirmesi gerekmektedir. Bu tür stratejik çerçeveler, yapay zekanın sağlık alanındaki faydalarını maksimize ederken, potansiyel riskleri minimize etmek amacıyla düzenleyici standartlar, profesyonel eğitim, veri gizliliği ve hasta güvenliğine öncelik vermelidir.

Anahtar Kelimeler: Sağlık hizmetleri, yapay zeka, metin analitiği, doğal dil işleme, gizli anlamsal analiz

Introduction

Modern healthcare is evolving from a volume-based model to a value-based one, emphasizing the use of health data to optimize resources, enhance quality of care, increase patient satisfaction, and improve health outcomes[1]. This transition highlights the growing importance of data analytics in healthcare, as it helps uncover valuable patterns in large, complex datasets to address new challenges. In this context, artificial intelligence (AI) has become a powerful tool in healthcare[2]. According to the common definition of AI as the simulation of human cognition, AI builds on advances in predictive modeling approaches, such as machine learning (ML), through which computer algorithms learn from training data without human guidance to enable algorithmic decision-making[2]. As technology advances, AI's ability to process and analyze data is becoming increasingly powerful, making it a key instrument in transforming healthcare[3]. Recent studies have focused on promoting innovative algorithms that integrate novel solutions into health systems to provide more robust healthcare[4]. Therefore, the integration of AI and big data has created opportunities to incorporate evidence-based decision-making more extensively into the healthcare system. Considering this horizon, novel developments offer new potential, such as augmented decision-making support for physicians across the spectrum of care, thereby contributing to enhanced patient safety and improved health outcomes[5,6].

In a technology-driven world, healthcare has witnessed breakthroughs and exponential advancements[7]. Data, advanced analytics, and computational tools have created a new paradigm for managing health data to predict disease occurrence and treatment outcomes[8]. In this context, AI has emerged through technological progress and is poised to revolutionize healthcare by enabling advanced algorithmic decision-making in domain-specific applications. These applications aim to mimic human thinking and cognitive functions, thereby transforming healthcare delivery. As is well known, healthcare and health services have traditionally been provided by experts and specialized institutions. Therefore, the convergence of AI and healthcare providers represents a paradigm-changing development, enabled by the increasing availability of healthcare data and rapid advances in analytic techniques. In particular, early detection and diagnosis, treatment optimization, and outcome prediction and prognostic evaluation are the three major areas of AI applications in healthcare[9]. AI applications can be understood through aggregated healthcare data, which produce more powerful models capable of automating diagnosis and enhancing personalized medicine solutions[10]. Generally, AI-driven technologies can be categorized into two main approaches. The first is the ML approach, which analyzes structured data such as imaging and genetic data. The second involves natural language processing (NLP) techniques, which extract information from unstructured data, such as clinical notes, to augment structured medical data for clinical use[9].

AI applications in healthcare domains generally utilize supervised learning techniques, including support vector machines, neural networks, decision trees, random forests, linear and logistic regression, Naive Bayes, discriminant analysis, and nearest neighbor algorithms. Clinically significant gains have been achieved, particularly through supervised learning methods such as neural networks and support vector machines, rather than unsupervised approaches such as clustering and principal component analysis[9]. Therefore, it is paramount to combine data with advanced systems, applications, and algorithms to generate meaningful knowledge[3]. The increasing volume and complexity of medical data underscore the rationale for using AI-driven technologies, especially deep learning and ML, in healthcare. The development of new AI applications based on deep learning demonstrates AI's substantial potential to transform healthcare. Kaku[11], author of Physics of the Future urges us to recognize the limits of AI in healthcare with the following words:

"In the near future, you will simply approach a wall-mounted screen and consult a robo-doctor. A friendly face will ask you many questions. Then you will answer orally rather than in writing. After a few initial questions, the robo-doctor will diagnose your disease based on the most robust clinical experience of doctors worldwide."

AI applications, particularly those based on deep learning, are already revolutionizing healthcare. For instance, International Business Machines's Watson platform enables healthcare providers to analyze complex health data more accurately and cost-effectively than human experts[3]. Numerous studies have demonstrated the effectiveness of AI in medical applications such as image-based diagnostics[12]. One study showed that AI could detect diabetic retinopathy, a leading cause of blindness among diabetic patients[13]. In the United States, AI models trained on over 128.000 images achieved excellent diagnostic performance in detecting diabetic retinopathy[14]. AI has also been utilized to predict pediatric diseases such as asthma and pneumonia, achieving results comparable to human expertise[12].

Pain prediction represents another promising domain for AI-driven algorithms, as pain directly affects patients' quality of life and informs the selection of optimal treatments. Given the potential bias in pain assessment tools arising from patient and physician subjectivity, Liu et al.[15] emphasize the importance of machine-based pain assessment that uses facial expression data to provide more accurate, less biased evaluations. In radiology, Li et al.[16] introduced a deep learning-based model combining multimodal brain data to improve diagnostic accuracy. The model, trained on magnetic resonance imaging (MRI) and positron emission tomography (PET) images, predicts missing PET patterns from MRI data to assist diagnosis of neurodegenerative diseases such as Alzheimer's disease. Similarly, in the field of surgery, robotic systems such as da Vinci have been used for minimally invasive procedures and have been associated with safe and effective patient outcomes[17-20]. As AI's predictive models continue to improve, even greater advancements in healthcare are expected in the near future.

Advancements in data, analytics, and computational tools are transforming how health data are utilized to predict disease occurrences and treatment outcomes[8]. AI plays a critical role in this transformation by harnessing data to guide medical decisions. However, the ethical and legal implications of AI in healthcare remain insufficiently explored[21]. While AI offers technical advantages, it also raises concerns regarding patient safety and privacy[22]. Issues such as fairness, autonomy, and accountability pose significant challenges to the integration of AI into healthcare, as legal systems continue to grapple with how to address these concerns[21]. A notable research gap exists concerning liability for AI-driven decisions that result in patient harm, underscoring the urgent need for clear ethical and legal guidelines governing AI applications in healthcare. Because AI requires access to sensitive patient data, the risk of misuse or data breaches increases[23]. According to the general data protection regulation, individuals have the right not to be subject to decisions based solely on automated processing (such as AI-driven systems) that could significantly affect them. Although AI can assist clinicians with clinical decision-making, it is unlikely to replace human clinicians in the near future[9,24]. The role of human clinicians, especially in critical decision-making, remains indispensable since AI is not yet capable of replicating human expertise[25,26].

It is unsurprising that little has been reported in the literature regarding liability for harm caused by AI-driven algorithmic decisions; this may call into question the adequacy of existing legal frameworks[27,28]. Hence, AI inevitably introduces various ethical and legal concerns, along with paradigmatic shifts in healthcare, in which harm to human health is unacceptable. Therefore, it is of paramount importance to analyze and evaluate AI-driven progress within an ethical and legal framework. Overall, these considerations highlight the need to clarify the legal nature of AI, a challenging responsibility

that includes identifying who bears liability for medical malpractice and ensuring appropriate compensation for patients harmed. Accordingly, it is essential that healthcare authorities-such as the Ministry of Health-along with governmental and regulatory bodies act responsibly, monitor emerging problem areas, and establish governance mechanisms to prevent conflicting outcomes. Therefore, scientific approaches developed using AI and healthcare research should form the foundation for decision-making, thereby highlighting the importance of relying on evidence-based findings. NLP techniques offer valuable opportunities for policymakers through text analytics. This study aims to generate insights for healthcare stakeholders by applying latent semantic analysis (LSA) to the most-cited scientific publications from the past decade in AI and healthcare. The study is considered original in both its focus and methodology and is expected to make a meaningful contribution to the literature.

Materials and Methods

Objective

The aim of the study is to generate insights for healthcare stakeholders by applying LSA to the most-cited scientific publications from the past decade in AI and healthcare.

Research Question

What insights and future horizons emerge from the convergence of AI and healthcare in scholarly literature?

Data Source and Research Unit

The data used in this study were obtained from the Web of Science (WoS) Core Collection, selected for its wide recognition and established reliability as a source of scientific literature, particularly for bibliometric and text-analytics research[29]. The research was conducted on the assumption that publications retrieved from the WoS database adequately represent developments in the scientific domain, allowing the study to effectively address its research question. Therefore, the research unit of this study comprises the 1.000 most-cited articles published in the last ten years (2015-2025), identified in the database.

Search Strategy

The publication retrieval process was completed in three stages within the WoS database. In the first stage, study titles were searched using specific keywords related to AI. The keywords related to AI were defined as follows: artificial intelligence, AI, machine learning, deep learning, neural networks, artificial neural networks, convolutional neural networks, recurrent neural networks, reinforcement learning, supervised learning, unsupervised learning, semi-supervised learning, transfer learning, federated learning, explainable AI, generative AI, natural language processing, computer vision, speech recognition, pattern recognition, knowledge representation, expert systems, cognitive computing, symbolic AI, evolutionary algorithms, swarm intelligence, fuzzy logic, rule-based systems, agent-based modeling, data mining, data preprocessing, feature extraction, feature selection, predictive modeling, classification, clustering, regression analysis, dimensionality reduction, big data, data visualization, knowledge discovery, decision trees, random forest, support vector machine, gradient boosting, ensemble learning, text mining, sentiment analysis, topic modeling, latent Dirichlet allocation, latent semantic analysis, word embeddings, Word2Vec, GloVe, transformer models, BERT, GPT models, text classification, named entity recognition, machine translation, question answering, speech-to-text, text summarization, decision-support systems, predictive analytics, intelligent automation, robotics, autonomous vehicles, recommender systems, chatbots, virtual assistants, image recognition, fraud detection, anomaly detection, medical diagnosis, smart healthcare, precision medicine, remote monitoring, health informatics, personalized medicine, clinical decision-support, AI ethics, algorithm, training data, test data, validation set, model accuracy, overfitting, underfitting, hyperparameter tuning, cross-validation, loss function, gradient descent, optimization, feature engineering, model interpretability, bias and variance, computational complexity, algorithmic bias, data privacy, data security, fairness, transparency, accountability, explainability, trustworthy AI, responsible AI, human-centered AI, AI governance, artificial general intelligence, quantum machine learning, edge AI, internet of things, internet of medical things, cyber-physical systems, human-AI interaction, human-computer interaction, digital twins, cognitive robotics, sustainable AI, green AI.

In the second stage, a search was conducted using keywords related to healthcare. The keywords defined for the healthcare domain included healthcare, health care, health system, healthcare system, health services, public health, global health, primary healthcare, secondary healthcare, tertiary healthcare, preventive healthcare, curative healthcare, rehabilitation services, health policy, health management,

health administration, health economics, health financing, health insurance, universal health coverage, health equity, health disparity, access to healthcare, quality of care, patient safety, health outcomes, health indicators, disease prevention, health promotion, telehealth, telemedicine, mobile health, mHealth, digital health, eHealth, virtual healthcare, remote monitoring, home healthcare, hospital care, primary care, ambulatory care, emergency care, intensive care, long-term care, palliative care, nursing care, mental health care, behavioral health, dental care, maternal health, child health, reproductive health, geriatric care, chronic disease management, preventive services, occupational health, environmental health, nutrition services, laboratory services, radiology services, pharmacy services, surgical services, health information systems, hospital information system, electronic health records, electronic medical records, health informatics, clinical decision-support systems, artificial intelligence in healthcare, machine learning in medicine, predictive analytics in healthcare, medical devices, digital therapeutics, wearable technology, smart healthcare, internet of medical things, precision medicine, personalized medicine, robotics in healthcare, remote diagnostics, virtual reality in healthcare, augmented reality in medicine, health expenditure, cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis, economic evaluation, health technology assessment, resource allocation, efficiency in healthcare, healthcare cost, payment systems, reimbursement, sustainability in healthcare, financial risk protection, value-based healthcare, hospital efficiency, budget impact analysis, health workforce, healthcare professionals, physicians, nurses, pharmacists, allied health professionals, workforce planning, health leadership, hospital management, healthcare quality management, performance measurement, patient-centered care, integrated care, care coordination, patient satisfaction, patient experience, healthcare marketing, hospital accreditation, strategic management in healthcare, medical ethics, bioethics, patient rights, informed consent, data privacy in healthcare, confidentiality, medical law, healthcare regulation, health governance, ethical decision-making, health system strengthening, sustainable health systems, global burden of disease, health security, epidemics, pandemics, COVID-19, vaccine distribution, health resilience, social determinants of health, environmental determinants, climate change and health, one health, sustainable development goals, planetary health, health research, clinical trials, observational studies, epidemiology, evidence-based medicine, health data analytics, big data in healthcare, health statistics, population health, health outcome measurement, patient-reported outcomes, quality indicators, data-driven healthcare. In the third and final stage, the publications obtained from the previous steps were combined using the AND operator, thereby completing the retrieval process. The inclusion criteria required that publications be written in English, categorized as research articles, published in citation-indexed journals (SCI, SSCI, SCI-Expanded, ESCI, or SCOPUS), and classified in health- or medical-related WoS categories. These parameters guided the final selection of studies included in the analysis.

Statistical Analysis

Latent Semantic Analysis

The ever-growing size and complexity of textual data make it increasingly challenging to identify meaningful documents and patterns, not only in other disciplines[30] but also in health and medicine. In this context, text analytics has recently emerged as a strategic tool for analyzing textual data in healthcare and medical research. Text analytics is a multi-stage process that aims to derive meaningful insights from unstructured textual data. The methods used in text analytics are primarily implemented using NLP techniques. When the objective of NLP is to understand the contextual meaning of words, semantic modeling approaches are employed. Within this framework, LSA represents a semantic modeling method that applies statistical and mathematical techniques to analyze textual data and extract significant insights[31]. Therefore, under the applied statistical framework, the primary goal of LSA is to uncover the contextual meaning of words within a large text corpus[32]. Originating in the 1980s, the LSA method was initially designed as an information retrieval technique[33], but it was later introduced into psychological research as a theory and method for discovering and representing the meanings of words[34,35]. Similarly, Landauer and Dumais[36] proposed that LSA constitutes a fundamental computational theory of knowledge acquisition and representation. Through its use, LSA can capture word-passage, passage-passage, and sentence-sentence relationships in ways that align with human cognition[32]. In this regard, LSA stands out as a fully automated method that employs mathematical and statistical techniques to infer information about the contextual use of words in textual data[37]. Foltz[38] discussed the applicability of the LSA method in text-based research in three domains, while Shen and Ho[39] demonstrated its usefulness in technology-assisted higher education, confirming that LSA is a highly effective tool in text analytics.

The workflow used for the LSA analysis in this study is shown in Figure 1.

The top 1.000 most-cited studies published in the last ten years were retrieved from the WoS database and exported to Microsoft Excel. A preliminary review of the publication titles revealed that the dataset contained no duplicate entries. The final dataset was then imported into the R programming environment[40] for LSA. The LSA procedure was conducted based on study abstracts, following similar approaches presented in the literature. Accordingly, the LSA process began with obtaining the text data from the WoS database. Before constructing the LSA model, the text corpus underwent data preprocessing. During this process, several operations were performed on the textual data, including tokenization, stopword removal, and lemmatization/stemming.

The first step in the LSA analysis involved transforming the text into a matrix in which each row represented a unique word and each column represented a passage or another content segment. The cells at the intersections of rows and columns indicated the frequency with which each word appeared in the text. In this step, a document-term matrix (DTM) was created, allowing textual data to be represented as vectors. Subsequently, word frequencies in the text data were transformed and weighted to reflect the relative importance of words within the corpus. For this purpose, a term frequency-inverse document frequency weighting scheme was applied to the DTM.

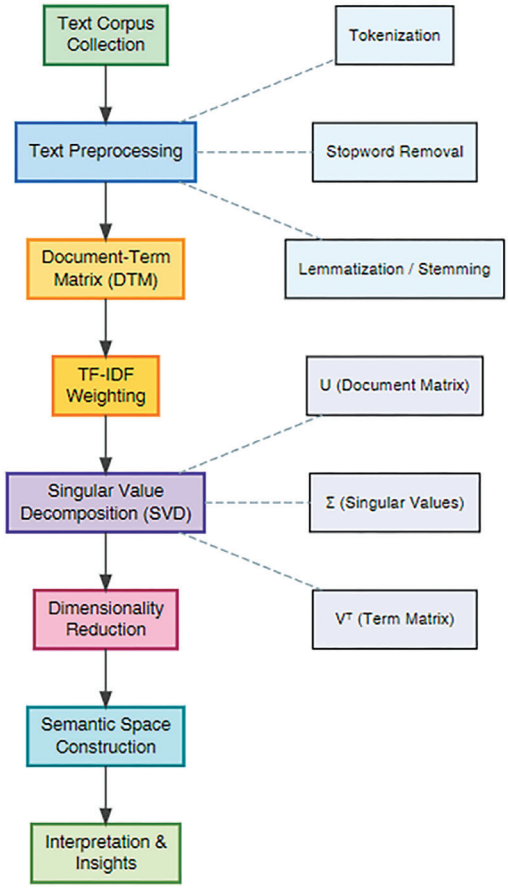


Figure 1. Flow diagram of the LSA analysis

LSA: Latent semantic analysis, TF: Term frequency, IDF: Inverse document frequency

The second step of the LSA involved applying singular value decomposition (SVD) to the matrix. As LSA is based on SVD-a mathematical matrix decomposition technique similar to factor analysis[32]-this operation can be interpreted as a form of factor extraction. In the SVD method, a rectangular matrix is decomposed into the product of three distinct matrices: one defining the rows as vectors of derived orthogonal factor values; another defining the columns in a similar manner; and a third, diagonal matrix containing the singular values that scale the two orthogonal matrices. When these three matrices are multiplied, the original matrix is reconstructed. Thus, after the weighting process, SVD was applied to the final matrix to perform decomposition. This process resulted in the decomposition of the original matrix into its constituent matrices. Using SVD, singular values and explained variances were obtained. The optimal number of dimensions to optimize model performance was determined graphically, thereby defining the appropriate dimensional space for the textual data and achieving dimensionality reduction.

Following these stages, construction of the semantic space, visualization, interpretation, and generation of insights were carried out as part of the LSA process. In the R environment[40], several packages are available for performing semantic analysis. In this study, the lsa and tm packages were used for modeling, while ggplot2 was employed for visualization[41]. Descriptive findings regarding the publication dataset were summarized as frequencies, whereas analytical results from the LSA process were presented as exploratory graphical visualizations to facilitate deeper interpretation.

Results

Descriptive findings regarding the publications are presented in Table 1.

According to Table 1, the most highly cited studies published in the past decade at the intersection of AI and healthcare

Table 1. Descriptive findings on the publications (top 25 publications)

Article title	Source title	Publication year	Times cited, all databases	WoS categories
A guide to deep learning in healthcare	Nature Medicine	2019	2714	Biochemistry and molecular biology, cell biology, medicine, research and experimental
Machine learning in medicine	Circulation	2015	2568	Cardiac and cardiovascular systems, peripheral vascular disease
Scalable and accurate deep learning with electronic health records	NPJ Digital Medicine	2018	1779	Health care sciences and services, medical informatics
AACR project GENIE: powering precision medicine through an international consortium	Cancer Discovery	2017	1544	Oncology
The future of digital health with federated learning	NPJ Digital Medicine	2020	1518	Health care sciences and services, medical informatics
COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks	Physical and Engineering Sciences in Medicine	2020	1472	Engineering, biomedical, radiology, nuclear medicine and medical imaging
Deep learning for health informatics	IEEE Journal of Biomedical and Health Informatics	2017	1344	Computer science, information systems, computer science, interdisciplinary applications, mathematical and computational biology, medical informatics
Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions	Journal of Thoracic Disease	2020	1062	Respiratory system
Classification of the cutaneous manifestations of COVID-19: a rapid prospective nationwide consensus study in Spain with 375 cases	British Journal of Dermatology	2020	1054	Dermatology
Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy	Radiology	2020	1046	Radiology, nuclear medicine and medical imaging
Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices	NPJ Digital Medicine	2018	1036	Health care sciences and services, medical informatics
Explainability for artificial intelligence in healthcare: a multidisciplinary perspective	BMC Medical Informatics and Decision Making	2020	992	Medical informatics
Machine learning in medicine: a practical introduction	BMC Medical Research Methodology	2019	883	Health care sciences and services
Treatable traits: toward precision medicine of chronic airway diseases	European Respiratory Journal	2016	856	Respiratory system
Personalized *in vitro* and *in vivo* cancer models to guide precision medicine	Cancer Discovery	2017	830	Oncology
The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care	Nature Medicine	2018	830	Biochemistry and molecular biology, cell biology, medicine, research and experimental

Table 1. Continued

Article title	Source title	Publication year	Times cited, all databases	WoS categories
Artificial intelligence (AI) applications for COVID-19 pandemic	Diabetes & Metabolic Syndrome-Clinical Research & Reviews	2020	799	Endocrinology and metabolism
Artificial intelligence-enabled rapid diagnosis of patients with COVID-19	Nature Medicine	2020	776	Biochemistry and molecular biology, cell biology, medicine, research and experimental
Federated learning for healthcare informatics	Journal of Healthcare Informatics Research	2021	724	Computer science, information systems, health care sciences and services, medical informatics
The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database	NPJ Digital Medicine	2020	701	Health care sciences and services, medical informatics
Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning	Medical Image Analysis	2020	689	Computer science, artificial intelligence, computer science, interdisciplinary applications, engineering, biomedical, radiology, nuclear medicine and medical imaging
Federated learning of predictive models from federated electronic health records	International Journal of Medical Informatics	2018	676	Computer science, information systems, health care sciences and services, medical informatics
A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)	European Radiology	2021	667	Radiology, nuclear medicine and medical imaging
CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19)	European Radiology	2020	616	Radiology, nuclear medicine and medical imaging
Deep learning COVID-19 features on CXR using limited training data sets	IEEE Transactions on Medical Imaging	2020	612	Computer science, interdisciplinary applications, engineering, biomedical, engineering, electrical and electronic, imaging science and photographic technology, radiology, nuclear medicine and medical imaging

WoS: Web of Science, AACR: American Association for cancer research, GENIE: Genomics evidence neoplasia information exchange, SEIR: Susceptible, exposed, infectious, recovered/removed, CT: Computed tomography, CXR: Chest X-ray, NPJ: Nature Partner Journal, IEEE: Institute of Electrical and Electronics Engineers, BMC: BioMed Central, COVID-19: Coronavirus disease-2019

primarily focus on the integration of deep learning, ML, and data-driven approaches into clinical applications. Studies published in high-impact journals such as Nature Medicine and Circulation indicate a highly multidisciplinary area of impact, encompassing fields including medical informatics, biomedical sciences, and oncology. A semantic examination reveals that research employing electronic health records to develop deep federated learning methods, along with AI applications designed for the diagnosis and prediction of coronavirus disease-2019 (COVID-19), has become particularly prominent. Furthermore, emerging approaches related to explainable AI and automation-based diagnostic systems demonstrate a clear orientation toward trustworthy, transparent, and clinically meaningful solutions. Overall,

Table 2. Findings on factor loadings (the first 25 factors and their loadings)

Dimension	Singular value	Explained variance	Cumulative variance
1	160.8859509	0.090776614	0.090776614
2	92.42407897	0.029957668	0.120734281
3	80.08099461	0.022490373	0.143224654
4	69.12955464	0.016759662	0.159984316
5	66.49003804	0.015504255	0.17548857
6	60.08236854	0.012659945	0.188148516
7	55.41135584	0.010768006	0.198916521
8	53.57346048	0.010065541	0.208982063
9	52.86730212	0.00980194	0.218784003
10	49.70429226	0.008664141	0.227448144
11	48.91963139	0.008392746	0.23584089
12	48.15019038	0.008130809	0.243971699
13	46.6449687	0.007630401	0.251602099
14	45.9798715	0.007414353	0.259016452
15	45.14492498	0.007147524	0.266163976
16	44.01776693	0.006795067	0.272959043
17	43.20480262	0.006546389	0.279505432
18	42.8560118	0.006441118	0.285946549
19	42.28367015	0.006270224	0.292216774
20	41.5990027	0.00606881	0.298285584
21	41.04925133	0.005909466	0.30419505
22	40.44950591	0.005738048	0.309933098
23	39.75590355	0.00554295	0.315476048
24	39.58423607	0.005495184	0.320971233
25	39.1316632	0.005370248	0.326341481

The LSA model includes SVD values corresponding to 160 dimensions. The first 25 components presented in this table account for approximately 33% of the total variance, while consideration of all components indicates that the model explains at least 80% of the variance overall

LSA: Latent semantic analysis, SVD: Singular value decomposition

the findings suggest a strong convergence between the domains of AI and healthcare, indicating that these fields are becoming increasingly intertwined. Technological advances in AI are being progressively reflected in clinical decision-support systems and patient-centered applications, signaling a significant transformation in healthcare delivery. Findings related to the factor loadings of the LSA model are presented in Table 2.

As shown in Table 2, the singular value corresponding to the first dimension was 160.88, which is considerably higher than that of the second dimension. This indicates that the first dimension represents the most dominant semantic axis within the text space. In subsequent dimensions, the singular values gradually decrease, suggesting that each successive dimension carries less information than the preceding one, reflecting the principle of diminishing marginal contribution. This trend can be interpreted as evidence of successful SVD performance in the LSA model. According to the findings, the first dimension explains approximately 9% of the total variance in the model. When the first 10 dimensions are considered, they collectively account for about 23% of the variance, while the first 25 dimensions explain approximately 33% of the model's total variance. This clearly demonstrates that roughly one-third of the semantic structure of the analyzed texts is represented by the first 25 dimensions. In terms of cumulative variance, the initial dimensions contribute significantly to the model, while the contribution of dimensions beyond the 20th progressively decreases. This pattern suggests that the model is approaching a point of saturation. This trend is consistent with the cumulative variance curve presented in Figure 2. Therefore, the findings of the LSA model highlight that the initial components capture a substantial portion of the semantic structure embedded within the text corpus. A systematic decline in singular values further indicates the model's effective dimensionality-reduction performance, demonstrating that each component contributes meaningfully to the model's explanatory capacity. Collectively, these results confirm that the LSA model effectively distinguishes the semantic structures within the text and attains optimal performance as the number of dimensions increases. Findings regarding the explained variance, optimal dimensionality, and document similarity distribution of the LSA model are presented in Figure 2.

The explained variance graph in the upper-left section shows that as the number of dimensions increases, the cumulative variance rises significantly, eventually reaching a plateau at around 160 dimensions. At this point, the LSA model explains approximately 80% of the total variance. This finding indicates that the optimal number of dimensions provides a balanced trade-off between model complexity and semantic representability. Therefore, obtaining additional dimensions beyond this threshold does not yield any meaningful improvement in the model's performance or in the proportion of variance explained. The document similarity distribution in the upper-right graph demonstrates that a significant portion of the similarity scores between document pairs is concentrated around zero, suggesting substantial semantic diversity among the documents in the dataset. This finding

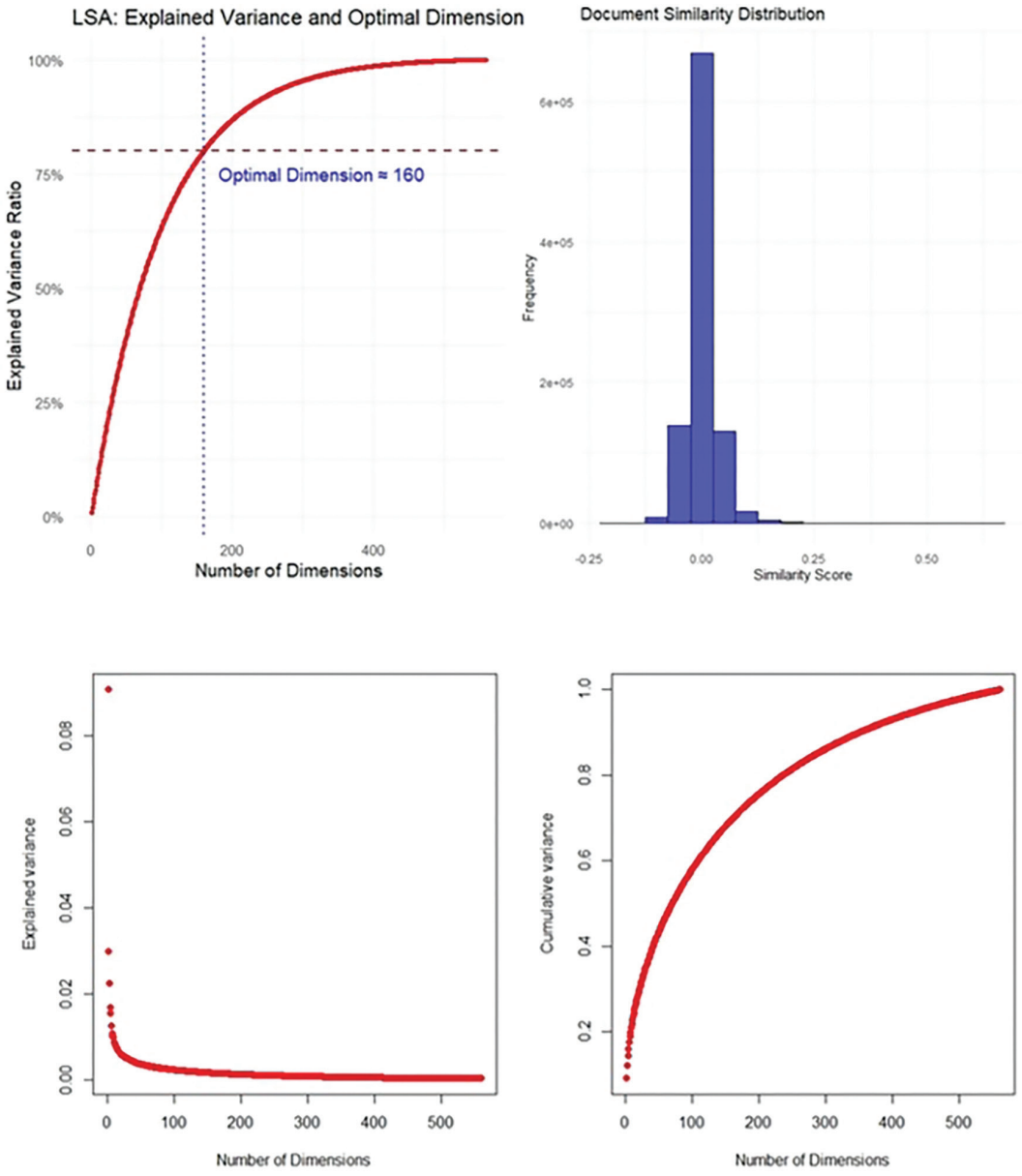


Figure 2. Performance findings related to the LSA model

LSA: Latent semantic analysis

indicates that the LSA model effectively distinguishes semantically distinct content within the corpus. The explained-variance graph in the lower-left section illustrates that the first components of the dataset account for a major portion of the variance in the LSA model, whereas the contributions of the subsequent components are markedly lower. This implies that the dominant semantic structure within the dataset is represented by a limited number of latent semantic components. The cumulative variance curve in the lower-right graph provides further clarity by showing that the incremental contribution of each additional dimension to the explained variance progressively decreases. Overall, the results demonstrate that the LSA model effectively captures the semantic structure of the dataset and produces outcomes that are dimensionally efficient. Accordingly, the number of dimensions obtained for the LSA model represents an optimal solution, effectively identifying semantic patterns while minimizing computational complexity. The correlation matrix illustrating the relationships among the documents is presented in Figure 3.

Figure 3 shows that correlations among documents vary considerably. This indicates that the documents included in

the LSA model are not highly correlated. The predominance of light-colored areas in the correlation matrix indicates that a substantial portion of the documents are either uncorrelated or weakly correlated. This finding is particularly important for the LSA model, as its main objective is to generate distinct semantic dimensions that differentiate between documents. Conversely, the presence of light-blue or light-red areas suggests that weak or even negative correlations may exist between certain pairs of documents. This implies that some documents may partially share common semantic patterns within the corpus. Overall, no high-level correlations were observed in the LSA model, indicating that no excessive semantic overlap or redundant components were produced. Consequently, the semantic dimensions derived from the LSA model are found to be meaningfully distinct from one another. On the other hand, the presence of minor correlations-although not statistically significant-suggests the emergence of shared themes or semantic clusters within the dataset. The subsequent findings related to the LSA model are presented in Figure 4.

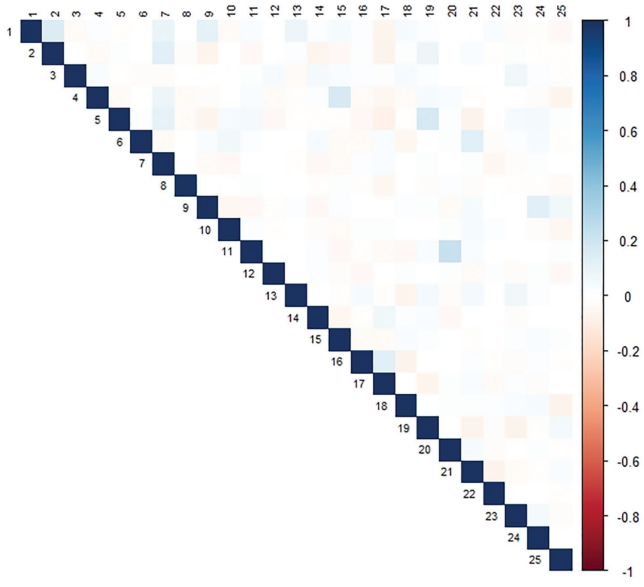


Figure 3. Document similarity matrix (the first 25 documents)

The upper-left graph shows changes in the number of publications over time. A noticeable increase in AI-related healthcare studies was observed in 2020, reflecting the exponential rise in AI research in healthcare during the COVID-19 pandemic. The upward trend in publication volume continued into 2021, but the number of publications appeared to stabilize thereafter, returning to pre-pandemic levels. Because this analysis is based on the top 1,000 most-cited publications, the observed trend must be compared with the overall publication trend to provide full contextual interpretation. The upper-right graph presents the two-dimensional representation of five document clusters

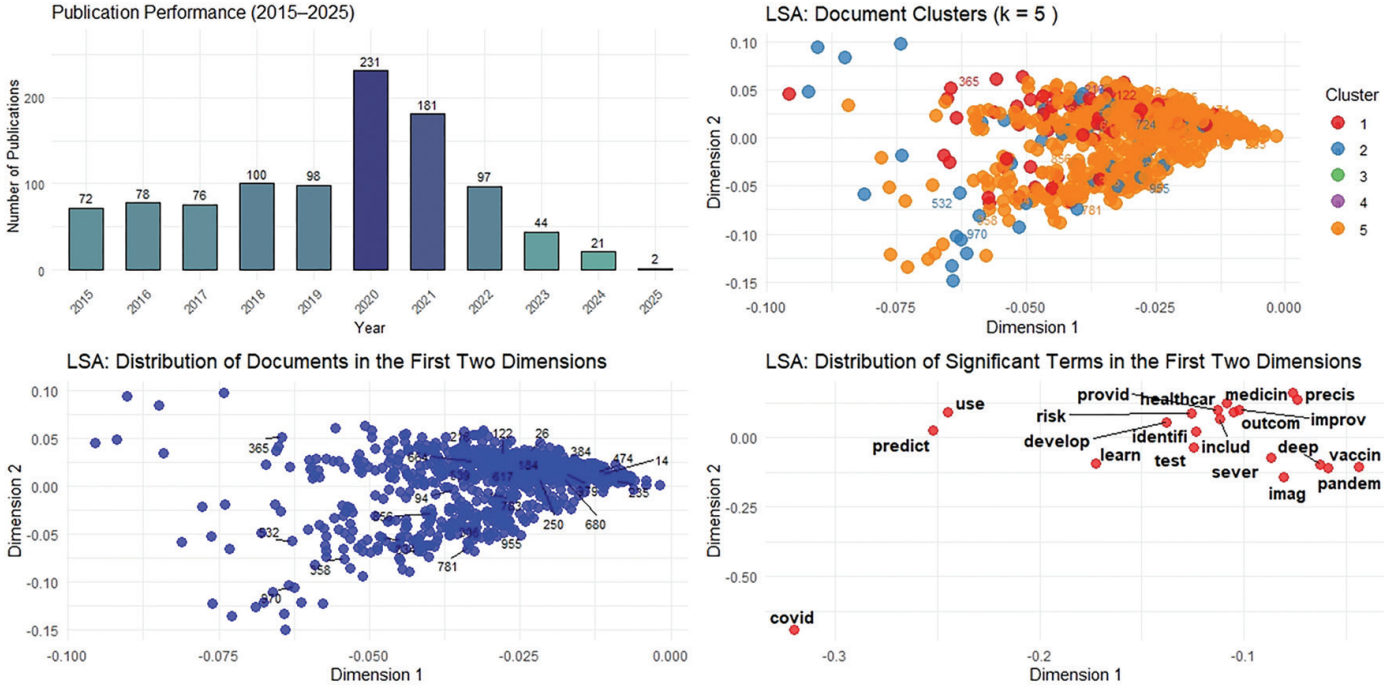


Figure 4. Descriptive and analytical findings related to the LSA analysis

LSA: Latent semantic analysis

generated by LSA. Each color in the graph represents a distinct cluster. Upon inspection, clusters 1 and 2 appear semantically similar, indicating that studies within these clusters share thematic content. In contrast, the scattered data points represent documents focusing on divergent thematic areas, demonstrating semantic diversity across the corpus. The group of points in the lower-left region of the graph clearly illustrates this dispersion.

The lower-left graph displays the distribution of documents across the first two semantic dimensions. The clustering pattern on the right side of the graph indicates that most documents share a semantic orientation. Areas with higher point density indicate documents that share strong conceptual similarity, whereas areas with points distributed toward the left or along the vertical axis likely represent unique or thematically distinct areas within the field. The lower-right graph shows the distribution of significant terms across the first two dimensions. The term "COVID" is clearly separated from the other terms, occupying a distinct region in semantic space. This finding suggests that the pandemic constitutes a thematically distinct subdomain within AI research in healthcare. Conversely, terms such as develop, deep, learn, pandemic, vaccine, healthcare, predict, outcome, identify, and risk are centrally clustered, implying a strong thematic association among studies focusing on AI-driven healthcare, deep learning, health outcomes, risk prediction, and vaccine-related research.

The isolated placement of the term "COVID" on the semantic map indicates that AI studies related to the pandemic constitute a distinct thematic domain within the LSA model. This also highlights the model's capacity to semantically distinguish pandemic-related AI research from other healthcare applications effectively. Overall, the LSA model identifies COVID-19 and pandemic-oriented AI research as a distinct thematic area. It also underscores the significance of AI-assisted clinical decision systems, particularly those focusing on risk assessment, prediction, learning, and diagnostic processes. Furthermore, the clustering of AI-related studies around the terms "healthcare" and "outcomes" clearly reveals that health analytics within the field of AI has gained an increasingly applied, data-driven orientation. In summary, the structure revealed by the LSA model demonstrates that, while COVID-19 constitutes a semantically distinct theme, most other related concepts cluster around the broader domains of health technologies and AI applications in healthcare.

Discussion

The LSA model developed for the top 1.000 most-cited studies at the intersection of AI and healthcare over the past decade demonstrates the model's ability to reveal the latent semantic structure embedded in the textual data. The findings derived from the SVD analysis show that the LSA model explains at least 80% of the total variance using 160 dimensions. Accordingly, the model achieved a high level of performance in distinguishing semantic clusters within the corpus. The LSA model results reveal that, in a two-dimensional semantic space, three distinct clusters form: two are semantically close to each other, whereas the third is positioned considerably farther away. Specifically, the words "predict" and "use" constituted the first cluster; "risk, outcome, deep, vaccine, medicine, precision, and image" formed the second cluster; and COVID represented the third and final cluster.

The concepts identified through the LSA model highlight the critical role of clinical integration between AI and healthcare domains, emphasizing their importance for this study. The positioning of the word "predict" as a partially separate cluster once again underscores the privileged role of prediction and foresight in clinical decision-making within healthcare. The distinct placement of the word "COVID", on the other hand, provides clear evidence of the model's accuracy and success, as the COVID-19 pandemic has been one of the most significant global health phenomena since December 2019[42]. Although studies focusing on the relationship between AI and COVID-19 were predominant during the pandemic, their frequency has gradually declined in subsequent years. Nevertheless, the COVID-19 theme remains a central yet semantically distinct topic, differing substantially from mainstream AI-healthcare research areas. This is clearly illustrated in the word cloud presented in Figure 5.

While COVID-19 remains a prominent topic in AI-related studies, the LSA model revealed that COVID-19 does not constitute a mainstream research domain in healthcare-related AI literature. Instead, it occupies a semantically distant and unique position in the two-dimensional semantic space. This finding demonstrates that the LSA model effectively differentiated COVID-19 from the broader semantic themes at the intersection of AI and healthcare. In this respect, the LSA model yielded unexpected yet insightful results by uncovering a hidden dimension of AI research in healthcare. The positioning of COVID-19 as a semantically distinct cluster highlights not only the value of text analytics

as a strategic tool for capturing developments in the health sector but also the potential for enhancing digital health capacity during global crises characterized by rapid growth in scientific publications. As highlighted in the literature, the pandemic witnessed a remarkable acceleration in telehealth initiatives and the implementation of remote healthcare services worldwide[43]. Thus, the findings of this study emphasize the importance of effectively leveraging scientific advancements emerging during crisis periods to strengthen the potential of AI in healthcare. For this reason, policymakers are encouraged to transform this momentum into sustainable progress by strategically investing in digital health initiatives, particularly chronic disease management, early diagnosis, and risk prediction.

The finding that the COVID-19 theme occupies a semantically distinct position in the study demonstrates that research orientations in healthcare can shift rapidly during crises. Accordingly, it can be inferred that the pandemic has acted as a significant catalyst in shaping the clinical reflections and practical adoption of AI within healthcare. This observation aligns with findings reported in the existing literature[44]. From a policymaking perspective, this underscores the necessity of flexible and adaptive research funding mechanisms that can respond swiftly to such thematic shifts during health crises such as COVID-19. Furthermore, the results of this study indicate that AI research in healthcare has become a sustained domain of inquiry. This, in turn, highlights the urgent need for a national health data strategy that establishes standards for data sharing, ethical frameworks, and interoperability across the healthcare system. The findings in the literature further support this conclusion[45]. In this context, strengthening national data infrastructures is expected to yield significant advancements in the field of AI. The study demonstrates that themes such as deep learning and prediction are of critical importance to the field of health analytics, as these concepts play a key role in the integration of AI applications into clinical practice. Within analytic frameworks that utilize patient data, AI exhibits a growing convergence with healthcare, spanning a wide spectrum-from disease diagnosis to measurement of health outcomes. These observations are strongly supported by findings reported in previous research[46]. Moreover, these developments highlight not only the need to enhance technological capacity but also the importance of strengthening human resource competencies in this field. Therefore, to maximize the benefits of AI applications in healthcare, it is recommended that AI-based health analytics training programs be developed collaboratively between universities and healthcare institutions, fostering multidisciplinary expertise.

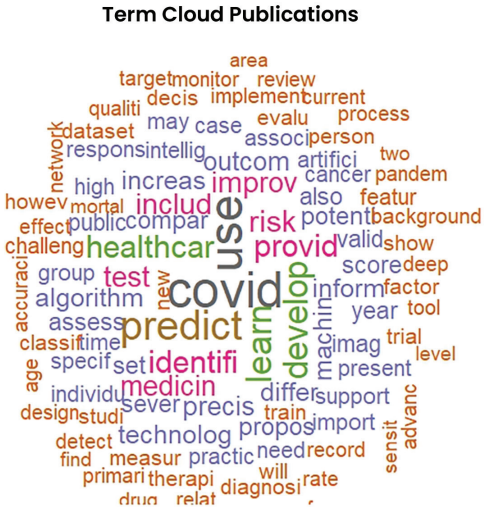


Figure 5. Word cloud of the LSA analysis

LSA: Latent semantic analysis

Study Limitations

This study used the WoS database, which may have contributed to the variability in findings regarding AI-healthcare convergence. The literature was analyzed from a healthcare management perspective using a holistic approach based on LSA, an NLP technique for text data. Future research could use alternative modeling approaches to compare results and improve generalizability. Given the study's limited dataset, collecting additional data in future research could strengthen the LSA models' representational capacity and provide more comprehensive insights into this domain. This study was limited to the top 1.000 most-cited publications from the past decade. Future research could diversify or expand the research unit, thereby enabling a broader scope of analysis and more generalizable findings. Additionally, incorporating full-text data instead of abstracts could yield results with greater representational depth and semantic richness. To achieve this, future studies may consider employing word-embedding-based models, which could enhance the precision and contextual understanding in the analysis of textual data.

Conclusion

Finally, the findings of the LSA model demonstrate that data derived from AI-related studies in healthcare can serve as a valuable resource for policy analysis. Policymakers can leverage such data-driven analytical approaches to adopt

evidence-based strategies in shaping future health policies. The LSA model used in this study revealed a significant convergence between the fields of AI and healthcare, indicating that these domains are becoming increasingly intertwined. It also showed that technological advancements in AI are increasingly reflected in clinical decision-support systems and patient-centered healthcare applications.

Drawing on Friedman's[47] fundamental theorem, it can be asserted that a healthcare system enhanced by AI will outperform one without AI. Therefore, policymakers should develop a strategic framework to safely and ethically integrate AI into healthcare-particularly within the domains of clinical decision-support and patient care. This framework must prioritize regulatory standards, professional training, data privacy, and patient safety to ensure responsible implementation. Moreover, the study demonstrates that text analytics can serve as a strategic tool for capturing and monitoring developments in healthcare. Consequently, policymakers can use such text-analytic approaches to adopt data-driven decision-making frameworks when shaping healthcare policies and strategies.

AI-Assisted Tools Disclosure

Since the study used publicly available data, ethical approval was not required. However, all other ethical principles were followed throughout the research process. The methods used and the study's main limitations have been clearly stated, and the results have been presented in an impartial and balanced manner. The article was written solely by the authors. OpenAI's ChatGPT model[48] was used to translate the manuscript into English, and Anthropic's Claude AI model[49] assisted in reviewing the translated text to ensure linguistic accuracy and professional quality.

Ethics

Ethics Committee Approval: Since the study used publicly available data, ethical approval was not required. However, all other ethical principles were followed throughout the research process.

Informed Consent: This study did not require informed consent as it involved no human participants, personal data collection, or interventional procedures.

Footnotes

Authorship Contributions

Concept: H.D., S.M., Design: H.D., S.M., Data Collection or Processing: H.D., Analysis or Interpretation: H.D., Literature Search: H.D., S.M., Writing: H.D., S.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. Abidi SSR, Abidi SR. Intelligent health data analytics: a convergence of artificial intelligence and big data. Healthc Manage Forum. 2019;32:178-82.
2. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. NPJ Digit Med. 2018;1:5.
3. Marr B. Veri stratejisi: büyük veri ve nesnelerin interneti nasıl kar getirir? İçinde: Gündüz B, (editör). İstanbul: MediCat Yayınları; 2019.
4. Tuli S, Sandhu R, Buyya R. Shared data-aware dynamic resource provisioning and task scheduling for data intensive applications on hybrid clouds using Aneka. Future Gener Comput Syst. 2020;106:595-606.
5. Bates DW, Levine DM, Syrowatka A, et al. The potential of artificial intelligence to improve patient safety: a scoping review. NPJ Digit Med. 2021;4:54.
6. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ. 2023;23:689.
7. Singhal S, Carlton S. The era of exponential improvement in healthcare? McKinsey & Company; 2019. Available from: https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/the-era-of-exponential-improvement-in-healthcare
8. Chen B, Baur A, Stepniak M, Wang J. Finding the future of care provision: the role of smart hospitals. 2019. Available from: https://www.scribd.com/document/432936514/Finding-the-future-of-care-provision-the-role-of-smart-hospital
9. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2:230-43.
10. Panch T, Mattie H, Celi LA. The "inconvenient truth" about AI in healthcare. NPJ Digit Med. 2019;2:77.
11. Kaku M. Geleceğin fiziği. 3. baskı. İçinde: Oymak YS, Oymak H, (editörler). Ankara: ODTÜ Geliştirme Vakfı Yayıncılık; 2015.
12. Liang H, Tsui BY, Ni H, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nat Med. 2019;25:433-8.
13. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med. 2018;1:39.
14. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016;316:2402-10.
15. Liu D, Cheng D, Houle TT, Chen L, Zhang W, Deng H. Machine learning methods for automatic pain assessment using facial expression information: protocol for a systematic review and meta-analysis. Medicine (Baltimore). 2018;97:e13421.
16. Li R, Zhang W, Suk H, et al. Deep learning based imaging data completion for improved brain disease diagnosis. Med Image Comput Comput Assist Interv. 2014;17:305-12.
17. Economopoulos KP, Mylonas KS, Stamou AA, et al. Laparoscopic versus robotic adrenalectomy: a comprehensive meta-analysis. Int J Surg. 2017;38:95-104.
18. Wei S, Chen M, Chen N, Liu L. Feasibility and safety of robot-assisted thoracic surgery for lung lobectomy in patients with non-small cell

lung cancer: a systematic review and meta-analysis. World J Surg Oncol. 2017;15:98.

19. Lauridsen SV, Tønnesen H, Jensen BT, Neuner B, Thind P, Thomsen T. Complications and health-related quality of life after robot-assisted versus open radical cystectomy: a systematic review and meta-analysis of four RCTs. Syst Rev. 2017;6:150.
20. Roh HF, Nam SH, Kim JM. Robot-assisted laparoscopic surgery versus conventional laparoscopic surgery in randomized controlled trials: a systematic review and meta-analysis. PLoS One. 2018;13:e0191628.
21. Schönberger D. Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. Int J Law Inf Technol. 2019;27:171-203.
22. Van Biesen W, Decruyenaere J, Sideri K, Cockbain J, Sterckx S. Remote digital monitoring of medication intake: methodological, medical, ethical and legal reflections. Acta Clin Belg. 2021;76:209-16.
23. Forcier MB, Gallois H, Mullan S, Joly Y. Integrating artificial intelligence into health care through data access: can the GDPR act as a beacon for policymakers? J Law Biosci. 2019;6:317-35.
24. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019;6:94-8.
25. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. J Med Internet Res. 2020;22:e15154.
26. Pereira KR, Sinha R. Welcome the "new kid on the block" into the family: artificial intelligence in oral and maxillofacial surgery. Br J Oral Maxillofac Surg. 2020;58:83-4.
27. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell. 2019;1:389-99.
28. Morley J, Machado CCV, Burr C, et al. The ethics of AI in health care: a mapping review. Soc Sci Med. 2020;260:113172.
29. Islam MM, Poly TN, Alsinglawi B, et al. Application of artificial intelligence in COVID-19 pandemic: bibliometric analysis. Healthcare (Basel). 2021;9:441.
30. Evangelopoulos N, Zhang X, Prybutok VR. Latent semantic analysis: five methodological recommendations. Eur J Inf Syst. 2012;21:70-86.
31. Wolfe MBW, Schreiner ME, Rehder B, et al. Learning from text: matching readers and texts by latent semantic analysis. Discourse Process. 1998;25:309-36.
32. Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. Discourse Process. 1998;25:259-84.
33. Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R. Indexing by latent semantic analysis. J Am Soc Inf Sci. 1990;41:391-407.
34. Landauer TK, Laham D, Derr M. From paragraph to graph: latent semantic analysis for information visualization. Proc Natl Acad Sci U S A. 2004;101(Suppl 1):5214-9.
35. Landauer TK. LSA as a theory of meaning. In: Landauer TK, McNamara DS, Dennis S, Kintsch W, (editors). Handbook of latent semantic analysis. Mahwah, NJ: Lawrence Erlbaum Associates; 2007:3-32.
36. Landauer TK, Dumais ST. A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychol Rev. 1997;104:211-40.
37. Valdez D, Pickett AC, Goodson P. Topic modeling: latent semantic analysis for the social sciences. Soc Sci Q. 2018;99:1665-79.
38. Foltz PW. Latent semantic analysis for text-based research. Behav Res Methods Instrum Comput. 1996;28:197-202.
39. Shen CW, Ho JT. Technology-enhanced learning in higher education: a bibliometric analysis with latent semantic approach. Comput Human Behav. 2020;104:106177.
40. R Core Team. R: A Language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2025. Available from: https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf
41. Günther F, Dudschig C, Kaup B. LSAfun--An R package for computations based on latent semantic analysis. Behav Res Methods. 2015;47:930-44.
42. Liu YC, Kuo RL, Shih SR. COVID-19: the first documented coronavirus pandemic in history. Biomed J. 2020;43:328-33.
43. Smolić Š, Blaževski N, Fabijančić M. Remote healthcare during the COVID-19 pandemic: findings for older adults in 27 European countries and Israel. Front Public Health. 2022;10:921379.
44. Adadi A, Lahmer M, Nasiri S. Artificial Intelligence and COVID-19: a systematic umbrella review and roads ahead. J King Saud Univ Comput Inf Sci. 2022;34:5898-920.
45. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J. 2021;8:e188-94.
46. Suman AP, Srivastava NK, Sharma P, Tyagi P. The convergence of artificial intelligence and healthcare in revolutionizing diagnosis, treatment, and personalization: a review. In: Proc 3rd Int Conf Disruptive Technol (ICDT). Greater Noida, India; 2025:1026-32.
47. Friedman CP. A "fundamental theorem" of biomedical informatics. J Am Med Inform Assoc. 2009;16:169-70.
48. OpenAI. ChatGPT (Version 4) [AI model]. OpenAI; 2025. Available from: https://www.openai.com/chatgpt
49. Anthropic. Claude AI (Version 1.0) [AI language model]. Anthropic; 2025. Available from: https://www.anthropic.com

2025 Referee Index - 2025 Hakem Dizini

Abdurrahman Hamdi İnan
Abdülaziz Doğan
Adnan Budak
Ahmet Acarer
Ahmet Deniz Uçar
Ahmet Karayiğit
Ahmet Kaya
Ahmet Özveren
Alaattin Yıldız
Ali Haydar Baykan
Alper İleri
Aslı Bayındır
Aslı Sade Memişoğlu
Atila Güngör
Atila Yıldırım
Atilla Ersen
Atilla Hikmet Çilengir
Ayça İnci
Ayfer Çolak
Aylin Demirci
Bahadır Öndeş
Beray Gelmez Taş
Berk Göktepe
Burcu Çakar
Cemile Canan Seçer
Cüneyt Arıkan
Çağla Ayer
Deniz Noyan Özlü
Dolunay Gürses
Durmuş Ali Çetin
Duygu Ayaz
Ejder Saylav Bora
Elanur Karaman
Elif Sarı
Elif Sıla Erim
Elif Türkan Aslan
Elif Uygur Küçükseymen
Emin Taşkıran
Erdal Yılmaz
Erhan Şimşek
Erman Alçı
Ersin Aksay
Eyüp Kebabçı
Eyüp Murat Yılmaz
Fatma Özge Kayhan Koçak
Figen Tokuçoğlu
Fikret Poyraz Çökmüş
Füsun Demirçivi Özer
Giray Kolcu
Gizem Gürsoy
Gökalp Okut
Göksever Akpınar
Gözde Derviş Hakim
Gülbin Aydoğdu Umaç
Günce Başarır
Güneş Özhan
Haluk Mergen
Haşim Çapar
Hülya Çakmur
Hülya Çetin Tunçez
Hülya Tosun Yıldırım
Hüseyin Aslan
İbrahim Akkoç
İlker Burak Arslan
İpek Polat
İrem Fatma Uludağ
İsmail Biçer
Kadir Aşçıbaşı
Kemal Erdinç Kamer
Mazlum Kılıç
Mehmet Özeren
Mehmet Şenoğlu
Mehmet Toptaş
Merve Tekinarslan
Murat Akşit
Murat Özdemir
Mustafa Tözün
Muzaffer Sancı
Neşe Çolak
Orkun Eroğlu
Ozan Barış Namdaroğlu
Özge Vural
Pakize Karaoğlu
Ragıp Kayar
Ragıp Ortaç
Selma Tekin
Seval Akay
Sibel Yavuz
Sinan Arı
Şermin Can
Şükrü Aydın
Talat Cem Özdemir
Timur Meşe
Tolga Acun
Tuğba Koca
Tuncay Küme
Turgay Yılmaz Kılıç
Umut Beylik
Übeyd Sungur
Ünsal Yılmaz
Yasemin Kılıç Öztürk
Yasin Aras
Zehra Hilal Adıbelli
Zeynep Altın