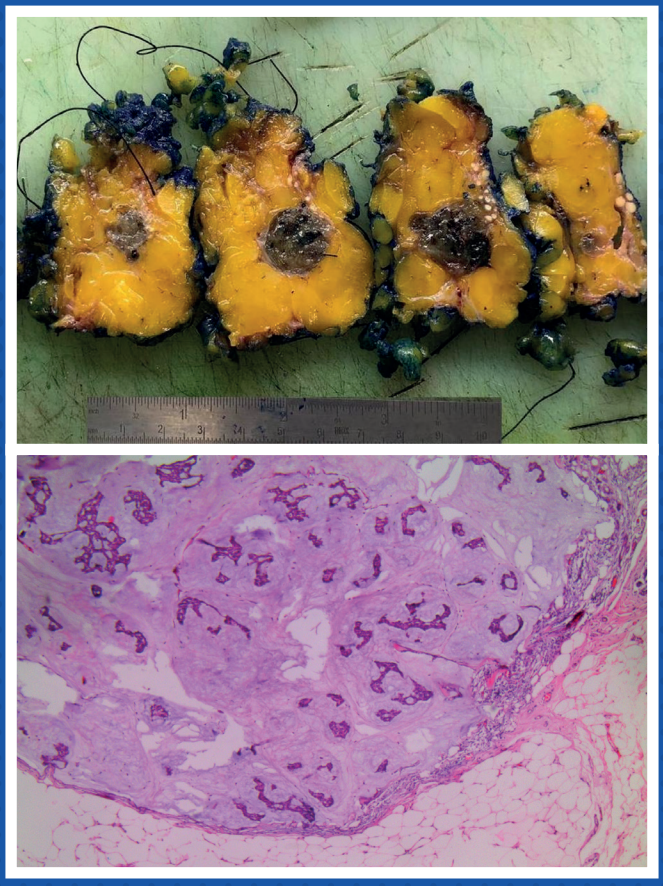
ISSN: 2979-9538

NORTHWESTERN MEDICAL JOURNAL

Year: 2026 ▪ Volume: 6 ▪ Issue: 3



Official Journal of Izzet Baysal Training and Research Hospital

NORTHWESTERN MEDICAL JOURNAL

Year: 2026 ▪ Volume: 6 ▪ Issue: 3

ISSN: 2979-9538

www.nwmedj.org

Official Journal of Izzet Baysal
Training and Research Hospital

Northwestern Medical Journal

ISSN: 2979-9538 (Online)

Owner
Owner on behalf of the Bolu İzzet Baysal Training and Research Hospital
Assoc. Prof. Süleyman Özsarı, MD
Director, Bolu İzzet Baysal Training and Research Hospital, Türkiye

Publication Type
International peer-reviewed journal

Publication Frequency and Language
Quarterly (January, April, July, October), English

Abstracting and Indexing
Northwestern Medical Journal indexed in TR Index, DOAJ, GALE, J-Gate, and TürkMedline.

Editor in Chief
Prof. Ahmet Ural, MD
Bolu Abant Izzet Baysal University, Faculty of Medicine, Department of Otorhinolaryngology, Head & Neck Surgery, Türkiye

Publisher
Izzet Baysal Training and Research Hospital

Publisher Address
Izzet Baysal Training and Research Hospital, Gölköy 14280 Bolu, Türkiye
Phone: +90 374 253 46 18 / +90 374 253 46 56
E-mail: nwmedj@gmail.com
Web: izzetbaysaleah.saglik.gov.tr

Publishing Services
Akdema Informatics and Publishing
Address: Kızılay Mah. Gazi Mustafa Kemal Bulvarı No: 23/8 06420 Çankaya/Ankara
Certificate number: 52576
E-mail: bilgi@akdema.com
Tel: +90 533 166 80 80
Web: www.akdema.com

You can reach all publication policies and author guidelines from **www.nwmedj.org**

Northwestern Medical Journal

Editorial Board

Northwestern Medical Journal

Advisory Board

Contents

Editorial

As we enter the second half of 2026, the medical landscape continues to evolve at a speed unprecedented in clinical history. The integration of advanced artificial intelligence, multi-omics biomarker profiling, and continuous digital health monitoring has fundamentally reshaped how we define, diagnose, and manage disease. Yet, as the boundaries of what is clinically possible expand daily, a fundamental question confronts our healthcare system: *How do we ensure that technological precision enhances, rather than eclipses, human-centered care?*

In July 2026 issue of the *Northwestern Medical Journal*, Karanfil et al. evaluated the range of motion after fasciectomy in dupuytren's contracture and the impact of rehabilitation. Karabekiroğlu et al. presented a case of mucinous carcinoma of the breast diagnosed incidentally during screening mammography. Özarslantürk et al. assessed the artifical inteligence chatbots for informing patients about cone beam computed tomography. Zorsu et al. investigated the effects of vitamin D replacement on disease symptoms, sleep, depression, and fatigue in restless leg syndrome patients. Ay et al. reported fertility and obstetric outcomes after complete and partial hydatidiform mole. Akbaş Güneş et al. studied th relationship between anthropometric measurements and frequency and severity of attack in migraine patients in Western Türkiye. Gökbulut et al. analyzed the cognitive performance in chronic postoperative hypoparathyroidism patients. Özer et al. reported ryan tumor regression scoring in rectal cancers operated after neoadjuvant therapy.

While the data-rich environment of modern clinical practice offers remarkable accuracy, it also risks increasing cognitive burden and screen-induced distance between clinicians and patients. Technology must function not as a substitute for clinical judgment or empathetic listening, but as an invisible framework that frees clinicians to focus on what matters most: understanding the human being behind the metrics.

As researchers, educators, and practitioners, our collective responsibility is to ensure that every technological leap forward is matched by a equal commitment to bioethical integrity and global health equity. The studies featured in this issue exemplify this essential balance, proving that cutting-edge science and compassionate healthcare are not mutually exclusive—they are fundamentally interdependent.

We invite our readers to engage deeply with the research and commentaries presented this month, as we work together to shape a future where medical innovation serves every patient, everywhere.

Sincerely,
Prof. Ahmet Ural, M.D.
Editor-in-chief

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):181-191
DOI: 10.54307/NWMJ.2026.206

RESEARCH ARTICLE

Assessment of a body shape index in patients with normal gated myocardial perfusion SPECT findings

Seda Sertel Meyvaci[1], Billur Caliskan[2], Hamdi Afsin[2], Handan Ankarali[3], Beyza Celik[1]

[1]Department of Anatomy, Faculty of Medicine, Bolu Abant İzzet Baysal University, Bolu, Türkiye
[2]Department of Nuclear Medicine, Faculty of Medicine, Bolu Abant İzzet Baysal University, Bolu, Türkiye
[3]Department of Biostatistics and Medical Informatics, Faculty of Medicine, İstanbul Medeniyet University, İstanbul, Türkiye

Cite as: Sertel Meyvaci S, Caliskan B, Afsin H, Ankarali H, Celik B. Assessment of a body shape index in patients with normal gated myocardial perfusion SPECT findings. Northwestern Med J. 2026;6(3):181-191.

ABSTRACT

Objective: This study aimed to evaluate the relationship between A Body Shape Index (ABSI) and normal values of gated myocardial perfusion SPECT (g-MPS) parameters.

Materials and Methods: This retrospective cross-sectional study analyzed 329 individuals (188 females and 141 males) who presented to the outpatient clinic of the Department of Nuclear Medicine, Faculty of Medicine, Bolu Abant Izzet Baysal University, and underwent myocardial perfusion scintigraphy (MPS). Anthropometric measurements were obtained before MPS, and g-MPS parameters were retrieved from clinical reports of patients with normal perfusion and gated functional findings. Descriptive statistics, including mean, standard deviation, median, interquartile range, frequency, and percentage, were calculated. Relationships between variables were evaluated using Spearman's rank correlation analysis.

Results: The analysis revealed the following mean values: age, 60.61 years; height, 161.15 cm; weight, 79.41 kg; waist circumference (WC), 95.73 cm; body mass index (BMI), 30.74 kg/m^2; and ABSI, 0.081. A positive correlation was found between BMI and WC ($r = 0.730$, $p < 0.001$). ABSI showed a weak negative correlation with heart rate ($r = -0.124$, $p = 0.025$). Metabolic equivalent of task (METS) values were negatively correlated with BMI ($r = -0.301$, $p < 0.001$) and WC ($r = -0.205$, $p < 0.001$). Additionally, negative correlations were observed between risk level and heart rate ($r = -0.121$, $p = 0.029$), and between ejection fraction (EF) and both end-diastolic volume (EDV; $r = -0.628$, $p < 0.001$) and end-systolic volume (ESV; $r = -0.884$, $p < 0.001$).

Conclusion: In individuals with normal myocardial perfusion findings, ABSI may serve as a complementary anthropometric index, providing additional insight into the relationship between abdominal fat distribution and cardiovascular physiological parameters.

Keywords: body mass index, body shape index, gated myocardial perfusion SPECT, left ventricular function, obesity

Corresponding author: Beyza Celik **E-mail:** beyzayilmazcelik1@gmail.com
Received: 12.05.2025 **Accepted:** 30.03.2026 **Published:** 31.07.2026

INTRODUCTION

Gated single-photon emission computed tomography (SPECT) is a myocardial perfusion scintigraphy technique that allows quantitative assessment of left ventricular functional and volumetric parameters in addition to myocardial perfusion (1,2). The precision and reproducibility of functional parameters derived from this technique have increasingly highlighted the clinical importance of gated SPECT data (3). Scintigraphy allows for the evaluation of the presence, severity, and distribution of ischemic myocardial segments, along with wall motion abnormalities and ejection fraction (EF), both visually and quantitatively (4,5). Obesity serves as a robust independent risk factor for cardiovascular diseases, with visceral adiposity emerging as a critical determinant of cardiovascular risk (6). Abdominal obesity, defined as abnormal fat accumulation posing health risks, can be assessed via anthropometric measurements (7). These measurements, which are simple, cost-effective, and non-invasive, are widely utilized to evaluate morbidity and mortality risks (8). Commonly used indices include body mass index (BMI), waist circumference (WC), waist-to-hip ratio, waist-to-height ratio, visceral fat area, body fat percentage, and the novel A Body Shape Index (ABSI) (9).

ABSI, a novel anthropometric measure designed to more accurately reflect abdominal adiposity and visceral fat distribution, has attracted increasing attention in recent studies as an alternative to BMI, as it demonstrates a stronger association with cardiometabolic risk than both BMI and WC (10-12).

Studies indicate that abdominal obesity, defined as WC ≥90 cm in men and ≥80 cm in women, doubles the risk of diabetes, hypertension, and atherosclerotic cardiovascular disease (10). Furthermore, ABSI has been proposed as a reliable anthropometric tool for population health surveillance due to its predictive efficacy for chronic diseases and mortality (13). Moreover, in various diseases in which visceral fat distribution is considered a risk factor, ABSI is widely used in clinical practice and has attracted attention due to its high predictive performance (11,14,15). This study aims to investigate the relationship between ABSI and normal parameters of gated myocardial perfusion SPECT (g-MPS).

MATERIALS AND METHODS

In this cross-sectional study, patients who underwent myocardial perfusion scintigraphy (MPS) at the Department of Nuclear Medicine, Faculty of Medicine, Bolu Abant Izzet Baysal University, between May 2023 and February 2024 were evaluated. Anthropometric measurements were obtained before the MPS procedure. Following the procedure, g-MPS images and clinical reports were reviewed, and patients with normal myocardial perfusion and normal gated functional parameters were included in the study.

All included patients had completely normal myocardial perfusion and gated functional findings, with no borderline or equivocal imaging results present in the study population. Patients meeting the inclusion criteria were enrolled in the study on the same day. Patients with paced rhythm, left bundle branch block, right bundle branch block, cardiomyopathy, structural cardiac abnormalities, or myocardial perfusion abnormalities were excluded.

This study received ethical approval from the Bolu Abant Izzet Baysal University Clinical Research Ethics Committee (Decision No: 2023/100).

The study analyzed the following parameters reported in the g-MPS reports: Ejection Fraction (EF), End-Systolic Volume (ESV), End-Diastolic Volume (EDV), Metabolic Equivalent of Task (METS), time to reach target heart rate (duration), and the percentage of achieved target heart rate (heart rate). Age, sex, height, and weight values were recorded from the examination reports. Additionally, WC was directly measured with a non-elastic tape measure for ABSI calculation. Based on the anthropometric measurements of all participants (including height, weight, and WC), BMI and ABSI were calculated. WC (cm) measurements were performed with the participants standing upright,

feet together, and at the end of a normal expiration. The measurement was taken at the approximate midpoint between the lower margin of the last palpable rib and the top of the iliac crest using a measuring tape positioned horizontally and without compressing the skin. This procedure was conducted in accordance with the World Health Organization (WHO) anthropometric measurement protocol (16).

BMI and ABSI were calculated using the following formulas:

$BMI = Weight / Height^2\ (kg/m^2)$.

$ABSI = WC \div (BMI^{2/3} \times Height^{1/2})$ (13,17).

Only cases with g-MPS values within normal physiological limits among those who underwent MPS were included in the analysis. The clinical and demographic parameters analyzed were: age, height, weight, WC, BMI, ABSI, heart rate, duration, METS, EF, ESV, and EDV. This study was conducted on 329 patients (188 female, 141 male) referred to the Nuclear Medicine Clinic for MPS due to suspected Coronary Artery Disease (CAD). Patients underwent a physiological exercise stress test using a treadmill device following the Bruce or Modified Bruce protocol. Upon reaching 85–100% of their age-predicted maximum heart rate, 10 mCi of 99mTc-MIBI was injected, and electrocardiography (ECG)-g-MPS was performed using a two-day protocol. After imaging, the parameters of EF, EDV, and ESV were obtained using QPS/QGS software (Cedars-Sinai Medical Center). Myocardial perfusion was evaluated using a 17-segment model and a five-point scoring system (0 = normal perfusion, 1 = equivocal hypoperfusion, 2 = moderate hypoperfusion, 3 = severe hypoperfusion, and 4 = absent perfusion). Summed stress score (SSS), summed rest score (SRS), and summed difference score (SDS) were calculated. An SDS <2 and an SSS ≤3 were considered indicative of normal myocardial perfusion. Gated-MPS images were interpreted independently by two observers, and in cases of discrepant interpretations, images were re-evaluated jointly and a final decision was reached by consensus.

Percentage of Achieved Target Heart Rate (Heart Rate): This parameter indicates the percentage of the patient's achieved heart rate relative to the age-predicted maximum value calculated as (220 - age). The target heart rate is calculated as (220 - age) × 0.85. If the patient cannot reach 85% due to capacity limitations or symptoms (e.g., chest pain, dyspnea), they are expected to reach at least 70–85%.

Duration: This is the time (in minutes) required to reach the target heart rate.

Metabolic Equivalent of Task (METS): This is a unit measuring myocardial oxygen consumption capacity. At rest, the heart muscle consumes approximately 3.5 mL/kg/min (1 MET) of oxygen. Under stress, the oxygen demand increases, and computer-assisted systems calculate how much higher the oxygen consumption is compared to the resting state. METS values were automatically calculated based on the treadmill exercise protocol.

Ejection Fraction: EF is a critical measure of the heart's stroke volume. It is calculated as the ratio of the difference between the EDV and ESV to the EDV: *EF = (EDV - ESV) / EDV.* An EF value of ≥50% was defined as normal. EF was calculated using EDV and ESV values derived from gated-MPS data. This ratio represents the percentage of the blood volume ejected from the left ventricle during systole. A high EF indicates effective cardiac pumping, while a low EF may suggest heart failure or other cardiac dysfunction.

End-Systolic Volume: ESV is the volume of blood remaining in the left ventricle at the end of systole, after the heart has contracted. It provides information about the efficiency of the heart's pumping function; a low ESV indicates more effective blood ejection, while a high ESV may signal potential heart failure.

End-Diastolic Volume: Represents the maximum volume of blood in the left ventricle at the end of diastole when the heart muscle is relaxed and filled with blood. This parameter is crucial in determining the heart's filling capacity and overall pumping efficiency. A high EDV suggests greater blood intake, which can enhance the strength of cardiac contraction.

Statistical analysis

Descriptive statistics of the measurements were calculated as the mean, standard deviation (SD), median, 25th and 75th percentiles (IQR), number, and percentage frequencies depending on the nature of the data. The conformity of numerical data to normal distribution was evaluated using the Shapiro-Wilk test, and it was determined that the data were not normally distributed. Relationships between measurements were analyzed using Spearman's rank correlation analysis. A p-value, $p \leq 0.05$ was considered statistically significant, and all calculations were performed using SPSS software (version 23).

RESULTS

The mean values of the 329 participants (188 female, 141 male) were as follows: age, 60.61 ± 11.35 years; height, 161.15 ± 8.95 cm; weight, 79.41 ± 13.88 kg; WC, 95.73 ± 12.35 cm; BMI, 30.74 ± 5.81 kg/m², and ABSI, 0.081 ± 0.056. Descriptive statistics of the participants' numerical characteristics are presented in Table 1.

The descriptive values for the categorical characteristics of the participants are presented in Table 2.

The correlations between the measurements are presented in Table 3. A statistically significant but weak negative correlation was observed between ABSI and heart rate (r = -0.124). In addition, a statistically significant moderate positive correlation was found between ABSI and risk level (r = 0.457). In contrast, no significant associations were detected between ABSI and g-MPS-derived volumetric or functional parameters.

A weak negative correlation was identified between risk level and heart rate (r = -0.121), whereas a weak positive correlation was observed between risk level and duration (r = 0.120). No significant relationship was detected between risk level and other g-MPS parameters.

When the relationships between g-MPS findings and BMI were examined, weak positive correlations were found with EF (r = 0.124) and heart rate (r = 0.155), while moderate negative correlations were observed with METS (r = -0.301) and duration (r = -0.284). In

Table 1. Descriptive statistics of participants' numerical characteristics

	N	Mean	SD	Min	Max	Percentiles		
						25	Median	75
Age	329	60.61	11.35	25.00	85.00	52.50	62.00	69.00
EF	328	67.33	11.82	21.00	98.00	60.00	68.00	75.00
EDV	328	69.24	24.83	18.00	192.00	51.00	66.00	82.00
ESV	328	24.87	17.09	1.00	131.00	13.25	21.00	32.00
Heart Rate	329	84.25	5.53	68.00	110.00	81.00	83.00	86.00
METS	329	7.75	2.50	1.18	14.79	5.17	6.87	9.54
Duration	329	4.72	2.25	1.00	11.45	3.01	4.44	6.36
Height	329	161.15	8.95	130.00	183.00	155.00	161.00	168.00
Weight	329	79.41	13.88	38.00	118.00	70.50	79.00	87.00
BMI	329	30.74	5.81	15.82	51.37	26.77	29.91	33.92
WC	329	95.73	12.35	60.00	130.00	88.50	96.00	104.00
ABSI	329	0.081	0.056	0.044	0.830	0.073	0.078	0.081

ABSI: a body shape index, BMI: body mass index, EDV: end-diastolic volume, EF: ejection fraction, ESV: end-systolic volume, METS: metabolic equivalents of task, WC: waist circumference, SD: standard deviation, Min: minimum, Max: maximum.

Table 2. Descriptive values for the categorical characteristics of the participants

Categorical characteristics		N	%
Gender	Female	188	57.1
	Male	141	42.9
Cholesterol	No	226	68.7
	Yes	103	31.3
Smoking	No	257	78.1
	Yes	72	21.9
Family History	No	154	46.8
	Yes	175	53.2
Hypertension	No	132	40.1
	Yes	197	59.9
Risk Level	Very low	193	58.7
	Low	65	19.8
	Average	39	11.9
	High	9	2.7
	Very high	23	7.0

the analysis of g-MPS findings in relation to WC, weak to moderate positive correlations were identified with EDV (r = 0.183) and ESV (r = 0.109), whereas weak negative correlations were found with METS (r = -0.205) and duration (r = -0.144). Other correlations not described in detail are presented in Table 3.

The strong positive correlation between BMI and WC (r = 0.730) indicates a close association between these two measures in relation to abdominal adiposity. A statistically significant but weak negative correlation was observed between ABSI and heart rate, indicating an association between these parameters. Moderate and weak negative correlations were observed between METS and BMI (r = -0.301) and WC (r = -0.205), respectively. A weak negative correlation was also present between risk level and heart rate (r = -0.121).

DISCUSSION

In this study, the relationship between g-MPS parameters and ABSI, an indicator of abdominal obesity, was investigated in individuals with normal myocardial perfusion imaging findings. The results provide insight into the associations between anthropometric indices and cardiac physiological parameters in a cardiovascularly normal population. Previous studies have reported that increased ABSI values are associated with coronary artery disease and cardiovascular risk. However, as all participants included in the present study had completely normal g-MPS findings, our results do not indicate the presence of coronary artery disease or an increased CAD risk. Instead, the observed associations reflect physiological or subclinical relationships rather than evidence of overt disease. In this context, our findings are in line with the existing literature addressing obesity-related cardiovascular risk factors, while emphasizing that such associations may also be observed in the absence of manifest cardiovascular disease (18-21)

Although BMI is widely used in the assessment of obesity, it has important limitations in accurately reflecting body composition, as it does not distinguish between fat and muscle components (22-24).

Obesity is not only related to body weight but also to fat distribution. Krakauer et al. reported that, because the waist circumference component in the ABSI formula reflects abdominal fat accumulation, ABSI is a more effective indicator for predicting mortality risk independently of BMI (13). Furthermore, studies have indicated that the use of alternative anthropometric measures, such as ABSI, as an alternative to BMI, is one of the factors that enhances the reliability of cardiovascular disease risk assessment (25,26).

In a study evaluating the performance of the A Body Shape Index (ABSI) in predicting hypertension, cardiovascular disease, type 2 diabetes, and all-cause mortality, ABSI demonstrated superior performance in predicting all-cause mortality compared with BMI and WC (10).

Table 3. Correlations between numerical measurements

		r	P	N	95% Confidence Interval [Lower; Upper]
Age	BMI	0.063	0.257	329	[-0.045 ; 0.170]
	WC	**0.210**	**<0.001**	**329**	[0.104 ; 0.311]
	ABSI	**0.332**	**<0.001**	**329**	[0.233 ; 0.424]
	Risk Level	0.015	0.792	329	[-0.093 ; 0.123]
	EF	0.063	0.257	328	[-0.045 ; 0.170]
	EDV	**-0.124**	**0.025**	**328**	[-0.229 ; -0.016]
	ESV	-0.102	0.065	328	[-0.208 ; 0.006]
	Heart Rate	**0.232**	**<0.001**	**329**	[0.127 ; 0.332]
	METS	**-0.478**	**<0.001**	**329**	[-0.557 ; -0.390]
	Duration	**-0.441**	**<0.001**	**329**	[-0.524 ; -0.350]
Height	WC	-0.053	0.335	329	[-0.160 ; 0.055]
	ABSI	**0.227**	**<0.001**	**329**	[0.122 ; 0.327]
	Risk Level	0.046	0.410	329	[-0.062 ; 0.153]
	EF	**-0.330**	**<0.001**	**328**	[-0.422 ; -0.231]
	EDV	**0.427**	**<0.001**	**328**	[0.334 ; 0.511]
	ESV	**0.411**	**<0.001**	**328**	[0.317 ; 0.497]
	Heart Rate	**-0.179**	**0.001**	**329**	[-0.282 ; -0.072]
	METS	**0.258**	**<0.001**	**329**	[0.154 ; 0.356]
	Duration	**0.312**	**<0.001**	**329**	[0.211 ; 0.406]
Weight	WC	0.727	**<0.001**	329	[0.672 ; 0.774]
	ABSI	-0.098	0.076	329	[-0.204 ; 0.010]
	Risk Level	-0.086	0.120	329	[-0.193 ; 0.022]
	EF	-0.056	0.311	328	[-0.163 ; 0.052]
	EDV	**0.297**	**<0.001**	**328**	[0.195 ; 0.392]
	ESV	**0.193**	**<0.001**	**328**	[0.087 ; 0.294]
	Heart Rate	0.044	0.423	329	[-0.064 ; 0.151]
	METS	**-0.164**	**0.003**	**329**	[-0.268 ; -0.056]
	Duration	**-0.110**	**0.050**	**329**	[-0.215 ; -0.002]
BMI	WC	0.730	<0.001	329	[0.675 ; 0.777]
	ABSI	**-0.208**	**<0.001**	**329**	[-0.309 ; -0.102]
	Risk Level	-0.090	0.102	329	[-0.196 ; 0.018]
	EF	**0.124**	**0.025**	**328**	[0.016 ; 0.229]

ABSI: a body shape index, BMI: body mass index, CI: confidence interval, Duration: time to reach the target heart rate, EDV: end-diastolic volume, EF: ejection fraction, ESV: end-systolic volume, Heart Rate: percentage of achieved target heart rate, METS: metabolic equivalents of task, N: number of participants, P: probability value, r: correlation coefficient, WC: waist circumference.

Table 3. Continued

		r	P	N	95% Confidence Interval [Lower; Upper]
	EDV	0.026	0.645	328	[-0.082 ; 0.134]
	ESV	-0.049	0.374	328	[-0.156 ; 0.059]
	Heart Rate	**0.155**	**0.005**	**329**	[0.048 ; 0.259]
	METS	**-0.301**	**<0.001**	**329**	[-0.404 ; -0.209]
	Duration	**-0.284**	**<0.001**	**329**	[-0.375 ; -0.176]
WC	ABSI	**0.410**	**<0.001**	**329**	[-0.677 ; -0.539]
	Risk Level	**0.342**	**<0.001**	**329**	[-0.274 ; -0.063]
	EF	-0.001	0.988	328	[0.072 ; 0.282]
	EDV	**0.183**	**0.001**	**328**	[-0.154 ; 0.063]
	ESV	**0.109**	**0.048**	**328**	[-0.217 ; -0.003]
	Heart Rate	0.061	0.267	329	[0.070 ; 0.280]
	METS	**-0.205**	**<0.001**	**329**	[-0.393 ; -0.194]
	Duration	**-0.144**	**0.009**	**329**	[-0.351 ; -0.147]
ABSI	Risk Level	**0.457**	**<0.001**	**329**	[0.368 ; 0.538]
	EF	-0.083	0.135	328	[-0.264 ; -0.053]
	EDV	0.079	0.154	328	[-0.101 ; 0.115]
	ESV	0.103	0.062	328	[-0.040 ; 0.175]
	Heart Rate	**-0.124**	**0.025**	**329**	[-0.099 ; 0.117]
	METS	0.017	0.766	329	[-0.220 ; -0.006]
	Duration	0.082	0.137	329	[-0.224 ; -0.010]
Risk Level	EF	0.016	0.778	328	[-0.087 ; 0.129]
	EDV	-0.001	0.991	328	[-0.048 ; 0.167]
	ESV	-0.005	0.923	328	[-0.060 ; 0.155]
	Heart Rate	**-0.121**	**0.029**	**329**	[-0.107 ; 0.109]
	METS	0.091	0.099	329	[-0.239 ; -0.026]
	Duration	**0.120**	**0.029**	**329**	[-0.243 ; -0.030]
EF	EDV	**-0.628**	**<0.001**	**328**	[-0.458 ; -0.274]
	ESV	**-0.884**	**<0.001**	**328**	[-0.677 ; -0.539]
	Heart Rate	0.047	0.399	328	[-0.052 ; 0.164]
	METS	**-0.178**	**0.001**	**328**	[-0.229 ; -0.016]
	Duration	**-0.243**	**<0.001**	**328**	[-0.227 ; -0.014]
EDV	ESV	**0.866**	**<0.001**	**328**	[0.952 ; 0.969]

ABSI: a body shape index, BMI: body mass index, CI: confidence interval, Duration: time to reach the target heart rate, EDV: end-diastolic volume, EF: ejection fraction, ESV: end-systolic volume, Heart Rate: percentage of achieved target heart rate, METS: metabolic equivalents of task, N: number of participants, P: probability value, r: correlation coefficient, WC: waist circumference.

Table 3. Continued

		r	P	N	95% Confidence Interval [Lower; Upper]
	Heart Rate	**-0.165**	**0.003**	**328**	[-0.233 ; -0.020]
	METS	**0.261**	**<0.001**	**328**	[0.100 ; 0.309]
	Duration	**0.312**	**<0.001**	**328**	[0.117 ; 0.324]
ESV	Heart Rate	**-0.141**	**0.011**	**328**	[-0.248 ; -0.036]
	METS	**0.250**	**<0.001**	**328**	[0.104 ; 0.313]
	Duration	**0.309**	**<0.001**	**328**	[0.124 ; 0.331]
Heart Rate	METS	**-0.444**	**<0.001**	**329**	[-0.265 ; -0.054]
	Duration	**-0.446**	**<0.001**	**329**	[-0.250 ; -0.039]
METS	Duration	**0.886**	**<0.001**	**329**	[0.955 ; 0.971]

ABSI: a body shape index, BMI: body mass index, CI: confidence interval, Duration: time to reach the target heart rate, EDV: end-diastolic volume, EF: ejection fraction, ESV: end-systolic volume, Heart Rate: percentage of achieved target heart rate, METS: metabolic equivalents of task, N: number of participants, P: probability value, r: correlation coefficient, WC: waist circumference.

Hung et al. investigated the role of abdominal obesity and circulating adipokines in predicting myocardial ischemia and found results supporting the impact of abdominal obesity on cardiovascular health. These findings highlight associations between abdominal fat accumulation and cardiovascular health reported in previous studies (4). Abdominal obesity has been associated with various health conditions, including heart disease, hypertension, and diabetes, in previous studies (27). In this study, the negative association observed between ABSI and heart rate in cardiovascularly healthy individuals suggests that abdominal fat distribution may be associated with cardiovascular physiological responses. Our findings are consistent with previous studies investigating the relationship between abdominal obesity and cardiovascular parameters (28). Given the associations reported between obesity-related measurements and cardiovascular parameters, the relationship between heart rate, body composition, and metabolism is a critical factor that should be considered in the treatment of obesity (29). According to the g-MPS findings, the positive correlations observed between EDV and ESV and waist circumference suggest that abdominal adiposity may be associated with left ventricular volume parameters. This relationship indicates that increased abdominal fat distribution may be related to early cardiac physiological changes in cardiovascularly healthy individuals and is consistent with previous studies reporting associations between waist circumference, visceral adiposity, cardiovascular risk factors, and cardiac parameters (29-32). Additionally, the significant negative correlation observed between waist circumference and heart rate suggests that abdominal fat distribution may be associated with cardiac physiological regulation. Chartrand et al. investigated the effects of visceral fat on cardiac health and reported associations between abdominal obesity and cardiovascular risk profiles in previous studies. Taken together, these findings indicate that waist circumference is an important anthropometric measure for assessing abdominal obesity and its association with cardiovascular parameters (6). There are significant differences between men and women regarding the effects of obesity on the body. The global prevalence of obesity in women is higher than in men (18). From a women's health perspective, obesity can lead to hormonal imbalances, insulin resistance, and an increased risk of metabolic syndrome (33). Exceeding the accepted waist circumference threshold values has been associated with an unfavorable cardiovascular risk profile, and waist circumference has been identified as an important anthropometric indicator for the assessment of abdominal adiposity (34). Women generally have a higher proportion of adipose tissue than men, with fat more commonly distributed in the lower extremities, a pattern that has been suggested to be relatively protective in terms of cardiometabolic

risk. Additionally, estrogen has been suggested to play a protective role against cardiovascular diseases in women (35).

But hormonal changes occurring after menopause may promote a shift toward abdominal fat accumulation in women, potentially increasing cardiometabolic risk. These observations highlight the importance of considering sex-specific fat distribution patterns, rather than BMI alone, when evaluating cardiometabolic risk (36-38).

In men, fat tissue typically accumulates in the abdominal area, which is a significant factor that increases the risk of heart disease. Therefore, abdominal obesity has been associated with a higher cardiovascular disease risk in men in previous studies (39).

Overall, this study demonstrated that anthropometric measures such as ABSI, BMI, and waist circumference exhibit distinct patterns of association with cardiovascular physiological parameters in individuals with normal g-MPS findings. These findings suggest that measures reflecting abdominal fat distribution may be associated with physiological variations within a cardiovascularly normal population. However, the study design and the inclusion of only individuals with normal g-MPS findings limit the ability to draw causal inferences from the observed associations.

CONCLUSION

In this study, the relationships between ABSI, conventional anthropometric measures, and gated myocardial perfusion SPECT parameters were evaluated in individuals with normal myocardial perfusion findings. The results suggest that ABSI, as a complementary index to traditional anthropometric measures, may contribute to the assessment of relationships between abdominal fat distribution and cardiovascular physiological characteristics. The use of ABSI in a clinically normal population may help provide a more comprehensive evaluation of body shape–related cardiovascular features.

Ethical approval

This study has been approved by the Bolu Abant Izzet Baysal University Clinical Research Ethics Committee (approval date 11/04/2023, number 2023/100). Written informed consent was obtained from the participants.

Author contribution

Concept: SSM, BiC, HanA; Design: SSM, BiC, HamA; Data Collection or Processing: SSM, BiC, HanA, BeC; Analysis or Interpretation: HanA; Literature Search: SSM, BiC, HamA, HanA, BeC; Writing: SSM, BiC, HamA, HanA, BeC. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Sciagrà R. The expanding role of left ventricular functional assessment using gated myocardial perfusion SPECT: the supporting actor is stealing the scene. Eur J Nucl Med Mol Imaging. 2007; 34(7): 1107-22. [Crossref]
2. Abidov A, Germano G, Hachamovitch R, Slomka P, Berman DS. Gated SPECT in assessment of regional and global left ventricular function: an update. J Nucl Cardiol. 2013; 20(6): 1118-43; quiz 1144-6. [Crossref]
3. Lee DS, Cheon GJ, Ahn JY, Chung JK, Lee MC. Reproducibility of assessment of myocardial function using gated 99Tc(m)-MIBI SPECT and quantitative software. Nucl Med Commun. 2000; 21(12): 1127-34. [Crossref]
4. Hung CS, Wu YW, Huang JY, Hsu PY, Chen MF. Evaluation of circulating adipokines and abdominal obesity as predictors of significant myocardial ischemia using gated single-photon emission computed tomography. PLoS One. 2014; 9(5): e97710. [Crossref]
5. Shaw LJ, Iskandrian AE. Prognostic value of gated myocardial perfusion SPECT. J Nucl Cardiol. 2004; 11(2): 171-85. [Crossref]
6. Chartrand DJ, Murphy-Després A, Alméras N, Lemieux I, Larose E, Després JP. Overweight, obesity, and CVD Risk: a focus on visceral/ectopic fat. Curr Atheroscler Rep. 2022; 24(4): 185-95. [Crossref]

7. Upadhyay MP, Upadhyay SK, Bhattarai JR, Khatri B, Shrestha R. Abdominal obesity among outpatients in a tertiary level eye ENT hospital: a descriptive cross-sectional study. JNMA J Nepal Med Assoc. 2022; 60(245): 63-7. [Crossref]
8. Jayant SS, Gupta R, Rastogi A, et al. Abdominal obesity and incident cardio-metabolic disorders in Asian-Indians: a 10-years prospective cohort study. Diabetes Metab Syndr. 2022; 16(2): 102418. [Crossref]
9. Gažarová M, Galšneiderová M, Mečiarová L. Obesity diagnosis and mortality risk based on a body shape index (ABSI) and other indices and anthropometric parameters in university students. Rocz Panstw Zakl Hig. 2019; 70(3): 267-75. [Crossref]
10. Ji M, Zhang S, An R. Effectiveness of A Body Shape Index (ABSI) in predicting chronic diseases and mortality: a systematic review and meta-analysis. Obes Rev. 2018; 19(5): 737-59. [Crossref]
11. Kunutsor SK, Bhattacharjee A, Jae SY, Laukkanen JA. A paradoxical association between A Body Shape Index and cardiometabolic multimorbidity: findings from the English longitudinal study of ageing. Nutr Metab Cardiovasc Dis. 2025; 35(11): 104167. [Crossref]
12. Tian J, Wang X, Wen X, Gao B. An optimized anthropometric index-a body shape index-cm-demonstrates superior performance in cardiovascular risk stratification. J Int Med Res. 2025; 53(5): 3000605251343018. [Crossref]
13. Krakauer NY, Krakauer JC. A new body shape index predicts mortality hazard independently of body mass index. PLoS One. 2012; 7(7): e39504. [Crossref]
14. Knap B, Žvanut B, Brezočnik L, Jurdana M. The usefulness of a body shape index in assessing muscle function and strength in older adults hemodialysis patients. Front Endocrinol (Lausanne). 2025; 16: 1585193. [Crossref]
15. Liu H, Fu H, Wang Z, et al. The predictive value of a body shape index as a novel obesity metric for cancer risk: a systematic review and meta-analysis. Front Nutr. 2025; 12: 1667466. [Crossref]
16. Consultation, WE. Waist circumference and waist-hip ratio. Report of a WHO Expert Consultation. Geneva: World Health Organization; 2008: 8-11.
17. Christakoudi S, Tsilidis KK, Evangelou E, Riboli E. A Body Shape Index (ABSI), hip index, and risk of cancer in the UK Biobank cohort. Cancer Med. 2021; 10(16): 5614-28. [Crossref]
18. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. Lancet. 2017; 390(10113): 2627-42. [Crossref]
19. Katta N, Loethen T, Lavie CJ, Alpert MA. Obesity and coronary heart disease: epidemiology, pathology, and coronary artery imaging. Curr Probl Cardiol. 2021; 46(3): 100655. [Crossref]
20. Wang ZJ, Zhou YJ, Galper BZ, Gao F, Yeh RW, Mauri L. Association of body mass index with mortality and cardiovascular events for patients with coronary artery disease: a systematic review and meta-analysis. Heart. 2015; 101(20): 1631-8. [Crossref]
21. Uretsky S, Messerli FH, Bangalore S, et al. Obesity paradox in patients with hypertension and coronary artery disease. Am J Med. 2007; 120(10): 863-70. [Crossref]
22. Nevill AM, Stewart AD, Olds T, Holder R. Relationship between adiposity and body size reveals limitations of BMI. Am J Phys Anthropol. 2006; 129(1): 151-6. [Crossref]
23. Heymsfield SB, Scherzer R, Pietrobelli A, Lewis CE, Grunfeld C. Body mass index as a phenotypic expression of adiposity: quantitative contribution of muscularity in a population-based sample. Int J Obes (Lond). 2009; 33(12): 1363-73. [Crossref]
24. Gómez-Ambrosi J, Silva C, Galofré JC, et al. Body mass index classification misses subjects with increased cardiometabolic risk factors related to elevated adiposity. Int J Obes (Lond). 2012; 36(2): 286-94. [Crossref]
25. He H, Chen Y, Liao Y, Hu L, Qin H, Yang R. Association between body shape index and coronary heart disease in individuals over 20 years old with obese. J Health Popul Nutr. 2024; 43(1): 123. [Crossref]
26. Aoki KC, Mayrovitz HN. Utility of a body shape index parameter in predicting cardiovascular disease risks. Cureus. 2022; 14(4): e23886. [Crossref]
27. Ruiz JR, Ortega FB. Physical activity and cardiovascular disease risk factors in children and adolescents. Curr Cardio Risk Rep. 2009; 3: 281-7. [Crossref]
28. Boden G, Salehi S. Why does obesity increase the risk for cardiovascular disease? Curr Pharm Des. 2013; 19(32): 5678-83. [Crossref]
29. Chuang HH, Li WC, Sheu BF, et al. Correlation between body composition and risk factors for cardiovascular disease and metabolic syndrome. Biofactors. 2012; 38(4): 284-91. [Crossref]
30. Poirier P, Després JP. Waist circumference, visceral obesity, and cardiovascular risk. J Cardiopulm Rehabil. 2003; 23(3): 161-9. [Crossref]
31. Rezende FAC, Rosado LEFPL, Ribeiro RDCL, et al. Body mass index and waist circumference: association with cardiovascular risk factors. Arq Bras Cardiol. 2006; 87(6): 728-34. [Crossref]
32. Du SM, Ma GS, Li YP, et al. Relationship of body mass index, waist circumference and cardiovascular risk factors in Chinese adult. Biomed Environ Sci. 2010; 23(2): 92-101. [Crossref]
33. Wilmot KA, O'Flaherty M, Capewell S, Ford ES, Vaccarino V. Coronary heart disease mortality declines in the United States from 1979 through 2011: evidence for stagnation in young adults, especially women. Circulation. 2015; 132(11): 997-1002. [Crossref]

34. Han TS, Lean ME. A clinical perspective of obesity, metabolic syndrome and cardiovascular disease. JRSM Cardiovasc Dis. 2016; 5: 2048004016633371. **[Crossref]**

35. Pepine CJ, Merz CNB, El Hajj S, et al. Heart failure with preserved ejection fraction: similarities and differences between women and men. Int J Cardiol. 2020; 304: 101-8. **[Crossref]**

36. Schorr M, Dichtel LE, Gerweck AV, et al. Sex differences in body composition and association with cardiometabolic risk. Biol Sex Differ. 2018; 9(1): 28. **[Crossref]**

37. Ibrahim MM. Subcutaneous and visceral adipose tissue: structural and functional differences. Obes Rev. 2010; 11(1): 11-8. **[Crossref]**

38. Mancuso P, Bouchard B. The impact of aging on adipose function and adipokine synthesis. Front Endocrinol (Lausanne). 2019; 10: 137. **[Crossref]**

39. Pi-Sunyer FX. The epidemiology of central fat distribution in relation to disease. Nutr Rev. 2004; 62(7 Pt 2): S120-6. **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):192-202
DOI: 10.54307/NWMJ.2026.229

RESEARCH ARTICLE

Evaluation of artificial intelligence chatbots in informing patients about CBCT: A comparative study

Savaş Özarslantürk[1], Seval Ceylan Şen[2], Özlem Saraç Atagün[2], Selin Gaş[3], Tuğçe Paksoy[4], Şeyma Çardakcı Bahar[2]

[1]Department of Oral and Maxillofacial Radiology, Gülhane Faculty of Dentistry, University of Health Sciences, Ankara, Türkiye
[2]Department of Periodontology, Gülhane Faculty of Dentistry, University of Health Sciences, Ankara, Türkiye
[3]Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, İstanbul Gelişim University, İstanbul, Türkiye
[4]Department of Periodontology, Hamidiye Faculty of Dentistry, University of Health Sciences, İstanbul, Türkiye

Cite as: Özarslantürk S, Ceylan Şen S, Saraç Atagün Ö, Gaş S, Paksoy T, Çardakcı Bahar Ş. Evaluation of artificial intelligence chatbots in informing patients about CBCT: A comparative study. Northwestern Med J. 2026;6(3):192-202.

ABSTRACT

Objective: This study aimed to explore common patient inquiries about Cone-Beam Computed Tomography (CBCT) and systematically assess the accuracy, quality, usefulness, and readability of outputs from four AI-based conversational platforms (ChatGPT-3.5, DeepSeek-V3.1, Gemini 1.5 Flash, and Microsoft Copilot Free).

Materials and Methods: Common questions about CBCT were gathered from online sources and expert contributions and submitted to ChatGPT-3.5, DeepSeek, Gemini, and Copilot under standardized conditions. Outputs were independently evaluated by three specialists using CLEAR, mGQS, accuracy, usefulness, DISCERN, and readability metrics (FRE and FKGL).

Results: Significant differences were observed among AI platforms regarding CLEAR scores ($p < 0.05$), with Gemini showing higher values than ChatGPT ($p = 0.016$). Across all four platforms, strong to a very strong positive correlations were found among CLEAR, mGQS, and Accuracy scores ($r \geq 0.77$, $p < 0.001$), while these measures were strongly to very strongly and inversely correlated with Usefulness ($r \leq -0.77$, $p < 0.001$). Flesch Reading Ease and Flesch-Kincaid Grade Level demonstrated a very strong negative correlations across platforms (r ranging from −0.85 to −0.93, $p < 0.05$). Within the Gemini group, Flesch-Kincaid Grade Level showed a moderate positive association with DISCERN scores ($r = 0.496$, $p = 0.043$).

Conclusions: Gemini and DeepSeek produced more accurate and clearer responses. The readability of all AI systems was below the recommended level for patient education. These systems may support patient education but require expert validation.

Keywords: CBCT, artificial intelligence, chatbots, patient education, readability

Corresponding author: Savaş Özarslantürk **E-mail:** savasozarslanturk@yahoo.com
Received: 10.11.2025 **Accepted:** 03.07.2026 **Published:** 31.07.2026

INTRODUCTION

Cone-beam computed tomography (CBCT) is a three-dimensional (3D) imaging technique that offers high spatial resolution at a lower radiation dose relative to standard computed tomography (CT) scans. These features form the basis for its widespread use in dentistry and maxillofacial imaging. It enables detailed evaluation of maxillofacial bone structures, facilitates accurate diagnosis and staging, and supports lesion follow-up (1).

The increasing use of CBCT has heightened patient demand for information about imaging procedures. As access to healthcare professionals may be limited due to increasing workloads, patients increasingly seek information from digital resources, including artificial intelligence (AI)-based chat systems (2).

Broadly speaking, AI encompasses computational systems built to approximate specific human cognitive abilities, drawing on subfields such as Natural Language Processing (NLP), Machine Learning, and Deep Learning to do so. Recent advances have enabled AI models to analyze complex information, interpret, perform reasoning, and generate predictions (3). ChatGPT, introduced by OpenAI in 2022, attracted broad interest with its advanced text-generation capabilities (4). In 2024, Google released Gemini, an AI platform built with multimodal capabilities that allow it to work across text, visual, and graphical content types (5). DeepSeek, developed in China in the same year, was primarily designed for academic and professional applications in the fields of mathematics, engineering, science, and technology (6). Microsoft Copilot, built on advanced GPT-based models, integrates with Microsoft 365 and Windows 11 and has potential applications in optimizing healthcare workflows (7).

Patients frequently inquire about the indications, applications, benefits, and radiation risks of CBCT. Therefore, evaluating the quality of AI-derived information is increasingly important. Despite these advances, the extent to which AI-generated answers can be considered accurate, dependable, clear, and user-friendly remains a subject of ongoing debate (8).

Four widely used AI platforms were evaluated in this study based on how they responded to patient questions commonly raised about CBCT, with comparisons drawn across accuracy, explanatory quality, understandability, readability, reliability, and usability. By doing so, the study seeks to clarify both the strengths and shortcomings of current AI technologies in a healthcare context—information that may prove useful to clinicians and developers alike.

MATERIALS AND METHODS

The study followed the ethical standards of the Declaration of Helsinki, and approval from an ethics committee was not deemed necessary. Common patient questions about CBCT, covering its indications, applications, procedure, diagnostic value, radiation risks, interpretation of findings, comparison with alternative imaging methods, and insurance-related issues, were submitted to ChatGPT-3.5, DeepSeek-V3.1, Gemini 1.5 Flash, and Microsoft Copilot Free (GPT-4o–based).

AI platform evaluation and data extraction protocol

A four-stage protocol (identification, selection, refinement, and finalization) was used to generate a standardized dataset of patient questions regarding CBCT. During the identification phase, potential questions were gathered from public health forums, frequently asked questions (FAQs) on dental imaging websites, online search trends, and clinical inquiries developed by an expert panel of dental specialists. These sources were selected to reflect both common patient concerns and clinically relevant topics related to CBCT examinations. In the selection phase, all questions were compiled into a master database (Microsoft Excel), and two authors (S.C.Ş., Ö.S.A.) independently screened entries to remove duplicates and semantically overlapping items, resolving discrepancies by consensus. During refinement, grammatical errors were corrected and technical terminology was adapted to patient-friendly language to reflect queries from individuals with average health literacy. In the finalization phase, all questions were formatted as open-ended prompts without additional context or instructions to ensure consistency across platforms.

The final dataset consisted of 17 open-ended questions designed to represent the most common patient concerns regarding CBCT examinations. The questions covered multiple domains, including general information about CBCT, clinical indications, procedural aspects, diagnostic applications, radiation safety, vulnerable patient populations (such as children and pregnant individuals), interpretation of findings, comparison with alternative imaging modalities, and insurance-related issues. This broad thematic coverage was intended to reflect real-world patient information needs and simulate typical patient–provider discussions regarding CBCT.

ChatGPT-3.5 (OpenAI; https://chat.openai.com/), Gemini 1.5 Flash (Google; https://gemini.google.com/), DeepSeek-V3.1 (https://chat.deepseek.com/), and Microsoft Copilot Free (https://copilot.microsoft.com/) were accessed in incognito mode using newly created email accounts and default settings. No prompt engineering, plugins, parameter adjustments, or response regeneration were used. Questions were entered sequentially in separate "New Chat" sessions and submitted once to each chatbot on August 24, 2025, between 09:00 and 16:00 (S.G.). Responses were recorded verbatim in Microsoft Word and analyzed in their original form. All data generated for this study are included in the Supplementary Materials to facilitate transparency and reproducibility.

Outcome measures and expert assessment criteria

Three independent experts (T.P., G.U., and S.Ö.), none of whom participated in the question development process, evaluated all responses produced by the AI models. Assessments were performed using the CLEAR criteria, Usefulness scale, modified Global Quality Score (mGQS), accuracy ratings, DISCERN instrument, and readability indices (FRE and FKGL).

Response reliability and overall quality were gauged through the CLEAR criteria, which rate five domains—completeness, misinformation, evidence-based support, appropriateness, and relevance—on a 5-point scale ranging from 1 (very poor) to 5 (excellent). Summed scores can range from 5 to 25 and are grouped into low- (5-11), moderate- (12-18), or high-quality (19-25) categories (9).

Content quality was quantified via the modified Global Quality Score (mGQS), a 5-point scale spanning from 1 (strongly disagree; inaccurate or irrelevant response) to 5 (strongly agree; comprehensive and fully accurate response) (10). Meanwhile, Accuracy was captured using a distinct 5-point Likert scale, where higher ratings indicated a greater degree of factual correctness (11).

A 4-point scale—very useful, useful, partially useful, or not useful—was used to evaluate Usefulness, with classification based on the completeness and accuracy of the information conveyed (12).

Readability metrics were evaluated via two indices, FRE and FKGL, generated through Readable Pro software (https://app.readable.com). The FRE scale runs from 0 to 100, with higher scores denoting greater readability, while the FKGL reflects the grade level a reader would need to comprehend the text. Sentence length and word difficulty underlie both measures. Given accessibility concerns, patient-facing health content is typically advised to stay at or below an eighth-grade level (13).

Statistical analysis

Descriptive statistics for continuous variables were reported as means with standard deviations (SDs), or as medians with 25th–75th percentiles, depending on distribution. The Shapiro-Wilk test was used to assess normality. The Kruskal-Wallis test was applied for between-group comparisons; where significant differences emerged, Bonferroni-adjusted post hoc analyses were subsequently conducted. The Intraclass Correlation Coefficient (ICC) was employed to evaluate agreement across the three raters, while relationships among continuous variables were assessed via Spearman's rank correlation. All analyses were conducted with IBM SPSS Statistics version 27. Given the exploratory nature of this study, which sought to compare the performance of commonly used large language models in responding to CBCT-related patient questions, no a priori sample size or power calculation was carried out.

RESULTS

Inter-rater agreement demonstrated high to excellent reliability across all subjective assessment tools. Single-measure ICC values for CLEAR, mGQS, Accuracy, Usefulness, and DISCERN ranged from 0.82 to 0.94 (all $p < 0.001$). These values demonstrate strong consistency among the three evaluators.

Response characteristics according to AI platform are presented in Table 1. Significant between-platform variation in CLEAR scores was detected using the Kruskal-Wallis test ($p < 0.05$). When Bonferroni-adjusted post hoc comparisons were applied, Gemini was found to score significantly higher than ChatGPT ($p = 0.016$). None of the other pairwise comparisons showed statistical significance (Table 1).

Spearman correlation analysis of ChatGPT responses demonstrated significant positive correlations between CLEAR scores and both mGQS and Accuracy ($r = 0.768$ for both, $p < 0.001$), whereas CLEAR was negatively correlated with Usefulness ($r = -0.768$, $p < 0.001$). mGQS showed a very strong positive correlation with Accuracy ($r = 1.000$, $p < 0.001$) and a very strong negative correlation with Usefulness ($r = -1.000$, $p < 0.001$) and FKGL ($r = -0.522$, $p = 0.010$). Similarly, Accuracy was very strongly negatively correlated with Usefulness ($r = -1.000$, $p < 0.001$) and moderately negatively correlated with FKGL ($r = -0.522$, $p = 0.010$), whereas Usefulness was moderately positively correlated with FKGL ($r = 0.522$, $p = 0.010$). In addition, FRE and FKGL were very strongly and negatively correlated ($r = -0.929$, $p < 0.001$) (Table 2).

Correlation patterns observed for Copilot closely mirrored those identified for ChatGPT. CLEAR scores were positively correlated with both mGQS and Accuracy ($r = 0.960$ for both, $p < 0.001$) and negatively correlated with Usefulness ($r = -0.960$, $p < 0.001$). A very strong positive correlation was found between mGQS and Accuracy ($r = 1.000$, $p < 0.001$), while both measures showed a very strong negative correlations with Usefulness ($r = -1.000$, $p < 0.001$). Readability indices also demonstrated a very strong inverse association between FRE and FKGL ($r = -0.927$; $p = 0.004$). Additionally, moderate positive correlation was observed between FKGL and DISCERN ($r = 0.488$; $p = 0.047$) (Table 3).

Table 1. Distribution and comparison of features of responses according to AI types

	ChatGPT		CoPilot		DeepSeek		Gemini			
	Mean ± SD	M. (%25-%75Q.)	Mean ± SD.	M. (%25-%75Q.)	Mean ± SD	M. (%25-%75Q.)	Mean ± SD.	M. (%25-%75Q.)	Test Statistics	p
Clear	17.47±8.85	24(11-25)	24.06±1.48	25(24-25)	22±4.14	25(19-25)	24.65±0.86	25(25-25)	9.394	0.024*
GQS	3.59±1.66	4(2-5)	4.59±0.51	5(4-5)	4.24±1.03	5(3-5)	4.82±0.39	5(5-5)	7.160	0.067
Accuracy	3.59±1.66	4(2-5)	4.59±0.51	5(4-5)	4.24±1.03	5(3-5)	4.82±0.39	5(5-5)	7.160	0.067
Usefulness	2.41±1.66	2(1-4)	1.41±0.51	1(1-2)	1.76±1.03	1(1-3)	1.18±0.39	1(1-1)	7.160	0.067
Flesch Reading Ease Score	59,19±25,7	63,1(42,6-77,3)	51,48±22,34	55,9(38,5-68,8)	54,2±18,74	52(44,4-68,4)	51,8±14,36	52,1(45,4-58,8)	1,426	0,699
Flesch Reading Grade Level	7,88±3,4	6,7(5,2-11,1)	15,76±24,74	9,6(8,2-11,5)	8,34±2,78	8(6,2-11,1)	9,64±1,84	10(8,4-11)	4,940	0,176
DISCERN	25.65±2.06	25(25-26)	27.88±2.39	28(27-29)	27.29±3.42	28(24-30)	27±3.1	28(24-28)	4.088	0.252

*p<0.05

Table 2. Correlations among ChatGPT-related measurements

		GQS	Accuracy	Usefulness	Flesch Reading Ease Score	Flesch Reading Grade Level	DISCERN
Clear	r	0.768	0.768	-0.768	0.440	-0.456	0.060
	p	<0.001*	<0.001*	<0.001*	0.077	0.066	0.819
GQS	r		1.000	-1.000	0.471	-0.522	0.305
	p			<0.001*	0.056	0.032*	0.234
Accuracy	r			-1.000	0.471	-0.522	0.305
	p			<0.001*	0.056	0.032*	0.234
Usefulness	r				-0.471	0.522	-0.305
	p				0.056	0.032*	0.234
Flesch Reading Ease Score	r					-0.929	-0.181
	p					<0.001*	0.486
Flesch-Kincaid Grade Level	r						-0.170
	p						0.515

*p<0.05, and r: correlation coefficient

Table 3. Correlations among Copilot-related measurements

		GQS	Accuracy	Usefulness	Flesch Reading Ease Score	Flesch Reading Grade Level	DISCERN
Clear	r	0.960	0.960	-0.960	-0.202	0.244	0.231
	p	<0.001*	<0.001*	<0.001*	0.436	0.345	0.373
GQS	r		1.000	-1.000	-0.244	0.256	0.178
	p		<0.001*	<0.001*	0.345	0.321	0.495
Accuracy	r			-1.000	-0.244	0.256	0.178
	p			<0.001*	0.345	0.321	0.495
Usefulness	r				0.244	-0.256	-0.178
	p				0.345	0.321	0.495
Flesch Reading Ease Score	r					-0.927	-0.373
	p					<0.001*	0.140
Flesch-Kincaid Grade Level	r						0.488
	p						0.047*

*p<0.05 ,and r: correlation coefficient

Similar findings were observed for DeepSeek responses. CLEAR scores were strongly positively associated with both mGQS and Accuracy (r = 0.980 for both, p < 0.001) and negatively associated with Usefulness (r = −0.980, p < 0.001). A very strong positive correlation was identified between mGQS and Accuracy (r = 1.000, p < 0.001), whereas both measures showed a very strong negative correlations with Usefulness (r = −1.000, p < 0.001). In addition, FRE and FKGL demonstrated a very strong inverse correlation (r = −0.923, p = 0.028) (Table 4).

Gemini responses exhibited the strongest overall correlation pattern among the evaluated platforms.

CLEAR scores were strongly positively associated with both mGQS and Accuracy (r = 0.994 for both, p < 0.001) and negatively associated with Usefulness (r = −0.994, p < 0.001). A very strong positive correlation was observed between mGQS and Accuracy (r = 1.000, p < 0.001), whereas both measures demonstrated a very strong negative correlations with Usefulness (r = −1.000, p < 0.001). Furthermore, FRE and FKGL were very strongly and inversely correlated (r = −0.849, p < 0.001) (Table 5).

Table 4. Correlations among DeepSeek-related measurements

		GQS	Accuracy	Usefulness	Flesch Reading Ease Score	Flesch Reading Grade Level	DISCERN
Clear	r	0.980	0.980	-0.980	0.104	0.079	-0.099
	p	<0.001*	<0.001*	<0.001*	0.690	0.763	0.706
GQS	r		1.000	-1.000	-0.025	0.190	-0.086
	p		<0.001*	<0.001*	0.924	0.466	0.742
Accuracy	r			-1.000	-0.025	0.190	-0.086
	p			<0.001*	0.924	0.466	0.742
Usefulness	r				0.025	-0.190	0.086
	p				0.924	0.466	0.742
Flesch Reading Ease Score	r				1.000	-0.923	0.072
	p					<0.001*	0.783
Flesch-Kincaid Grade Level	r						-0.139
	p						0.595

*p<0.05, and r: correlation coefficient

Table 5. Correlations among Gemini-related measurements

		GQS	Accuracy	Usefulness	Flesch Reading Ease Score	Flesch Reading Grade Level	DISCERN
Clear	r	0.994	0.994	-0.994	0.356	-0.094	0.078
	p	<0.001*	<0.001*	<0.001*	0.161	0.719	0.766
GQS	r		1.000	-1.000	0.346	-0.063	0.115
	p		<0.001*	<0.001*	0.173	0.810	0.659
Accuracy	r			-1.000	0.346	-0.063	0.115
	p			<0.001*	0.173	0.810	0.659
Usefulness	r				-0.346	0.063	-0.115
	p				0.173	0.810	0.659
Flesch Reading Ease Score	r				1.000	-0.849	-0.310
	p					<0.001*	0.225
Flesch-Kincaid Grade Level	r						0.335
	p						0.189

*p<0.05, and r: correlation coefficient

DISCUSSION

Diagnostic decision-making is based on clinical history, physical examination, together with laboratory and imaging findings, to identify the underlying pathology. In dentistry, CBCT has become an essential imaging modality that enhances diagnostic accuracy by providing precise three-dimensional visualization of oral and maxillofacial structures (14). Considering these factors, ChatGPT cannot independently diagnose medical or dental conditions. However, it may serve as an assisting tool for clinicians. While it cannot replace medical practitioners who rely on communication, observation, and questioning to obtain patient history, ChatGPT can help generate clinical notes, summaries, and documentation, thereby saving time and reducing the risk of human error. Furthermore, it may offer therapy recommendations based on patient data (15). With technological advancement, patients increasingly seek information from AI tools, which may provide quick but sometimes inaccurate responses. Therefore, caution is needed when using such technologies in medical contexts. The empathy and human connection between physician and patient remain irreplaceable aspects of healthcare that artificial systems cannot replicate (16). Artificial intelligence chatbots can process a wide range of query formats, including multiple-choice, open-ended, and yes/no questions (17). The accuracy, quality, usefulness, and readability of responses produced by four AI-powered conversational platforms—ChatGPT-3.5, DeepSeek, Gemini, and Copilot—were comprehensively evaluated in this study, alongside an effort to identify the most commonly asked patient questions concerning CBCT. In our study design, questions were asked in an open-ended manner because this better reflects the conversation between the patient and the doctor.

To thoroughly evaluate the clinical safety and educational efficacy of AI-generated responses regarding CBCT, a multi-dimensional assessment framework was intentionally constructed. Each selected evaluation instrument was justified by its ability to capture a distinct and non-overlapping dimension of information quality. While the Accuracy Scale strictly audited the medical and technical correctness of the dental data provided, the modified Global Quality Score (mGQS) focused on the structural integrity, flow, and contextual completeness of the information. To prevent clinical misalignment, the Usefulness Scale was integrated to assess the pragmatic value of responses from a clinician's perspective, distinguishing between completely accurate answers that are clinically helpful and potentially vague to a patient. Furthermore, the CLEAR criteria provided an objective, domain-specific evaluation of reliability across five specific criteria, including evidence-based support and relevance, whereas the DISCERN instrument uniquely quantified the broader systemic reliability and quality of the text as an educational consumer health resource. Finally, because technically accurate and structurally high-quality text can remain inaccessible to the lay public, readability indices (FRE and FKGL) were utilized to provide an objective, mathematical computation of linguistic complexity based on syllable count and sentence length. Together, this comprehensive matrix of metrics ensured that the chatbots were not merely judged on superficial correctness but were rigorously screened for structural quality, clinical utility, trustworthiness, and actual patient accessibility.

In our study, ChatGPT demonstrated strong, positive correlations between the Clear score and both mGQS and Accuracy, with a negative correlation with Usefulness. CoPilot, DeepSeek, and Gemini exhibited progressively stronger positive correlations with mGQS and Accuracy, alongside negative correlations with Usefulness, with Gemini showing the strongest overall associations. The findings of our study are broadly consistent with the available literature on AI chatbot performance in healthcare. Hassona et al.(18) reported that ChatGPT achieved strong results in oral cancer detection based on the CLEAR assessment. Similarly, Sallam et al.(9) found that ChatGPT-4 and Microsoft Bing outperformed ChatGPT-3 and Google Bard (Gemini), which received comparatively lower evaluation scores. Esmailpour et al.(19) observed comparable GQS values across DeepSeek-R1, ChatGPT-o1, ChatGPT-4, and Dental GPT, indicating robust performance. Similarly, a recent comparative study on dental implant failure found that Gemini achieved the highest accuracy, ChatGPT-4 the highest reliability, and DeepSeek-R1 the best readability, underscoring performance differences among AI models (20). Güven et al.(21) also reported that ChatGPT-4 and Google Gemini outperformed

ChatGPT-3.5 in traumatic dental injury scenarios, with ChatGPT-4 providing the most comprehensive treatment recommendations. Hassona et al.(18) further noted high GQS scores for ChatGPT, confirming the reliability of its outputs. Some studies reported differing results. Dursun et al.(22) found CoPilot achieved the highest GQS scores, outperforming ChatGPT-4 and ChatGPT-3.5. Mohammad-Rahimi et al.(23) observed that Claude obtained the highest mean GQS, followed by GPT-4, while Bing consistently underperformed; they also highlighted that a notable proportion of chatbot-generated citations were fabricated. Ozcivelek et al.(24) and Kinikoğlu (3) emphasized that specialized models, such as Dental GPT, ChatGPT-o1, and Gemini 2.0, achieved superior accuracy compared to general-purpose models. Yilmaz et al.(25) similarly found that ChatGPT-o1 had the highest accuracy, followed by Claude, Gemini 2, and DeepSeek, while CoPilot ranked lowest. Çitir (26) reported that ChatGPT-3 and 3.5 frequently produced unverifiable citations, reflecting limitations in real-time access to data.

Singal et al.(27) compared the performance of Google Gemini 2.0 Flash, Google Gemini 2.5 Flash, ChatGPT-4o, and DeepSeek V3 on multiple-choice anatomy questions and reported an overall accuracy rate of 95.9%. Among the models evaluated, Google Gemini 2.5 Flash achieved the highest overall accuracy rate, followed in descending order by ChatGPT-4o, DeepSeek V3, and Gemini 2.0 Flash. These findings further support evidence indicating that successive generations of large language models continue to demonstrate improvements in performance across healthcare-related knowledge and reasoning tasks.

Overall, our results and the literature indicate that newer AI models offer high clarity, accuracy, and response quality, but caution is warranted due to occasional inaccuracies and unverifiable outputs, underscoring the continued need for expert oversight.

From a dental practice perspective, these findings have important implications for patient communication and education regarding CBCT examinations. As patients increasingly seek information from AI-based platforms before or after dental consultations, chatbots may serve as supplementary educational resources that help explain the purpose, indications, benefits, and limitations of CBCT imaging. In particular, they may assist patients in understanding how CBCT contributes to the evaluation of endodontic conditions, cystic lesions, impacted teeth, and maxillofacial trauma. However, the variability observed among chatbot responses, together with the presence of occasional inaccuracies and readability challenges, suggests that AI-generated information should be interpreted as a supportive resource rather than a replacement for dentist-led clinical evaluation. Dentists should therefore be aware that patients may arrive with information obtained from AI chatbots and should be prepared to verify, clarify, and contextualize this information within the framework of individualized clinical decision-making.

In our study, ChatGPT demonstrated strong relationships among readability, quality, and usefulness metrics. A moderate negative correlation was observed between GQS and Flesch Reading Grade Level, and Accuracy showed a very strong negative correlation with Usefulness. Additionally, Accuracy negatively correlated with the Flesch Reading Grade Level, while Usefulness positively correlated with it. A very strong negative correlation was also found between Flesch Reading Ease and Flesch-Kincaid Grade Level. Similarly, CoPilot, DeepSeek, and Gemini displayed a very strong negative correlation between Accuracy and Usefulness. These negative correlations between Accuracy and Usefulness should be interpreted in light of the inverse scoring structure of the Usefulness scale, where lower scores indicate greater usefulness. Therefore, higher Accuracy scores was associated with lower Usefulness scores.

A very strong inverse correlation between Flesch Reading Ease and Flesch-Kincaid Grade Level scores was evident across CoPilot, DeepSeek, and Gemini alike ($r = -0.927$, -0.923, and -0.849, respectively). For CoPilot specifically, Flesch-Kincaid Grade Level and DISCERN scores were also moderately and positively correlated, pointing to a possible link between higher reading grade levels and greater content reliability.

These findings are consistent with previous studies evaluating AI-generated readability. Hassona et al.(18) reported that ChatGPT's responses on oral cancer

detection were difficult to understand, with Flesch-Kincaid scores indicating a higher reading grade level. According to Şahin et al.(28), patients with low health literacy may find it difficult to understand content generated by ChatGPT-4, as the responses often exceed a Year 6 reading level. A comparable pattern emerged in Helvacıoğlu-Yiğit et al.'s (13) work, where none of the three models—ChatGPT-3.5, Gemini, or Copilot—produced text below an eighth-grade reading level. Güven et al.(21) also noted that Gemini had slightly higher readability than ChatGPT models, while Özcivelek et al.(24) observed that DeepSeek-R1 achieved the highest Flesch Reading Ease scores, and Dental GPT performed the worst among the tested models. Esmailpour et al.(19) further confirmed that ChatGPT responses were more complex than Gemini or Copilot, with Gemini providing the most patient-friendly outputs.

Sezer et al.(29) reported that DeepSeek achieved the highest Flesch Reading Ease Score (FRES) among the evaluated models. DeepSeek's output was nevertheless rated as "difficult" in terms of readability, placing it at a high school reading level. In addition, they reported that most of the chatbot outputs did not align with the 6th–8th grade readability standard widely advised for patient-facing educational content. Similarly, Taşyürek et al.(30) compared different question types and versions of large language models (LLMs) and found that DeepSeek-V3.1 produced the highest FRES. Despite this, none of the evaluated chatbots generated text at the recommended 6th–8th grade readability level. Nonetheless, DeepSeek-V3.1 was identified as the model that produced results closest to these thresholds. The authors suggested that this finding may be attributed to DeepSeek-V3.1 being a free, publicly accessible platform, which is likely trained on a broader, more diverse dataset encompassing various question formats and narrative styles.

This research is subject to a number of limitations. Notably, even with standardized phrasing and evaluation criteria, factors such as query wording, timing, and context may still lead to variation in AI-generated responses. Second, the full range of diversity, complexity, and information needs across all patient populations may not be captured by the selected questions, even though these were designed to reflect common patient inquiries about CBCT. As such, the scope of the included question set should be taken into account when interpreting these findings. Third, the reproducibility and generalizability of these findings may be limited, given that the chatbots assessed here undergo continuous updates that can alter their performance and response quality over time. Fourth, the evaluation was based on simulated patient inquiries rather than real-world patient interactions. Consequently, the actual comprehensibility, usefulness, and educational value of the responses from a patient perspective could not be directly assessed. Although inter-rater agreement was high, several assessment tools, including Accuracy, Usefulness, mGQS, CLEAR, and DISCERN, involve expert judgment and may therefore retain an inherent degree of subjectivity. A further limitation lies in the readability indices themselves. Metrics like the Flesch Reading Ease and Flesch–Kincaid Grade Level standardize the measurement of textual complexity, yet they do not account for factors such as health literacy, conceptual understanding, cultural background, or actual comprehension by patients. Additionally, this analysis was restricted to English-language material; readability outcomes and interpretive nuances of AI-generated content are likely to differ in other linguistic and cultural contexts.

CONCLUSION

Within the scope of this study, the evaluated AI-based conversational platforms generally achieved favorable scores across measures of accuracy, quality, readability, and educational value when responding to patient-oriented questions about CBCT. Performance varied across the platforms examined; in particular, more recently released models such as Gemini and DeepSeek-V3.1 tended to score higher across multiple assessment criteria. Nevertheless, no chatbot reliably generated content that fell within the 6th–8th grade readability band advised for patient-facing education materials, suggesting this may pose a challenge for individuals with lower health literacy.

A number of methodological limitations should be taken into account when interpreting these findings—namely, the specific question set used, reliance on expert-based assessment tools, the lack of testing with

real patients, and the fast-changing nature of large language models. While AI chatbots hold potential as supportive tools in patient education, information provision, and clinical assistance, these findings do not suggest that such systems are capable of delivering reliable clinical guidance on their own or standing in for professional expertise. Maintaining accuracy, appropriateness, and ethical standards in AI-generated health information still requires human review. Future research should incorporate real-world patient interactions, multilingual evaluations, and longitudinal assessments of evolving AI systems to better define their role in healthcare communication and education.

Generative AI statement

AI-assisted technologies (ChatGPT-3.5, OpenAI; Claude, Anthropic) were used solely for language editing and translation support during manuscript preparation; all scientific content, analysis, and interpretation remain the sole responsibility of the authors.

Ethical approval

This study, which did not involve human subjects, personal information, or animal experiments, was conducted entirely through AI-based data analysis using open and publicly available sources. Therefore, the authors state that ethical exemption was granted for this work, with the study carried out in line with the Declaration of Helsinki (2013).

Author contribution

Concept: SÖ, SCŞ; Design: SÖ, ÖSA; Data Collection or Processing: SÖ, SG, TP, ŞÇP; Analysis or Interpretation: SÖ, SCŞ, ÖSA; Literature Search: SÖ, SCŞ, ÖSA, SG, TP, ŞÇB; Writing: SÖ. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Ko YY, Yang WF, Leung YY. The role of Cone Beam Computed Tomography (CBCT) in the diagnosis and clinical management of Medication-Related Osteonecrosis of the Jaw (MRONJ). Diagnostics (Basel). 2024; 14(16): 1700. **[Crossref]**
2. Holmgren AJ, Byron ME, Grouse CK, Adler-Milstein J. Association between billing patient portal messages as e-visits and patient messaging volume. JAMA. 2023; 329(4): 339-42. **[Crossref]**
3. Kinikoglu I. Evaluating ChatGPT and Google Gemini performance and implications in Turkish dental education. Cureus. 2025; 17(1): e77292. **[Crossref]**
4. Miah ASM, Tusher MMR, Hossain MM, et al. ChatGPT in research and education: a SWOT analysis of its academic impact. CMES - Computer Modeling in Engineering and Sciences. 2025; 143(3): 2573-614. **[Crossref]**
5. Carlà MM, Gambini G, Baldascino A, et al. Large language models as assistance for glaucoma surgical cases: a ChatGPT vs. Google Gemini comparison. Graefes Arch Clin Exp Ophthalmol. 2024; 262(9): 2945-59. **[Crossref]**
6. Diniz-Freitas M, Diz-Dios P. DeepSeek: another step forward in the diagnosis of oral lesions. J Dent Sci. 2025; 20(3): 1904-7. **[Crossref]**
7. Semeraro F, Gamberini L, Carmona F, Monsieurs KG. Clinical questions on advanced life support answered by artificial intelligence. A comparison between ChatGPT, Google Bard and Microsoft Copilot. Resuscitation. 2024; 195: 110114. **[Crossref]**
8. Lewis P, Perez E, Piktus A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. NeurIPS. 2020: 9459–9474.
9. Sallam M, Barakat M, Sallam M. Pilot testing of a tool to standardize the assessment of the quality of health information generated by artificial intelligence-based models. Cureus. 2023; 15(11): e49373. **[Crossref]**
10. Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. Am J Gastroenterol. 2007; 102(9): 2070-7. **[Crossref]**
11. Hatia A, Doldo T, Parrini S, et al. Accuracy and completeness of ChatGPT-Generated information on interceptive orthodontics: a multicenter collaborative study. J Clin Med. 2024; 13(3): 735. **[Crossref]**
12. Yeo YH, Samaan JS, Ng WH, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. Clin Mol Hepatol. 2023; 29(3): 721-32. **[Crossref]**
13. Helvacioglu-Yigit D, Demirturk H, Ali K, Tamimi D, Koenig L, Almashraqi A. Evaluating artificial intelligence chatbots for patient education in oral and maxillofacial radiology. Oral Surg Oral Med Oral Pathol Oral Radiol. 2025; 139(6): 750-9. **[Crossref]**

14. Hampton JR, Harrison MJ, Mitchell JR, Prichard JS, Seymour C. Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. Br Med J. 1975; 2(5969): 486-9. **[Crossref]**

15. Khan RA, Jawaid M, Khan AR, Sajjad M. ChatGPT - Reshaping medical education and clinical management. Pak J Med Sci. 2023; 39(2): 605-7. **[Crossref]**

16. Qaiser N, Arshad R. Pandora box of possibilities and limitations: unleashing the power of chatgpt in healthcare. Pak J Phsyiol. 2025; 21(3): 58-63. **[Crossref]**

17. Bashah A, Salem A, Al-Waqeerah A, et al. Evaluation of deepseek, gemini, ChatGPT-4o, and perplexity in responding to salivary gland cancer. BMC Oral Health. 2025; 25(1): 1358. **[Crossref]**

18. Hassona Y, Alqaisi D, Al-Haddad A, et al. How good is ChatGPT at answering patients' questions related to early detection of oral (mouth) cancer? Oral Surg Oral Med Oral Pathol Oral Radiol. 2024; 138(2): 269-78. **[Crossref]**

19. Esmailpour H, Rasaie V, Babaee Hemmati Y, Falahchai M. Performance of artificial intelligence chatbots in responding to the frequently asked questions of patients regarding dental prostheses. BMC Oral Health. 2025; 25(1): 574. **[Crossref]**

20. Ceylan Şen S, Çardakci Bahar Ş, Saraç Atagün Ö, Ustaoğlu G, Yildiz ZH, Toker H. Quality and reliability of ai information on dental implant failure: a comparative multi-model analysis. J Craniofac Surg. 2026; 37(3-4): 675-80. **[Crossref]**

21. Guven Y, Ozdemir OT, Kavan MY. Performance of artificial intelligence chatbots in responding to patient queries related to traumatic dental injuries: a comparative study. dent traumatol. 2025; 41(3): 338-47. **[Crossref]**

22. Dursun D, Bilici Geçer R. Can artificial intelligence models serve as patient information consultants in orthodontics? BMC Med Inform Decis Mak. 2024; 24(1): 211. **[Crossref]**

23. Mohammad-Rahimi H, Khoury ZH, Alamdari MI, et al. Performance of AI chatbots on controversial topics in oral medicine, pathology, and radiology. Oral Surg Oral Med Oral Pathol Oral Radiol. 2024; 137(5): 508-14. **[Crossref]**

24. Özcivelek T, Özcan B. Comparative evaluation of responses from DeepSeek-R1, ChatGPT-o1, ChatGPT-4, and dental GPT chatbots to patient inquiries about dental and maxillofacial prostheses. BMC Oral Health. 2025; 25(1): 871. **[Crossref]**

25. Yilmaz BE, Gokkurt Yilmaz BN, Ozbey F. Artificial intelligence performance in answering multiple-choice oral pathology questions: a comparative analysis. BMC Oral Health. 2025; 25(1): 573. **[Crossref]**

26. Çi Ti R M. ChatGPT and oral cancer: a study on informational reliability. BMC Oral Health. 2025; 25(1): 86. **[Crossref]**

27. Singal A, Goyal S. Comparative evaluation of AI platforms "Google Gemini 2.5 Flash, Google Gemini 2.0 Flash, DeepSeek V3 and ChatGPT 4o" in solving multiple-choice questions from different subtopics of anatomy. Surg Radiol Anat. 2025; 47(1): 193. **[Crossref]**

28. Sahin S, Erkmen B, Duymaz YK, Bayram F, Tekin AM, Topsakal V. Evaluating ChatGPT-4's performance as a digital health advisor for otosclerosis surgery. Front Surg. 2024; 11: 1373843. **[Crossref]**

29. Sezer B, Aydoğdu T. Performance of advanced artificial intelligence models in traumatic dental injuries in primary dentition: a comparative evaluation of ChatGPT-4 Omni, DeepSeek, Gemini Advanced, and Claude 3.7 in terms of accuracy, completeness, response time, and readability. Appl Sci. 2025; 15(14): 7778. **[Crossref]**

30. Taşyürek M, Adıgüzel Ö, Gündoğan M, Goncharuk-Khomyn M, Ortaç H. Comparative evaluation of the responses from ChatGPT-5, Gemini 2.5 Flash, and DeepSeek-V3.1 chatbots to patient inquiries about endodontic treatment in terms of accuracy, understandability, and readability. Int Dent Res. 2025; 15(3). **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):203-208
DOI: 10.54307/NWMJ.2026.233

RESEARCH ARTICLE

The relationship between vitamin D supplementation and symptoms, sleep, depression, and fatigue in restless legs syndrome

Orkun Zorsu[1], **Ahmet Yabalak**[2], **Halime Şahan**[2]

[1]Department of Neurology, Serdivan State Hospital, Sakarya, Türkiye
[2]Department of Neurology, Faculty of Medicine, Düzce University, Düzce, Türkiye

Cite as: Zorsu O, Yabalak A, Şahan H. The relationship between vitamin D supplementation and symptoms, sleep, depression, and fatigue in restless legs syndrome. Northwestern Med J. 2026;6(3):203-208.

ABSTRACT

Background: Restless Legs Syndrome (RLS) is a neurological disorder primarily linked to dopaminergic dysfunction. Since vitamin D plays a role in modulating the dopaminergic system, this study aimed to evaluate the impact of vitamin D supplementation on symptom severity, sleep quality, depression, and fatigue in patients with idiopathic RLS and vitamin D deficiency.

Methods: This study retrospectively evaluated patients with idiopathic RLS and vitamin D deficiency who presented to our clinic between April 2022 and March 2023. Participants received weekly oral supplementation of 50,000 IU vitamin D and were assessed at baseline and at 3 months using the International Restless Legs Syndrome Study Group (IRLS) Symptom Scale, Fatigue Severity Scale (FSS), Beck Depression Inventory (BDI), and Pittsburgh Sleep Quality Index (PSQI).

Results: Twenty-five patients (10 females and 15 males) meeting the eligibility criteria were enrolled. Following vitamin D replacement, mean serum 25-hydroxyvitamin D [25(OH)D] levels increased significantly from 16.91 ± 8.09 ng/mL to 38.08 ± 9.58 ng/mL ($p< 0.001$). A statistically significant reduction was observed in IRLS symptom severity scores (15.36 ± 6.75 vs. 14.28 ± 6.22, $p = 0.019$). However, no statistically significant changes were found in FSS, PSQI, and BDI scores ($p = 0.058$, 0.242, and 0.846, respectively).

Conclusion: Vitamin D supplementation appears to reduce symptom severity in patients with idiopathic RLS and vitamin D deficiency; however, its impact on sleep quality, depression, and fatigue warrants further investigation. Large-scale, randomized controlled trials are necessary to confirm these preliminary findings and define the role of vitamin D in RLS management.

Keywords: depression, fatigue, restless legs syndrome, sleep quality, vitamin D

Corresponding author: Orkun Zorsu **E-mail:** dr.orkunzorsu@gmail.com
Received: 06.01.2026 **Accepted:** 16.06.2026 **Published:** 31.07.2026

INTRODUCTION

Restless legs syndrome (RLS), also known as Willis-Ekbom disease, is a neurological disorder characterized by an urge to move the legs, typically accompanied by uncomfortable sensations (1). These sensations usually occur during rest, where movement is limited, especially in sitting or lying positions, and may worsen in the evening or night hours (2). Most patients report experiencing this discomfort deep within their limbs rather than superficially. Although symptoms typically present bilaterally, they can be asymmetrical or, less frequently, unilateral (3). In patients with RLS, the exacerbation of symptoms upon lying down and the subsequent urge to move the legs significantly impair sleep duration, daily quality of life, and occupational performance (3,4).

The prevalence of RLS in the general population ranges from 1% to 13%, with higher rates observed in women and older populations (5). The condition is classified into primary and secondary forms. Primary RLS is largely considered a genetic disorder, with a positive family history reported in 50% to 70% of cases (6). Conversely, secondary RLS is associated with underlying conditions such as iron deficiency, renal failure, anemia, diabetes mellitus, and pregnancy (7).

The underlying pathophysiological mechanism of RLS involves a decrease in striatal D2 receptor binding potential and circadian fluctuations in extracellular dopamine levels, which lead to sensorimotor disinhibition in susceptible individuals. Rather than a structural loss of dopaminergic neurons, this dysfunction represents a modulation error (2,5). Emerging research suggests that vitamin D plays a critical role in the central dopaminergic system, which is central to the pathophysiology of RLS. Specifically, vitamin D has been shown to exert regulatory effects on the nigrostriatal pathway, increasing dopamine levels and modulating dopaminergic transmission (8). While several studies have demonstrated that low serum vitamin D levels correlate with increased symptom severity and impaired sleep quality, the literature remains inconsistent. Some cross-sectional studies have reported that although RLS prevalence may be higher in vitamin D-deficient cohorts, symptom scores do not differ significantly from those with normal vitamin D levels (9-11).

Given these conflicting findings in the literature, further investigation into the therapeutic impact of vitamin D supplementation is required. Therefore, this study aimed to evaluate the effects of vitamin D replacement therapy on disease symptoms, sleep quality, depression, and fatigue in patients diagnosed with idiopathic RLS and vitamin D deficiency.

METHODS

Study design and population

This study was conducted following approval from the Düzce University Clinical Research Ethics Committee (Decision No: 2023/51). A retrospective review was performed on the medical records of patients who visited the Neurology Outpatient Clinic at Düzce University Hospital between April 2022 and March 2023.

In the movement disorders department of our clinic, vitamin D levels are routinely checked alongside other biochemical parameters. Additionally, patient follow-up routinely includes the assessment of IRLS symptom scale (3), Beck Depression Inventory (BDI) (12), Fatigue Severity Scale (FSS) (13), and Pittsburgh Sleep Quality Index (PSQI) (14) scores. In patients presenting with low vitamin D levels, these parameters are re-evaluated three months after initiating replacement therapy.

This study retrospectively included patients with baseline vitamin D deficiency who started 50,000 IU/week replacement therapy, completed a 3-month follow-up showing normalization of vitamin D levels, and underwent complete BDI, IRLS, FSS, and PSQI evaluations.

Inclusion criteria were: diagnosis of RLS according to the 2014 International Restless Legs Syndrome Study Group (IRLS) criteria (3), receiving a stable dose of RLS medication for at least three months, laboratory-confirmed vitamin D deficiency (serum 25(OH)D levels <30 ng/mL), age of 18 years or older, literacy, and provision of written informed consent.

Patients with renal failure, iron deficiency, liver dysfunction, osteoporosis, secondary RLS due to pregnancy, or polyneuropathy were excluded from the study.

Statistical analysis

Statistical analyses were performed using IBM SPSS version 21 software. The Shapiro-Wilk test was applied to evaluate the normality of data distribution. For normally distributed continuous variables, the Paired Sample T-test was used to compare consecutive measurements between dependent groups. For non-normally distributed data, the Wilcoxon signed-rank test was utilized. Statistical significance was defined as $p < 0.05$.

RESULTS

A total of 25 patients (10 females, 15 males) with a mean age of 45.96 ± 15.41 years were included in the study. Baseline serum 25(OH)D levels ranged from 3.48 ng/mL to 29.48 ng/mL, with a mean baseline concentration of 16.91 ± 8.09 ng/mL. Following three months of vitamin D replacement therapy, post-treatment levels ranged from 30.00 to 61.24 ng/mL, yielding a mean concentration of 38.08 ± 9.58 ng/mL (Figure 1).

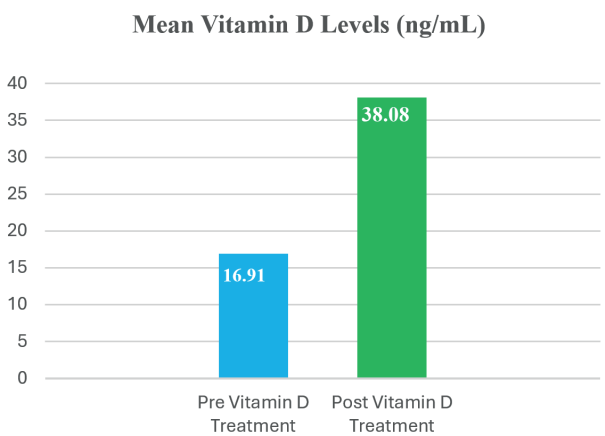


Figure 1. Vitamin D Levels before and after treatment.

RLS symptom severity

At baseline, the mean IRLS Severity Scale score was 15.36 ± 6.75. After three months of vitamin D replacement therapy, the mean IRLS score decreased to 14.28 ± 6.22. Post-replacement IRLS scores were statistically significantly lower than baseline values ($p = 0.019$). This suggests that vitamin D supplementation may reduce clinical symptoms in restless legs syndrome patients with concomitant vitamin D deficiency, although the reduction in symptom scores is modest.

Fatigue

The mean baseline FSS score was 19.32 ± 12.34. After three months of vitamin D replacement therapy, the mean FSS score was 18.24 ± 11.32. Comparing baseline and 3-month values, a slight numerical decrease in FSS scores was observed; however, this change did not reach statistical significance ($p = 0.058$).

Depression and sleep quality

The mean baseline BDI score was 6.80 ± 5.64. Following three months of vitamin D replacement, the mean BDI score was 6.84 ± 5.91, representing no statistically significant change ($p = 0.846$).

The mean baseline PSQI score was 11.76 ± 1.98. The mean post-treatment PSQI score was 11.52 ± 2.30, showing no statistically significant difference between baseline and 3-month follow-up values ($p = 0.242$). A comprehensive comparison of clinical scores (depression, symptom severity, sleep quality, and fatigue) at baseline and at three months is summarized in Table 1.

DISCUSSION

This study demonstrated that three months of vitamin D replacement therapy may have a symptom-reducing effect in patients with idiopathic restless legs syndrome (RLS) and concomitant vitamin D deficiency. However, no statistically significant improvements were observed in sleep quality, depression, or fatigue during the same period.

Table 1. Comparison of clinical assessment scores at baseline and post-treatment (3rd month)

Assessment Scale	Baseline (Mean ± SD)	3rd Month (Mean ± SD)	p-value
IRLS (Symptom Severity)	15.36 ± 6.75	14.28 ± 6.22	0.019*
FSS (Fatigue)	19.32 ± 12.34	18.24 ± 11.32	0.058
PSQI (Sleep Quality)	11.76 ± 1.98	11.52 ± 2.30	0.242
BDI (Depression)	6.80 ± 5.64	6.84 ± 5.91	0.846

*Statistically significant at $p < 0.05$. Abbreviations: BDI: Beck Depression Inventory; IRLS: International Restless Legs Syndrome Study Group Symptom Scale; PSQI: Pittsburgh Sleep Quality Index; FSS: Fatigue Severity Scale.

Although the exact pathophysiology of RLS remains partially understood, dopaminergic dysfunction is widely accepted as the primary mechanism. Vitamin D is known to exert neuroprotective effects on dopaminergic neurons in the nigrostriatal pathway and plays a vital role in modulating dopamine synthesis (8,15,16). The reduction in symptom severity observed in our cohort may be attributed to the supportive role of vitamin D in maintaining the functional integrity of the dopaminergic system.

In a study conducted by Atalar involving 152 patients, 25(OH)D levels were low in 89 patients and normal in 63 (9). RLS symptom scores were significantly higher in the vitamin D-deficient cohort. Similar to our findings, their study demonstrated that RLS severity is heightened in patients with 25(OH)D deficiency. Furthermore, Atalar reported significantly higher PSQI scores in patients with low 25(OH)D levels (9). In contrast, our study did not find a significant change in PSQI scores. In a study by Wali et al., which included 78 patients with primary RLS and 123 controls, vitamin D deficiency was detected in 75.6% (n = 59) of RLS patients compared to 42.3% (n = 52) of control subjects. The risk of developing RLS was significantly higher in individuals with vitamin D levels < 50 nmol/L (< 20 ng/mL) than in those with levels ≥ 50 nmol/L. RLS symptom severity was also greater in vitamin D-deficient patients than in those with normal levels (11). In a cross-sectional study of 102 cases by Cakir et al., the prevalence of RLS was significantly higher in patients with vitamin D deficiency; however, contrary to our findings, no significant differences in symptom severity scores were found when comparing RLS patients with low and normal vitamin D levels (17).

In a study by Liu et al., serum vitamin D levels in 57 patients with primary RLS were significantly lower than those of healthy controls. Additionally, a negative correlation was observed between serum vitamin D levels and IRLS scores, with symptom severity being significantly higher in patients with insufficient vitamin D levels (18). Similarly, a prospective study by Tutuncu and Tutuncu reported that vitamin D replacement led to significant improvements in both total IRLS scores and symptom severity sub-scores (19). Nevertheless, the literature contains conflicting findings. A randomized controlled trial by Wali et al. indicated no significant difference in RLS symptom severity between the group receiving vitamin D supplementation and the placebo cohort (10). These discrepancies suggest that clinical response to vitamin D replacement may depend on the baseline severity of the deficiency or specific patient cohort characteristics.

Regarding secondary outcomes, the lack of significant changes in sleep quality (PSQI) and depression (BDI) scores in our study contrasts with some existing reports. While Liu et al. suggested that low vitamin D levels correlate with poor sleep quality and depressive symptoms (18), our data did not show a statistically significant shift. This lack of change may be explained by the stable baseline treatment of our patients with pramipexole, which potentially masked the additional therapeutic benefits of vitamin D on sleep and mood. Additionally, psychological parameters and sleep architecture often require follow-up periods longer than three months to show measurable clinical recovery.

Although our study demonstrated a downward numerical trend in fatigue scores, this reduction did not achieve statistical significance. Future investigations

with larger sample sizes or extended intervention periods may demonstrate a significant impact of vitamin D replacement on fatigue, potentially offering a valuable adjunctive therapeutic avenue.

Despite the importance of our findings, several limitations must be acknowledged. First, the study was conducted with a relatively small sample size at a single center, which limits the generalizability of the results. Second, the retrospective design and the absence of a control group make it difficult to rule out placebo effects or external variables. Finally, the short follow-up duration prevents assessment of the long-term sustainability of the observed improvements. To confirm these findings and establish definitive causality, prospective, multicenter randomized controlled trials are warranted.

In conclusion, while our study indicates that vitamin D supplementation may improve restless legs syndrome symptoms, the small cohort and lack of a comparative control arm preclude broad clinical application. To validate these preliminary results and optimize therapeutic protocols, more robust, large-scale randomized controlled trials are necessary.

Ethical approval

This study has been approved by the Düzce University Clinical Research Ethics Committee (approval date 20/03/2023, number 2023/51). Written informed consent was obtained from the participants.

Author contribution

Surgical and Medical Practices: OZ, AY, HŞ; Concept: OZ, AY; Design: OZ, AY; Data Collection or Processing: OZ, AY, HŞ; Analysis or Interpretation: OZ, AY; Literature Search: OZ, AY; Writing: OZ, AY, HŞ. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Silber MH, Buchfuhrer MJ, Earley CJ, et al. The management of restless legs syndrome: an updated algorithm. Mayo Clinic Proceedings. 2021; 96(7): 1921-37. **[Crossref]**
2. Wijemanne S, Ondo W. Restless Legs Syndrome: clinicasl features, diagnosis and a practical approach to management. Pract Neurol. 2017; 17(6): 444-52. **[Crossref]**
3. Allen RP, Picchietti DL, Garcia-Borreguero D, et al. Restless legs syndrome/Willis-Ekbom disease diagnostic criteria: updated International Restless Legs Syndrome Study Group (IRLSSG) consensus criteria-history, rationale, description, and significance. Sleep Med. 2014; 15(8): 860-73. **[Crossref]**
4. Ferri R, Koo BB, Picchietti DL, Fulda S. Periodic leg movements during sleep: phenotype, neurophysiology, and clinical significance. Sleep Med. 2017; 31: 29-38. **[Crossref]**
5. Wang S, Zheng X, Huang J, et al. Restless legs syndrome and cognitive function among adults: a systematic review and meta-analysis. J Neurol. 2023; 270(3): 1361-70. **[Crossref]**
6. Sayin S, Atilla FD. Demir eksikliğinin gözden kaçan klinik prezantasyonu; huzursuz bacak sendromu. Ankara Medical Journal. 2019; 19(3): 694-7.
7. Filiz MB, Çakir T. Restless legs syndrome with current diagnostic criteria. Turk J Osteoporos. 2015; 21(2): 87-95 **[Crossref]**
8. Mansourian M, Rafie N, Khorvash F, Hadi A, Arab A. Are serum vitamin D, calcium and phosphorous associated with restless leg syndrome? A systematic review and meta-analysis. Sleep Med. 2020; 75: 326-34. **[Crossref]**
9. Atalar AÇ. The relationship between 25 (OH) vitamin D level and the severity of Disease and Sleep Quality in Restless Legs Syndrome. Turkish Journal of Neurology. 2019; 25(2): 087-91. **[Crossref]**
10. Wali SO, Abaalkhail B, Alhejaili F, Pandi-Perumal SR. Efficacy of vitamin D replacement therapy in restless legs syndrome: a randomized control trial. Sleep Breath. 2019; 23(2): 595-601. **[Crossref]**
11. Wali S, Alsafadi S, Abaalkhail B, et al. The association between vitamin d level and restless legs syndrome: a population-based case-control study. J Clin Sleep Med. 2018; 14(4): 557-64. **[Crossref]**
12. Beck At, Ward Ch, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry. 1961; 4: 561-71. **[Crossref]**
13. Armutlu K, Korkmaz NC, Keser I, et al. The validity and reliability of the Fatigue Severity Scale in Turkish multiple sclerosis patients. Int J Rehabil Res. 2007; 30(1): 81-5. **[Crossref]**
14. Ağargün MY, Kara H, Anlar Ö. Pittsburgh uyku kalitesi indeksinin gecerligi ve guvenirligi. Turk Psikiyatri Dergisi. 1996; 7(2): 107-15.
15. Allen RP. Controversies and challenges in defining the etiology and pathophysiology of restless legs syndrome. Am J Med. 2007; 120(1 Suppl 1): S13-21. **[Crossref]**

16. Prono F, Bernardi K, Ferri R, Bruni O. The role of vitamin D in sleep disorders of children and adolescents: a systematic review. Int J Mol Sci. 2022; 23(3): 1430. **[Crossref]**

17. Çakır T, Doğan G, Subaşı V, et al. An evaluation of sleep quality and the prevalence of restless leg syndrome in vitamin D deficiency. Acta Neurol Belg. 2015; 115(4): 623-7. **[Crossref]**

18. Liu HM, Chu M, Liu CF, Zhang T, Gu P. Analysis of serum vitamin D level and related factors in patients with restless legs syndrome. Front Neurol. 2021; 12: 782565. **[Crossref]**

19. Tutuncu M, Tutuncu M. The effect of vitamin D on restless legs syndrome: prospective self-controlled case study. Sleep Breath. 2020; 24(3): 1101-6. **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):209-217
DOI: 10.54307/NWMJ.2026.235

RESEARCH ARTICLE

Reproductive and obstetric outcomes after complete and partial hydatidiform mole: A single-center retrospective cohort study

Orhan Ay[1], **Fethiye Şahin**[1], **Pelin Bahçeci**[1], **Deniz Taşkıran**[2], **Ali Acar**[1]

[1]Department of Obstetrics and Gynecology, Faculty of Medicine, Necmettin Erbakan University, Konya, Türkiye
[2]Department of Obstetrics and Gynecology, Faculty of Medicine, Giresun University, Giresun, Türkiye

Cite as: Ay O, Şahin F, Bahçeci P, Taşkıran D, Acar A. Reproductive and obstetric outcomes after complete and partial hydatidiform mole: A single-center retrospective cohort study. Northwestern Med J. 2026;6(3):209-217.

ABSTRACT

Aim: To evaluate the reproductive and obstetric outcomes after complete hydatidiform mole (CHM) and partial hydatidiform mole (PHM) in a tertiary referral center.

Materials and Methods: This retrospective cohort included 90 women with histologically confirmed molar pregnancy (65 CHM and 25 PHM) who were managed between 2015 and 2024. Cases with ectopic localization, uncertain diagnosis, missing essential data, or insufficient follow-up were excluded. Women with uterine preservation, no permanent contraception, reproductive age, and adequate follow-up were defined as the subgroup with potential for subsequent pregnancy and were included in the reproductive outcome analysis. Postmolar gestational trophoblastic neoplasia (GTN), occurrence of at least one subsequent pregnancy, outcomes of the first postmolar pregnancy, and obstetric outcomes of the first live birth were compared between CHM and PHM.

Results: Among women with potential for subsequent pregnancy (43 CHM and 14 PHM), at least one subsequent pregnancy occurred in 22 (51.2%) and 8 (57.1%) women, respectively (p=0.697). When only the first subsequent pregnancy was considered, live birth, miscarriage, and recurrent molar pregnancy rates did not differ significantly between CHM and PHM groups (live birth: 63.6% vs 87.5%, p=0.374; miscarriage: 18.2% vs 12.5%, p=1.000; recurrent mole: 18.2% vs 0%, p=0.550). Obstetric outcomes of the first subsequent live birth were also comparable between groups, including gestational age at delivery (p=0.336), birthweight (p=0.924), Apgar scores (all p>0.05), and mode of delivery (p=1.000); all deliveries occurred at term. Fetal growth restriction, preeclampsia, uterine atony, and oligohydramnios did not differ significantly between groups (all p>0.05). No cases of preterm premature rupture of membranes (PPROM), premature rupture of membranes (PROM), placenta previa, placenta accreta spectrum, or placental abruption were observed.

Corresponding author: Orhan Ay **E-mail:** orhan.ay.5561@gmail.com
Received: 24.01.2026 **Accepted:** 23.06.2026 **Published:** 31.07.2026

Conclusion: In this single-center cohort, fertility potential and obstetric outcomes appeared generally favorable after both CHM and PHM. However, subgroup comparisons should be interpreted cautiously, as the limited sample size, particularly in the PHM group, may have reduced the statistical power to detect meaningful between-group differences.

Keywords: gestational trophoblastic disease, hydatidiform mole, postmolar pregnancy, pregnancy outcome, reproductive outcome

INTRODUCTION

Gestational trophoblastic disease (GTD) encompasses a spectrum of disorders arising from placental trophoblastic tissue, ranging from premalignant lesions to invasive neoplasms. Hydatidiform mole is the most common form within this spectrum and is classified into two main subtypes: complete hydatidiform mole (CHM) and partial hydatidiform mole (PHM) (1,2). CHM is most often androgenetic and diploid, whereas PHM is typically diandric and triploid. These differences determine the risk of malignant transformation and guide follow-up strategies. Although hydatidiform mole has a relatively low incidence of 1–3 per 1,000 pregnancies in high-income countries, it remains important as a cause of early pregnancy loss and as a treatable form of gestational trophoblastic neoplasia (GTN) (1,3).

With advances in ultrasonography, sensitive β-human chorionic gonadotropin (β-hCG) assays and expert pathological evaluation, molar pregnancies are now recognized earlier in gestation and with fewer complications. Current guidelines, taking into account the difference in GTN risk between CHM and PHM, recommend subtype-specific durations of hCG surveillance and particularly support shortening the follow-up period after PHM (2). With treatment protocols based on the International Federation of Gynaecology and Obstetrics (FIGO) risk scoring system, cure rates in postmolar GTN have exceeded 95%, and preservation of reproductive potential, in addition to oncologic control, has therefore become a primary therapeutic goal (1,2).

Despite this success, many patients still seek clear information regarding the likelihood of conceiving again after a molar pregnancy, the safe timing of a new pregnancy, the impact of chemotherapy on fertility, the risk of recurrent mole, and potential obstetric complications. Existing studies, including low-risk GTN cases treated with single-agent chemotherapy in most series after conservative management, indicate that successful pregnancy and live birth rates are largely preserved, and that the rates of miscarriage, preterm birth, hypertensive disorders, and congenital anomalies are generally comparable to those in the general population (4-6). Although the risk of recurrent mole is increased after a history of CHM or PHM, it is reported to be around 1% in most series, and the majority of subsequent pregnancies result in healthy outcomes (5). However, many studies are based on selected referral-center populations, do not provide detailed analyses by molar subtype, or are unable to distinguish pregnancy desire and contraceptive status among women who do not conceive. This may cause the true fertility potential to appear lower than it actually is and may contribute to uncertainty in clinical counseling (5-7).

In this study, we aimed to evaluate the postmolar course and subsequent reproductive and obstetric outcomes following complete and partial molar pregnancies managed in a tertiary referral center. In particular, among women with uterine preservation and potential for subsequent pregnancy after postmolar surveillance, we sought to characterize subsequent pregnancy rates and live birth rates. Obstetric outcomes were further assessed among women who achieved a subsequent live birth. In addition, we examined first subsequent pregnancy outcomes, including miscarriage, recurrent molar pregnancy, preterm birth, preeclampsia, fetal growth restriction, and other major obstetric complications, in relation to molar subtype. In this way, we aimed to provide up-to-date, clinically applicable data that can be directly used in counseling after a molar pregnancy.

MATERIALS AND METHODS

This retrospective cohort study included women who were followed after a diagnosis of molar pregnancy at a tertiary referral center between January 2015 and December 2024, and whose pathology reports confirmed complete or partial hydatidiform mole. Cases with ectopic localization, uncertain diagnosis or missing key clinical data, as well as those with insufficient follow-up duration, were excluded. From the medical records, data on maternal age, obstetric history, pathology reports regarding mole type, treatment modality, postmolar follow-up, development of GTN, and pregnancy outcomes after the molar pregnancy were collected.

Definitions and subgroup classification

Molar pregnancy subtypes

The diagnoses of complete and partial hydatidiform mole were based on final histopathological reports prepared according to standard histopathological criteria. Cases explicitly reported as complete hydatidiform mole or partial hydatidiform mole were assigned to the corresponding subgroup. Cases with uncertain, equivocal, or non-classifiable pathological diagnoses were excluded from the analysis.

Subgroup of women with potential for subsequent pregnancy

For the evaluation of postmolar reproductive outcomes, we defined a subgroup of women with potential for subsequent pregnancy within the overall cohort. This subgroup was intended to identify women with a realistic biological and clinical possibility of a subsequent pregnancy during follow-up. The inclusion criteria were preservation of the uterus (no hysterectomy performed), no definitive contraception (e.g. tubal ligation) initiated before completion of hCG surveillance, being of reproductive age during the postmolar period, and adequate follow-up. In women meeting these criteria, pregnancies occurring after the molar gestation and their obstetric outcomes were analyzed. Women were considered to have adequate follow-up if they had documented completion of postmolar hCG surveillance and available follow-up records regarding subsequent reproductive status, or if a subsequent pregnancy was documented during follow-up. However, because of the retrospective design, voluntary postponement of pregnancy could not be reliably distinguished from reduced fertility or unsuccessful attempts to conceive based on the medical records.

Postmolar GTN

Postmolar GTN was diagnosed on the basis of clinical and laboratory findings according to internationally accepted criteria. In the presence of a plateau or rise in serial hCG measurements after a molar pregnancy, the detection of new lesions, or evidence of distant metastasis, a diagnosis of GTN was considered and these cases were classified within the GTN subgroup.

Postmolar reproductive and obstetric outcome assessment

In the subgroup of women with potential for subsequent pregnancy, pregnancies achieved after the molar gestation were evaluated. For statistical analyses, only the first pregnancy occurring after the molar pregnancy was taken into account for each woman (Figure 1). For this first subsequent pregnancy, live birth, miscarriage, and recurrent molar pregnancy were assessed. Obstetric outcomes were evaluated among women who achieved a subsequent live birth and included gestational age at delivery, birthweight, Apgar scores, mode of delivery, fetal growth restriction, preeclampsia, uterine atony, oligohydramnios, and major obstetric complications. The occurrence of preterm premature rupture of membranes (PPROM), premature rupture of membranes (PROM), placenta previa, placenta accreta spectrum, and placental abruption was also recorded.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics 22.0 (IBM Corp., Armonk, NY, USA). Continuous variables were presented as mean ± standard deviation or median (interquartile range, IQR), according to their distributional characteristics. For comparisons between two independent groups, parametric tests (independent-samples t-test) were used when appropriate, and non-parametric tests (Mann–Whitney U) were applied otherwise.

Categorical variables were expressed as numbers (percentages) and compared using the Pearson chi-square test or Fisher's exact test when expected cell counts were insufficient. A p-value < 0.05 was considered statistically significant for all analyses. No formal a priori power analysis was performed, as this was a retrospective study and the sample size was determined by the number of eligible cases available during the study period.

Ethical approval

The study was approved by the Necmettin Erbakan University Ethics Committee for Non-Drug and Non-Medical Device Research at its meeting held on January 24, 2025 (Approval No: 2025/5472). Individual informed consent was waived because of the retrospective design and the use of anonymized data. The study was conducted in accordance with the principles of the Declaration of Helsinki.

RESULTS

In our study, 65 of the 90 cases (72.2%) were complete moles and 25 (27.8%) were partial moles. Among women with potential for subsequent pregnancy, 43 had a history of complete mole and 14 had a history of partial mole. At least one subsequent pregnancy occurred in 22 (51.2%) and 8 (57.1%) women, respectively. Reproductive outcomes following molar pregnancy in these patients are summarized in Figure 1. There were no significant differences between the complete and partial mole groups in terms of age, gravidity, parity, or previous miscarriage history. Serum β-hCG levels at diagnosis were markedly higher in complete mole cases (median 212,330 vs 85,000 mIU/mL, p<0.001) (Table 1). None of the patients had a prior history of molar pregnancy.

The primary treatment approach and postmolar course are presented in Table 2. Uterine evacuation was the main treatment modality in both groups and was performed in 59 (90.8%) complete mole cases and 23 (92.0%) partial mole cases, while primary hysterectomy was performed in 6 (9.2%) and 2 (8.0%) cases, respectively, with no significant difference between groups (p=1.000). Postmolar invasive mole/GTN developed in one patient with complete

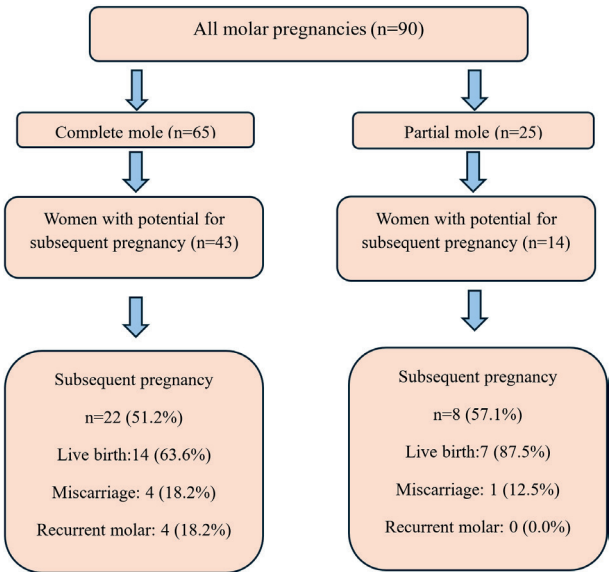


Figure 1. Flow diagram of the study cohort and subsequent pregnancy outcomes after molar pregnancy, overall and stratified by mole type.

mole (1.5%) and in no patients with partial mole (p=1.000). Although the duration of contraception after evacuation appeared longer in the complete mole group than in the partial mole group (4 [1–8] months vs 2 [1–7] months), this difference was not statistically significant (p=0.129).

Reproductive outcomes after molar pregnancy are presented in Table 3. The subgroup of women with potential for subsequent pregnancy comprised 43 complete mole and 14 partial mole cases. The proportion of women achieving at least one subsequent pregnancy was similar after complete and partial mole (51.2% and 57.1%, respectively; p=0.697). When only the first subsequent pregnancy was taken into account, miscarriage, live birth, and recurrent molar pregnancy rates did not differ significantly between the prior complete and partial mole groups (miscarriage: 18.2% vs 12.5%, p=1.000; live birth: 63.6% vs 87.5%, p=0.374; recurrent molar pregnancy: 18.2% vs 0.0%, p=0.550). Recurrent molar pregnancy was observed only in women with a previous complete mole.

Obstetric outcomes of the first subsequent live birth are shown in Table 4. There were 14 live births after complete mole and 7 live births after partial mole.

Table 1. Baseline characteristics according to molar subtype

Variable	Complete mole (n=65)	Partial mole (n=25)	p-value
Age (years), median (IQR)	28 (23-35)	25 (21-32)	0.417
Serum β-hCG at diagnosis (mIU/mL)	212,330 (86,352-387,124)	85,000 (28,700-138,067)	**<0.001**
Gravidity, median (IQR)	1 (1-3)	2 (1-3)	0.344
Parity, median (IQR)	0 (0-1)	0 (0-1)	0.741
Number of previous miscarriages, median (IQR)	0 (0-0)	0 (0-0)	0.573
Previous normal vaginal delivery, median (IQR)	0 (0-0)	0 (0-0)	0.870
Previous cesarean section, median (IQR)	0 (0-0)	0 (0-0)	0.506

Data are presented as median (IQR). P-values were calculated using the Mann–Whitney U test; $p < 0.05$ was considered statistically significant.
β-hCG: beta-human chorionic gonadotropin.

Table 2. Primary treatment modality, postmolar course, and contraception duration according to molar subtype

Variable	Complete mole (n=65)	Partial mole (n=25)	p-value
Primary treatment modality			
Uterine evacuation only, n (%)	59 (90.8)	23 (92.0)	1.000
Primary hysterectomy, n (%)	6 (9.2)	2 (8.0)	
Postmolar invasive mole / GTN, n (%)	1 (1.5)	0 (0.0)	1.000
Contraception duration after molar evacuation (months), median (IQR)	4 (1-8)	2 (1-7)	0.129

Categorical variables are presented as n (%) and compared using Fisher's exact test; continuous variables are presented as median (IQR) and compared using the Mann–Whitney U test. p-value < 0.05 was considered statistically significant.

Table 3. Reproductive outcomes after molar pregnancy according to prior molar subtype

Outcome	Prior complete mole	Prior partial mole	p-value[a]
Women with potential for subsequent pregnancy[b]	n=43	n=14	
Achieved ≥1 subsequent pregnancy, n (%)	22 (51.2)	8 (57.1)	0.697
No subsequent pregnancy[c], n (%)	21 (48.8)	6 (42.9)	0.697
Outcome of the first subsequent pregnancy[d]	(n= 22)	(n= 8)	
Miscarriage, n (%)	4 (18.2)	1 (12.5)	1.000
Live birth, n (%)	14 (63.6)	7 (87.5)	0.374
Recurrent molar pregnancy, n (%)	4 (18.2)	0 (0.0)	0.550

[a]P-values were calculated using the chi-square or Fisher's exact test, as appropriate; $p < 0.05$ was considered statistically significant.

[b] Women with uterine preservation, no permanent contraception, reproductive age, and adequate follow-up.

[c]Includes women remaining non-pregnant during follow-up (does not distinguish voluntary vs involuntary).

[d]Only the first subsequent pregnancy per woman was considered

In both groups, all deliveries occurred at term and no preterm birth was observed. Gestational age at delivery, birthweight, and 1-, 5-, and 10-minute Apgar scores were similar between women with a history of complete and partial mole. There was no significant difference in mode of delivery between the groups (cesarean section: 57.1% vs 57.1%; vaginal delivery: 42.9% vs 42.9%; p=1.000). Fetal growth restriction (7.1% vs 0.0%, p=1.000), preeclampsia (0.0% vs 14.3%, p=0.333), uterine atony (0.0% vs 14.3%, p=0.333), and

Table 4. Obstetric outcomes of the first subsequent live birth according to prior molar subtype

Outcome	Prior complete mole (n=14)	Prior partial mole (n=7)	p-value
Gestational age at delivery (weeks), mean ± SD	38.0 ± 0.68	37.71 ± 0.49	0.336
Birthweight (g), mean ± SD	3,238.57 ± 442.63	3,258.57 ± 452.38	0.924
Mode of delivery, n (%)			1.000
Cesarean section	8 (57.1)	4 (57.1)	
Vaginal delivery	6 (42.9)	3 (42.9)	
1-min Apgar score, mean ± SD	6.93 ± 0.99	6.57 ± 0.98	0.446
5-min Apgar score, mean ± SD	7.93 ± 0.99	7.57 ± 0.98	0.446
10-min Apgar score, mean ± SD	8.86 ± 0.86	8.57 ± 0.98	0.502
Fetal growth restriction, n (%)	1 (7.1)	0 (0.0)	1.000
Preeclampsia, n (%)	0 (0.0)	1 (14.3)	0.333
Uterine atony, n (%)	0 (0.0)	1 (14.3)	0.333
Oligohydramnios, n (%)	3 (21.4)	3 (42.9)	0.354

Continuous variables were compared using the independent-samples t-test; categorical variables were compared using Fisher's exact test. A p-value < 0.05 was considered statistically significant.

oligohydramnios (21.4% vs 42.9%, p=0.354) did not differ significantly between women with a history of complete and partial mole. None of the pregnancies were complicated by PPROM (preterm premature rupture of membranes), PROM (premature rupture of membranes), placental abruption, placenta previa, or PAS (placenta accreta spectrum).

DISCUSSION

In this single-center retrospective cohort, we evaluated the postmolar course and subsequent reproductive and obstetric outcomes following complete and partial molar pregnancies. Although our study is based on a small and selected patient group, it provides practical data that may be useful for clinical counseling, as it reflects real-life experience from a tertiary referral center.

First, the differences in baseline characteristics between complete and partial molar pregnancies support the expected biological distinction. The markedly higher serum β-hCG levels at diagnosis in complete mole cases, with other clinical parameters remaining similar, are consistent with the more pronounced trophoblastic proliferation and higher malignant transformation potential of complete mole (4,8). In large series in the literature, postmolar GTN rates after complete mole have also been reported to be significantly higher than after partial mole. In the centralized surveillance program by Jiao et al., the incidence of postmolar GTN was approximately 12% following complete mole and around 0.4% following partial mole (9). Recent evidence further supports this subtype-specific biological distinction. In a 2025 retrospective cohort from Shanghai including 506 pathologically confirmed hydatidiform moles, 42 women (8.3%) developed postmolar GTN, and both pathological subtype and pre-evacuation lesion diameter were independently associated with malignant progression (10). Notably, all patients who developed postmolar GTN achieved complete response after treatment. These findings are in line with the 2025 update on gestational trophoblastic disease, which continues to emphasize careful risk-adapted surveillance after molar evacuation (2). In our series, the postmolar GTN rate after complete mole was 1.5%, and no cases of GTN were observed after partial mole. The rate observed in the complete mole group appears lower than that reported in several large series. However, this finding was based on a single GTN event and should therefore be interpreted cautiously. The limited sample size, center-specific referral and surveillance patterns, early diagnosis and

evacuation in some cases, and statistical imprecision related to the very small number of events may have contributed to this lower observed rate. Therefore, this finding should not be interpreted as evidence of a lower GTN risk after complete mole.

In the subgroup of women with potential for subsequent pregnancy, the proportions achieving at least one subsequent pregnancy after complete and partial mole were 51.2% and 57.1%, respectively, with no significant difference between the two groups. When only the first subsequent pregnancy was considered, live birth rates were 63.6% after complete mole and 87.5% after partial mole. Although numerically higher in favor of partial mole, this difference did not reach statistical significance due to the sample size. These findings are in line with large cohort and review studies reporting that successful pregnancy and live birth rates in conservatively treated GTD patients are broadly comparable to those in the general population (11,12). In the meta-analysis by Capozzi et al., live birth rates after complete and partial mole were reported as 53.6% and 51.0%, respectively, and no significant difference was demonstrated between the two subtypes (4). Similarly, Joneborg et al. reported that, among women wishing to conceive after hydatidiform mole and across the GTD spectrum including low-risk GTN, pregnancy and live birth rates were preserved, while the prevalence of infertility was comparable to that in the general population (6).

Gestational age at delivery, birthweight, and 1, 5, and 10-minute Apgar scores for the first subsequent live birth did not differ between women with a history of complete versus partial mole, and all subsequent live births occurred at term with favorable neonatal outcomes. In addition, the low frequencies of major complications such as preeclampsia, fetal growth restriction, uterine atony, and oligohydramnios in both groups, together with the absence of severe events such as PPROM, placenta previa, and placental abruption, are broadly consistent with the existing literature suggesting no clear increase in obstetric risk after a molar pregnancy (4,12,13). However, these findings should be interpreted cautiously, as the sample size, particularly in subgroup analyses and rare obstetric outcomes, was limited. Similarly, miscarriage and preterm birth rates were low in our cohort. Nevertheless, given the small number of cases and the very small number of postmolar GTN cases, our results should be viewed as reassuring but preliminary rather than definitive evidence of favorable obstetric prognosis after molar pregnancy.

Recurrent molar pregnancy was observed only in women with a history of complete mole and in their first subsequent pregnancy (4/22; 18.2%). Although this proportion is higher than the recurrence risk reported in most large series, including the approximately 1–1.5% risk described by Eagles et al.(5), it was based on only four events and a small denominator. Therefore, this finding should be interpreted cautiously and should not be considered a precise recurrence risk estimate for the entire cohort. Potential selection effects related to the retrospective tertiary referral-center design and the inclusion of women with available follow-up data may also have influenced the observed case mix. Overall, this descriptive finding supports the importance of early β-hCG monitoring and ultrasound assessment in subsequent pregnancies after a molar pregnancy.

With regard to postmolar surveillance and contraception practice, the longer contraception interval observed after complete mole than after partial mole in our cohort is broadly consistent with current subtype-specific follow-up recommendations (2,14). Contemporary guidance does not uniformly require a universal 1-year delay after every molar pregnancy. Rather, follow-up is adapted according to molar subtype, with shorter surveillance after partial mole and longer monitoring after complete mole. In particular, current recommendations support hCG surveillance for 1 month after the first normal hCG value in partial mole and monthly surveillance for 6 months after the first normal hCG value in complete mole (2,14). In addition, a large retrospective cohort of 7,761 women reported postmolar GTN after hCG normalization in 0% of histologically confirmed partial moles and 0.36% of complete moles, supporting a risk-adapted rather than uniformly prolonged follow-up approach (15). Nevertheless, in our cohort, the observed contraception interval still appears shorter than the recommended surveillance duration in many patients, particularly after complete mole, and this should be taken into account when interpreting subsequent reproductive outcomes. Therefore, our findings should

not be interpreted as evidence against guideline-based surveillance and contraception practices; instead, they should be considered cautiously in light of the limited sample size, the small number of recurrence events, and the retrospective design of the study.

Findings from our complete and partial mole cases are consistent with FIGO guidance and recent reviews. They show that fertility preservation is a realistic goal when postmolar follow-up and treatment are appropriate. In well-treated and regularly monitored women, the chances of conceiving and achieving a healthy live birth are high. Major obstetric complications do not appear to differ markedly from those in the general population. These results support a reassuring approach when counseling young women who plan future pregnancies.

Recent Turkish publications also support the relevance of reporting contemporary single-center experience from Türkiye. In a 2025 Turkish benign GTD cohort, complete mole accounted for 57.4% and partial mole for 42.6% of cases; vaginal bleeding remained the leading presenting symptom (68.9%), and 26.2% of diagnoses were made during follow-up visits, illustrating that case mix and timing of diagnosis remain highly center-dependent (16). Likewise, a 2025 Turkish tertiary-center series reported high complete response rates across standard GTD treatment regimens, reinforcing the generally favorable outcomes achievable with referral-center care (17).

This study has some limitations. It is a retrospective, single-center study with a relatively small sample size. These findings should be interpreted with caution, as the subgroup sizes, particularly in the PHM group and in analyses restricted to subsequent live births, were small, limiting the statistical power to detect potentially meaningful between-group differences. Therefore, nonsignificant findings should not be interpreted as evidence of equivalence between complete and partial mole groups. Because the number of outcome events was limited, particularly for recurrent molar pregnancy, postmolar GTN, and live-birth subgroup analyses, multivariable modeling was considered statistically unreliable and was therefore not performed. In addition, although we identified a subgroup of women with potential for subsequent pregnancy, the retrospective design did not allow us to reliably distinguish voluntary postponement of pregnancy from reduced fertility or unsuccessful attempts to conceive. Moreover, focusing only on the first subsequent pregnancy may not fully capture the overall reproductive and obstetric experience of women with more than one postmolar pregnancy. The number of postmolar GTN cases was also very small. Nevertheless, the inclusion of pathologically confirmed complete and partial mole cases, detailed subtype-specific analyses, the focus on women with potential for subsequent pregnancy, and the comprehensive recording of obstetric outcomes in completed pregnancies represent important strengths of our study. These data may contribute to local clinical counseling and provide supportive real-world evidence for subtype-specific postmolar follow-up discussions.

CONCLUSION

This study suggests that fertility potential and obstetric outcomes appear generally favorable after both complete and partial molar pregnancy. However, given the limited sample size and the small number of outcome events, potential differences between subtypes should be interpreted with caution. These findings may contribute to more balanced and realistic counseling of women with a history of molar pregnancy regarding their future reproductive outcomes.

Ethical approval

This study has been approved by the Ethics Committee for Non-Drug and Non-Medical Device Research of Necmettin Erbakan University (approval date anuary 24, 2025, number 2025/5472). As the study was conducted retrospectively and all data were analyzed after removal of identifying information, the Ethics Committee for Non-Drug and Non-Medical Device Research of Necmettin Erbakan University determined that written informed consent was not required. Accordingly, individual patient consent forms were not obtained.

Author contribution

Surgical and Medical Practices: OA, FŞ, PB, AA; Concept: OA,DT; Design: OA, PB; Data Collection or Processing: FŞ, PB, AA; Analysis or Interpretation: OA,

PB, DT, AA; Literature Search: OA; Writing: OA, FŞ, PB, DT, AA. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Soper JT. Gestational trophoblastic disease: current evaluation and management. Obstet Gynecol. 2021; 137(2): 355-70. [Crossref]
2. Ngan HYS, Seckl MJ, Berkowitz RS, et al. Diagnosis and management of gestational trophoblastic disease: 2025 update. Int J Gynaecol Obstet. 2025; 171 Suppl 1(Suppl 1): 78-86. [Crossref]
3. Seckl MJ, Sebire NJ, Berkowitz RS. Gestational trophoblastic disease. Lancet. 2010; 376(9742): 717-29. [Crossref]
4. Capozzi VA, Butera D, Armano G, et al. Obstetrics outcomes after complete and partial molar pregnancy: review of the literature and meta-analysis. Eur J Obstet Gynecol Reprod Biol. 2021; 259: 18-25. [Crossref]
5. Eagles N, Sebire NJ, Short D, Savage PM, Seckl MJ, Fisher RA. Risk of recurrent molar pregnancies following complete and partial hydatidiform moles. Hum Reprod. 2015; 30(9): 2055-63. [Crossref]
6. Joneborg U, Coopmans L, van Trommel N, Seckl M, Lok CAR. Fertility and pregnancy outcome in gestational trophoblastic disease. Int J Gynecol Cancer. 2021; 31(3): 399-411. [Crossref]
7. Turgal M, Beksac MS. Molar pregnancy and poor obstetrics outcome: retrospective evaluation of six cases. J Med Cases. 2016; 7(6): 239-41. [Crossref]
8. Al-Talib AA. Clinical presentation and treatment outcome of molar pregnancy: ten years experience at a Tertiary Care Hospital in Dammam, Saudi Arabia. J Family Community Med. 2016; 23(3): 161-5. [Crossref]
9. Jiao L, Wang Y, Jiang J, et al. Centralized surveillance of hydatidiform mole: 7-year experience from a regional hospital in China. Int J Gynecol Cancer. 2022; 32(2): 147-52. [Crossref]
10. Zhu T, Zhu G, Ma J, et al. Demographics, risk factors associated with malignant progression, and pregnancy outcomes of hydatidiform mole: a retrospective cohort study in Shanghai, China. Front Oncol. 2025; 15: 1643590. [Crossref]
11. Vargas R, Barroilhet LM, Esselen K, Diver E, Bernstein M, Goldstein DP, et al. Subsequent pregnancy outcomes after complete and partial molar pregnancy, recurrent molar pregnancy, and gestational trophoblastic neoplasia: an update from the New England Trophoblastic Disease Center. J Reprod Med. 2014; 59(5-6): 188-94.
12. Garrett LA, Garner EIO, Feltmate CM, Goldstein DP, Berkowitz RS. Subsequent pregnancy outcomes in patients with molar pregnancy and persistent gestational trophoblastic neoplasia. J Reprod Med. 2008; 53(7): 481-6.
13. Tranoulis A, Georgiou D, Sayasneh A, Tidy J. Gestational trophoblastic neoplasia: a meta-analysis evaluating reproductive and obstetrical outcomes after administration of chemotherapy. Int J Gynecol Cancer. 2019; 29(6): 1021-31. [Crossref]
14. Eiriksson L, Dean E, Sebastianelli A, et al. Guideline No. 408: Management of gestational trophoblastic diseases. J Obstet Gynaecol Can. 2021; 43(1): 91-105.e1. [Crossref]
15. Descargues P, Hajri T, Massardier J, et al. Gestational trophoblastic neoplasia after human chorionic gonadotropin normalization in a retrospective cohort of 7761 patients in France. Am J Obstet Gynecol. 2021; 225(4): 401.e1-9. [Crossref]
16. Turan K, Sargın BA. Retrospective evaluation of patients diagnosed with gestational trophoblastic disease in our clinic. Van Med J. 2025;32(4):227-32. [Crossref]
17. Özberk U, Aktürk Esen S, Seven I, Bal O, Algın E, Bilgin B, et al. Clinicopathological features and treatment outcomes of gestational trophoblastic disease: a retrospective study from a tertiary care center in Türkiye. ACH Med J. 2025;4(2):40-6. [Crossref]

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):218-224
DOI: 10.54307/NWMJ.2026.237

RESEARCH ARTICLE

The relationship between anthropometric measurements and the frequency and severity of attacks in migraine patients in western Türkiye

Nurcan Akbaş Güneş[1], **Canan Akünal**[2], **Çağla Özdemir**[3]

[1]Department of Family Medicine, Faculty of Medicine, Bolu Abant İzzet Baysal University, Bolu, Türkiye
[2]Department of Neurology, Faculty of Medicine, Bolu Abant İzzet Baysal University, Bolu, Türkiye
[3]Kütahya University of Health Sciences, Kütahya, Türkiye

Cite as: Akbaş Güneş N, Akünal C, Özdemir Ç. The relationship between anthropometric measurements and the frequency and severity of attacks in migraine patients in western Türkiye. Northwestern Med J. 2026;6(3):218-224.

ABSTRACT

Objective: Migraine affects approximately 20% of the general population. This study evaluated the effects of anthropometric measurements on migraine. We aimed to investigate the relationship between migraine and various anthropometric indices as superior indicators of regional body fat distribution.

Methods: This study included 190 participants. Height, weight, waist circumference (WC), hip circumference (HC), and neck circumference (NC) were measured. These measurements were used to calculate body mass index (BMI), waist-to-hip ratio (WHR), waist-to-height ratio (WHtR), and a body shape index (ABSI). The relationships between these anthropometric parameters and the Migraine Disability Assessment (MIDAS) scale were evaluated.

Results: BMI, WHR, and HC were significantly associated with migraine status ($p < 0.001$, $p < 0.001$, and $p = 0.040$, respectively). Among migraine patients, attack frequency correlated significantly with BMI, WHtR, WC, NC, and HC, whereas ABSI showed no significant association with attack frequency or severity ($p = 0.746$).

Conclusion: BMI, WHR, and HC are significantly associated with migraine. No correlation was observed between ABSI and migraine attack frequency or severity. Furthermore, a significant relationship was identified between MIDAS scores, HC, and WHtR.

Keywords: migraine, a body shape index, Migraine Disability Assessment, waist-to-hip ratio, body mass index

Corresponding author: Nurcan Akbaş Güneş **E-mail:** drnurak@hotmail.com
Received: 16.02.2026 **Accepted:** 12.05.2026 **Published:** 31.07.2026

INTRODUCTION

Migraine is a neurovascular disorder characterized by recurrent headache attacks and associated symptoms, including photophobia, phonophobia, osmophobia, nausea, vomiting, myalgia, and cutaneous allodynia (1). It affects up to 20% of the general population (2). Although its pathophysiology is complex, it involves interactions within the trigeminovascular system and various peptides, including sex hormones, oxytocin, and calcitonin gene-related peptide (CGRP). While its etiology remains not fully elucidated, it is influenced by genetic, hormonal, and environmental factors. Its prevalence is higher in women than in men, peaking during middle age. Migraine susceptibility and attacks are influenced by environmental factors such as alcohol and tobacco use, caffeine consumption, a sedentary lifestyle, dietary habits, and psychological distress (3,4). Additionally, obesity has been demonstrated to impact migraine frequency and severity (5). Large-scale studies have shown that the prevalence of migraine in women with severe obesity is twice as high as in those with a normal weight (6).

Several studies have shown that obesity affects other pain disorders, including various types of headache (7). Furthermore, migraine and obesity share overlapping pathophysiological pathways, such as endothelial dysfunction, chronic systemic inflammation, and elevated levels of inflammatory mediators (8). A Body Shape Index (ABSI) is calculated using waist circumference, height, and weight to assess an individual's body structure and health risk. Higher ABSI values are associated with greater central obesity, a higher risk of cardiovascular disease, and increased all-cause mortality (9,10). Although body mass index (BMI) is widely used in clinical practice to define obesity, its utility is limited as it does not reflect regional body fat distribution. Therefore, other parameters, including waist circumference (WC), hip circumference (HC), neck circumference (NC), waist-to-hip ratio (WHR), waist-to-height ratio (WHtR), and ABSI, may provide a more comprehensive representation of body fat distribution.

In previous studies, the effect of obesity on migraine was evaluated using BMI only (11,12). In this study, the effect of other anthropometric measurements (WC, HC, NC, WHR, WHtR, and ABSI) as a better indicator of body fat distribution on migraine was evaluated.

METHODS

Study design

Prior to initiating the study, necessary ethical approvals and permissions were obtained (Approval No: 2019/324). The sample size was calculated using the G*Power 3.1.9.2 software. For an independent samples t-test with a medium effect size (Cohen's d = 0.50), a significance level of $\alpha = 0.05$, and 90% power, the calculation indicated that a minimum of 172 participants was required.

The study prospectively enrolled 95 migraine patients and 95 healthy controls aged 18–50 years who presented to the Neurology outpatient clinic and agreed to participate. Migraine was diagnosed according to the criteria of the International Classification of Headache Disorders, 3rd edition (ICHD-3) (13). Participants with other major neurological or psychiatric disorders (e.g., bipolar disorder, schizophrenia) were excluded from the study.

Data collection tools

Anthropometric Measurements: Written informed consent was obtained from all participants prior to enrollment and physical assessment. Height, weight, WC, NC, and HC were measured using standardized protocols. These measurements were used to calculate BMI, WHR, and WHtR. Additionally, A Body Shape Index (ABSI), a novel anthropometric index, was calculated using the standard formula:

BMI = Weight / Height2 (kg/m2)

WHR = Waist Circumference / Hip Circumference

WHtR = Waist Circumference / Height

ABSI = WC ÷ (BMI2/3 × Height1/2)

Migraine Disability Assessment Scale (MIDAS): This scale, which evaluates migraine-associated work productivity and daily activity loss, was developed by Stewart et al., who reported a Cronbach's alpha of 0.87

(14). The Turkish validity and reliability of this scale was established, with a reported Cronbach's alpha of >0.8 (15). The MIDAS score is calculated by summing the total days of activity or work loss due to migraine headaches over the preceding three months. Scores are categorized into four grades of disability: Grade I (little or no disability, scores 0–5), Grade II (mild disability, scores 6–10), Grade III (moderate disability, scores 11–20), and Grade IV (severe disability, scores 21 or more).

Statistical analysis

Statistical analyses were performed using the IBM SPSS Statistics software package. Continuous variables are expressed as mean ± standard deviation, while categorical variables are presented as frequencies and percentages (%). The normality of the distribution of continuous variables was evaluated using standard normality tests and graphical methods. For comparisons between two independent groups, the independent samples t-test was used when normal distribution assumptions were satisfied. The relationships between continuous variables were examined using Pearson correlation analysis. For all analyses, a p-value of less than 0.05 was considered statistically significant.

RESULTS

A total of 190 participants (n = 95 patients; n = 95 healthy controls) were included in this study. The mean age of the migraine patient group was 33.18 ± 8.8 years, while that of the control group was 32.3 ± 1.5 years. Females constituted 75.8% (n = 72) of the migraine patient group. The sociodemographic characteristics of the cohorts are summarized in Table 1.

Significant differences in HC, BMI, and WHR were observed between the migraine group and the controls ($p < 0.001$, $p < 0.001$, and $p = 0.040$, respectively). Specifically, HC and BMI were significantly higher in the migraine group, whereas WHR was significantly lower compared to the control group. A detailed comparison of the anthropometric measurements between the groups is presented in Table 2.

There was no significant association between MIDAS scores and ABSI in the patient group ($p = 0.746$). In terms of sociodemographic factors, a statistically significant relationship was found between migraine attack frequency and marital status, with married patients experiencing a higher frequency of attacks

Table 1. Distribution of participants' sociodemographic characteristics

	Diseased Group	n (%)	Control Group	n (%)
Gender	Male	23 (24.2%)	Male	46 (48.4%)
	Female	72 (75.8%)	Female	49 (51.6%)
Marital Status	Single	37 (38.9%)	Single	79 (83.2%)
	Married	58 (61.1%)	Married	16 (16.8%)
Education Status	None	1 (1.1%)	None	0 (0%)
	Primary	16 (16.8%)	Primary	12 (12.5%)
	High School	49 (51.6%)	High School	47 (49.5%)
	University	29 (30.5%)	University	36 (38%)
Family History	No	43 (45.3%)	No	68 (71.6%)
	Yes	52 (54.7%)	Yes	27 (28.4%)
MIDAS	Between 1-5 days	2 (2.1%)		
	Between 6-10 days	13 (13.7%)		
	Between 11-20 days	27 (28.4%)		
	21 days and more	53 (55.8%)		

* MIDAS: Migraine Disability Assessment

Table 2. Evaluation of anthropometric measurements of the diseased and control group

		n	Mean±SD	p	95% CI [Lower, Upper]
Gender	Diseased	95	1.76±0.43	**<0.001**	[0.10, 0.37]
	Control	95	1.52±0.50		
Height	Diseased	95	163.59±8.95	**<0.001**	[-9.85, -4.86]
	Control	95	170.95±8.45		
Weight	Diseased	95	72.41±11.35	**0.018**	[0.80, 8.54]
	Control	95	67.74±15.38		
WC	Diseased	95	84.36±14.04	0.128	[-0.86, 6.82]
	Control	95	81.38±12.80		
NC	Diseased	95	35.64±2.47	0.101	[-0.14, 1.58]
	Control	95	34.92±3.47		
HC	Diseased	95	103.96±9.70	**<0.001**	[2.89, 7.56]
	Control	95	98.73±6.24		
BMI	Diseased	95	27.32±4.89	**<0.001**	[3.22, 5.61]
	Control	95	22.90±3.29		
ABSI	Diseased	95	0.08±0.09	0.297	[-0.00, 0.02]
	Control	95	0.07±0.01		
WHR	Diseased	95	0.78±0.13	0.040	[-0.07, -0.00]
	Control	95	0.82±0.10		
WHtR	Diseased	95	0.49±0.11	**0.066**	[-0.00, 0.05]
	Control	95	0.47±0.05		

* WC: Waist circumference, NC: Neck Circumference, HC: Hip Circumference, BMI: Body Mass Index, ABSI: A Body Shape Index, WHR: waist/hip ratio, WHtR: waist/height ratio

compared to single patients (p = 0.010). Conversely, there was no statistically significant relationship between family history of migraine and the frequency or severity of attacks (p = 0.053). The correlations between individual anthropometric measurements and migraine attack frequency and severity within the patient group are detailed in Table 3.

DISCUSSION

This study demonstrated that BMI, WHR, and HC are significantly associated with migraine status. Furthermore, BMI was found to correlate with both the frequency and severity of migraine attacks. In addition, positive correlations were observed between the frequency of migraine attacks and WHtR, NC, WC, and HC. Conversely, no correlation was found between ABSI, a relatively novel anthropometric index, and either the frequency or severity of migraine attacks.

Existing literature suggests that obesity increases the risk of migraine (5). Consistently, higher BMI has been associated with increased migraine frequency (11). A large-scale study involving over 3,800 participants reported that the risk of migraine was 81% higher in obese individuals than in their normal-weight peers (16). These findings align with our results. Nonetheless, conflicting findings exist in the literature (17). For instance, in a study of 15,105 individuals, no direct association was found between BMI and the presence of migraine; however, a strong correlation was identified between elevated BMI and migraine attack frequency (18).

Table 3. Evaluation of attack frequency and severity and anthropometric measurements

Variable	Attack frequency		Attack severity		MIDAS	
	r	p	r	p	r	p
BMI	0.273	**0.007**	-0.245	**0.017**	0.057	0.584
ABSI	-0.005	0.965	0.100	0.333	-0.027	0.796
WHR	0.032	0.758	-0.024	0.819	0.264	**0.010**
WHtR	0.209	**0.042**	-0.101	0.328	0.321	**0.002**
WC	0.275	**0.007**	-0.119	0.252	0.218	0.034
NC	0.232	**0.024**	0.109	0.293	0.053	0.611
HC	0.259	**0.011**	-0.067	0.516	0.102	0.327

* WC: Waist circumference, NC: Neck Circumference, HC: Hip Circumference, BMI: Body Mass Index, ABSI: A Body Shape Index, WHR: waist to hip ratio, WHtR: waist to height ratio, MIDAS: Migraine Disability Assessment

While some studies have reported no relationship between headache duration or frequency and overweight/obesity status, other research has shown that headache frequency, duration, severity, and MIDAS scores are significantly elevated in obese patients compared to those with a normal BMI (19). In agreement with the latter findings, we observed a significant positive correlation between migraine attack frequency and BMI.

Intriguingly, our study revealed an inverse relationship between attack severity and BMI. In contrast, some previous studies have demonstrated that overweight or obese status is positively associated with pain intensity (20), while others indicated that BMI had no influence on the severity or duration of migraine attacks (21). Thus, our findings regarding attack severity diverge from these previous reports.

Our findings also showed that WC, NC, HC, and WHtR have significant correlations with migraine attack frequency. Prior literature has reported that the frequency of migraine attacks is higher in patients with an increased WC compared to those with a normal WC, though no significant relationship was found with attack severity or duration (22). In a study conducted in Türkiye, Özcan et al. identified associations of both BMI and WC with migraine (23). Consistent with our findings, a study from Iran demonstrated significant correlations between WC, WHR, WHtR, and both migraine attack frequency and severity (24).

A significant relationship was also found between MIDAS scores, HC, and WHtR in the present study; specifically, HC and WHtR were significantly higher in patients with severe disability (MIDAS Grade IV, indicating 21 or more lost days). In contrast, no significant correlations were found between BMI or other anthropometric parameters and MIDAS scores. Consistent with our findings, previous studies have reported no significant association between BMI and MIDAS scores (25). Similarly, Yu et al. demonstrated no association between obesity and MIDAS scores (26).

We did not observe a significant relationship between ABSI and migraine. Although ABSI has been widely studied, its potential association with migraine has not been previously investigated. To our knowledge, the present study is among the first to evaluate the relationship between ABSI and migraine-related clinical features. In the existing literature, ABSI has been linked to various adverse health outcomes, including cerebrovascular events, cardiovascular diseases, and malignancies (9,10).

This study has several limitations. First, manual anthropometric measurements were selected for their ease of use and rapid implementation in a clinical setting. While body composition was estimated using standard formulas based on these physical measurements, advanced imaging techniques such as computed tomography (CT) and magnetic resonance imaging (MRI) provide more precise evaluations of body

fat distribution. Future studies utilizing these imaging modalities are needed to clarify the relationship between precise body fat distribution and migraine. Second, this study was conducted at a single center, which may limit the generalizability of the findings to wider populations.

CONCLUSION

In conclusion, this study demonstrates that BMI, WHR, and HC are significantly associated with migraine. While no correlation was found between ABSI and migraine attack frequency or severity, a significant relationship was established between MIDAS scores, HC, and WHtR. These findings contribute to the growing body of evidence linking regional body fat distribution to migraine clinical features. As one of the first studies to evaluate the relationships of NC and ABSI with migraine, this work may serve as a valuable reference for future large-scale, multicenter trials.

Ethical approval

This study has been approved by the Bolu Abant Izzet Baysal University Clinical Researches Ethics Committee (approval date: 19.12.2019, number: 2019/324). Written informed consent was obtained from the participants.

Author contribution

Concept: NAG, CA; Design: CA, NAG; Analysis or interpretation: NAG, CO; Literature search: NAG, CO; Manuscript writing: NAG, CA, CO. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Steiner TJ, Paemeleire K, Jensen R, et al. European principles of management of common headache disorders in primary care. J Headache Pain. 2007; 8 Suppl 1: S3-47.
2. Natoli JL, Manack A, Dean B, et al. Global prevalence of chronic migraine: a systematic review. Cephalalgia. 2010; 30(5): 599-609. [Crossref]
3. Yuan S, Daghlas I, Larsson SC. Alcohol, coffee consumption, and smoking in relation to migraine: a bidirectional mendelian randomization study. Pain. 2022; 163(2): e342-8. [Crossref]
4. Calabrò RS. Gender effect on migraine: nature or culture? J Integr Neurosci. 2022; 21(1): 18. [Crossref]
5. Ornello R, Ripa P, Pistoia F, et al. Migraine and body mass index categories: a systematic review and meta-analysis of observational studies. J Headache Pain. 2015; 16: 27. [Crossref]
6. Vo M, Ainalem A, Qiu C, Peterlin BL, Aurora SK, Williams MA. Body mass index and adult weight gain among reproductive age women with migraine. Headache. 2011; 51(4): 559-69. [Crossref]
7. Scher AI, Stewart WF, Ricci JA, Lipton RB. Factors associated with the onset and remission of chronic daily headache in a population-based study. Pain. 2003; 106(1-2): 81-9. [Crossref]
8. Peterlin BL, Bigal ME, Tepper SJ, Urakaze M, Sheftell FD, Rapoport AM. Migraine and adiponectin: is there a connection? Cephalalgia. 2007; 27(5): 435-46. [Crossref]
9. Allais G, Chiarle G, Sinigaglia S, Benedetto C. Menstrual migraine: a review of current and developing pharmacotherapies for women. Expert Opin Pharmacother. 2018; 19(2): 123-36. [Crossref]
10. Delaruelle Z, Ivanova TA, Khan S, et al. Male and female sex hormones in primary headaches. J Headache Pain. 2018; 19(1): 117. [Crossref]
11. Sadeghi O, Askari G, Maghsoudi Z, Ghiasvand R, Khorvash F. The association between abdominal obesity and characteristics of migraine attacks in Iranian adults. Iran J Nurs Midwifery Res. 2016; 21(3): 271-7. [Crossref]
12. Miri A, Nasiri M, Zonoori S, et al. The association between obesity and migraine in a population of Iranian adults: a case-control study. Diabetes Metab Syndr. 2018; 12(5): 733-6. [Crossref]
13. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. Cephalalgia. 2018;38(1):1-211. [Crossref]
14. 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(2): 107-14; discussion 74. [Crossref]
15. Ertaş M, Siva A, Dalkara T, et al. Validity and reliability of the Turkish Migraine Disability Assessment (MIDAS) questionnaire. Headache. 2004; 44(8): 786-93. [Crossref]
16. Recober A, Peterlin BL. Migraine and obesity: moving beyond BMI. Future Neurol. 2014; 9(1): 37-40. [Crossref]

17. Mattsson P. Migraine headache and obesity in women aged 40-74 years: a population-based study. Cephalalgia. 2007; 27(8): 877-80. **[Crossref]**

18. Santos IS, Goulart AC, Passos VM, Molina MDC, Lotufo PA, Bensenor IM. Obesity, abdominal obesity and migraine: a cross-sectional analysis of ELSA-Brasil baseline data. Cephalalgia. 2015; 35(5): 426-36. **[Crossref]**

19. Togha M, Haghdoost F, Khorsha F, Razeghi Jahromi S, Ghorbani Z. Body mass index and its association with migraine characteristics in female patients. Arch Iran Med. 2019; 22(10): 554-9.

20. Galioto R, O'Leary KC, Gunstad J, et al. The role of migraine headache severity, associated features and interactions with overweight/obesity in inhibitory control. Int J Neurosci. 2018; 128(1): 63-70. **[Crossref]**

21. Khan, M. M., Huda, M. N., Bhattacharjee, M., et al. Relationship of Migraine and Body Mass Index (BMI). Journal of Enam Medical College, 2016;6(2), 80-87. **[Crossref]**

22. Rossoni de Oliveira V, Camboim Rockett F, Castro K, da Silveira Perla A, Chaves MLF, Schweigert Perry ID. Body mass index, abdominal obesity, body fat and migraine features in women. Nutr Hosp. 2013; 28(4): 1115-20. **[Crossref]**

23. Özcan RK, Özmen SG. The Association between migraine, metabolic syndrome, insulin resistance, and obesity in women: a case-control study. Sisli Etfal Hastan Tip Bul. 2019; 53(4): 395-402. **[Crossref]**

24. Afshinmajd S, Davati A, Akbari F. The effects of body mass index on the treatment of the patients with migraine headaches. Iran J Neurol. 2011; 10(3-4): 35-8.

25. Téllez-Zenteno JF, Pahwa DR, Hernandez-Ronquillo L, García-Ramos G, Velázquez A. Association between body mass index and migraine. Eur Neurol. 2010; 64(3): 134-9. **[Crossref]**

26. Yu S, Liu R, Yang X, et al. Body mass index and migraine: a survey of the Chinese adult population. J Headache Pain. 2012; 13(7): 531-6. **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):225-235
DOI: 10.54307/NWMJ.2026.239

RESEARCH ARTICLE

Cognitive performance in chronic postoperative hypoparathyroidism: A cross-sectional study

Püren Gökbulut[1], Şerife Ezgi Doğan[2], Çağatay Emir Önder[1], Işılay Taşkaldıran[1], Cavit Çulha[1], Tülay Omma[3]

[1]Department of Endocrinology and Metabolic Diseases, Ankara Training and Research Hospital, Ankara, Türkiye
[2]Department of Endocrinology and Metabolic Diseases, Sincan Training and Research Hospital, Ankara, Türkiye
[3]Department of Endocrinology and Metabolic Diseases, Lokman Hekim University Akay Hospital, Ankara, Türkiye

Cite as: Gökbulut P, Doğan ŞE, Önder ÇE, Taşkaldıran I, Çulha C, Omma T. Cognitive performance in chronic postoperative hypoparathyroidism: A cross-sectional study. Northwestern Med J. 2026;6(3):225-235.

ABSTRACT

Objective: To evaluate cognitive performance in patients with chronic postoperative hypoparathyroidism and to investigate its association with biochemical parameters.

Methods: This prospective, observational, cross-sectional study included 56 patients aged 40–65 years with chronic postoperative hypoparathyroidism, 28 post-thyroidectomy hypothyroid patients without hypoparathyroidism, and 28 healthy controls. Cognitive function was assessed using the Mini-Mental State Examination (MMSE). Biochemical parameters, including albumin-corrected serum calcium and intact parathyroid hormone (iPTH), were measured concurrently. Multivariable analyses were performed to identify independent determinants of cognitive performance.

Results: Although no significant differences were observed among the groups in terms of total MMSE scores, selective differences were detected in the registration and language subdomains in patients with chronic postoperative hypoparathyroidism. Albumin-corrected serum calcium levels and age independently predicted language performance. In addition, a positive correlation was identified between iPTH levels and registration memory scores.

Conclusion: Chronic postoperative hypoparathyroidism may be associated with mild, domain-specific cognitive alterations rather than global cognitive decline. The association of albumin-corrected serum calcium and age with language performance suggests a potential role of calcium homeostasis in cognitive function. Cognitive assessment may therefore be considered as part of long-term follow-up in these patients.

Keywords: hypoparathyroidism, calcium, mini-mental state examination, cognition

Corresponding author: Püren Gökbulut **E-mail:** puren_gokbulut@hotmail.com
Received: 26.02.2026 **Accepted:** 08.05.2026 **Published:** 31.07.2026

INTRODUCTION

Hypoparathyroidism (hypoPT) is a rare disorder of mineral metabolism characterized by hypocalcemia and hyperphosphatemia resulting from absent or inappropriately low levels of parathyroid hormone (PTH) (1,2). The most common etiology is the iatrogenic form, which develops secondary to accidental removal or damage to the parathyroid glands during neck surgery (3). Clinically, hypoparathyroidism may present as a transient condition that resolves within the first months after surgery; however, persistence beyond 12 months is defined as chronic hypoPT (4). In high-volume centers with experienced endocrine surgeons, the reported rates of chronic hypoPT following thyroid surgery range between 0.9% and 1.6%. Although conventional treatment with calcium supplementation and active vitamin D analogs constitutes the cornerstone of management in chronic hypoPT, this approach does not fully restore physiological calcium homeostasis (5). HypoPT may lead to a spectrum of neurological, psychiatric, and cognitive disturbances. Patients frequently report muscle cramps, spasms, paresthesia, sleep disturbances, anxiety, depression, and reduced quality of life (6-10). Recent evidence also suggests that hypoPT may adversely affect cognitive function (5).

Cognitive functioning encompasses a broad range of mental processes, including learning, memory, reasoning, problem-solving, decision-making, and attention (11). Patients with hypoPT commonly describe symptoms such as slowed thinking, difficulty performing daily activities, and "brain fog." Although the underlying mechanisms remain incompletely understood, intracranial calcifications and hypocalcemia have been associated with impairments in attention, memory, and information-processing speed. Moreover, cognitive dysfunction has been reported in association with extrapyramidal symptoms in this population (12-16).

PTH receptors are expressed in multiple tissues, including the central nervous system. PTH has been shown to cross the blood–brain barrier and bind to PTH2 receptors expressed in the brain, where it may contribute to the regulation of intracellular calcium homeostasis (17,18). The PTH2 receptor has also been proposed to play a role in limbic, neuroendocrine, and sensory-protective functions (19). Calcium ions are fundamental to synaptic transmission and neuronal cell function (5). Furthermore, PTH may influence regional cerebral capillary blood flow, and alterations observed in hypoPT have been suggested to be associated with disturbances in neurovascular coupling (20).

Data regarding the objective assessment of cognitive function in patients with hypoPT are limited, and most available studies rely predominantly on patients' subjective evaluations of well-being. Therefore, in the present study, we aimed to evaluate the relationship between cognitive function and biochemical parameters in patients with chronic hypoPT using the Mini-Mental State Examination (MMSE).

MATERIALS AND METHODS

Study population

This study was designed as a prospective, observational, single-center study. A total of 112 individuals aged between 40 and 65 years who presented for routine follow-up at the Endocrinology Outpatient Clinic of Ankara Training and Research Hospital between January 2021 and July 2023 were enrolled. The postoperative hypoparathyroidism group (Po-hypoPT [+]) consisted of patients who had undergone total thyroidectomy at least 12 months prior to enrollment, had documented hypocalcemia accompanied by low or inappropriately normal PTH levels, and were receiving calcium supplementation and/or active vitamin D therapy (n = 56). The comparison group included patients with a history of total thyroidectomy who had hypothyroidism but did not develop hypoparathyroidism (Po-hypoPT [–]) (n = 28). The control group comprised healthy individuals aged 40–65 years with no history of thyroid surgery, no chronic medical conditions, and no regular medication use (n = 28). The exclusion criteria were as follows: pregnancy; incomplete or insufficient clinical data; history of parathyroidectomy due to primary, secondary, or tertiary hyperparathyroidism; follow-up duration of less than 12 months after total thyroidectomy; previously diagnosed psychiatric disease or cognitive impairment; use of psychotropic medications; severe renal insufficiency (glomerular filtration rate <30

mL/min); chronic liver disease; presence of diabetes mellitus or hypertension; inadequate thyroid hormone control under levothyroxine therapy; anemia; vitamin B12 or iron deficiency; serum 25-hydroxyvitamin D levels <20 ng/mL; other systemic diseases affecting calcium–phosphorus metabolism or PTH levels; use of medications that may pharmacologically interact with calcium; and any condition precluding administration of the Mini-Mental State Examination (e.g., illiteracy, refusal to undergo testing).

Study protocol

Demographic and clinical data, including age, sex, educational level, time since thyroidectomy, presence of hypocalcemia-related symptoms, and doses of calcium, calcitriol, and levothyroxine were recorded for all participants. Height and body weight were measured during physical examination, and body mass index (BMI) was calculated accordingly. Laboratory assessments included measurements of serum glucose, alanine aminotransferase (ALT), urea, creatinine, alkaline phosphatase (ALP), thyroid-stimulating hormone (TSH), free thyroxine (fT4), parathyroid hormone (PTH), total calcium, albumin, phosphorus, 25-hydroxyvitamin D [25(OH)D], complete blood count, ferritin, and vitamin B12 levels. Corrected serum calcium levels were calculated using the following formula: Corrected calcium (mg/dL) = Total serum calcium (mg/dL) + 0.8 × [4 – serum albumin (g/dL)]. In patients diagnosed with hypoparathyroidism, 24-hour urinary calcium and creatinine levels were additionally evaluated. For patients who had undergone cranial computed tomography (CT) within the previous two years, reports retrieved from the hospital's electronic medical record system were reviewed for the presence of calcifications in the basal ganglia. MMSE was administered to all participants on the day of clinical evaluation by the same physician to ensure standardization. Blood samples for biochemical analyses were obtained within one hour following completion of the cognitive assessment. This approach was adopted to ensure that biochemical parameters and neurocognitive performance reflected the participants' concurrent clinical status.

Mini-mental state examination

MMSE was first described by Folstein et al. (21). It is a brief (5–10 minutes), practical, and standardized screening instrument designed to provide a quantitative assessment of cognitive performance. The MMSE has been widely used across different cultural and ethnic populations worldwide and has been translated and culturally adapted from its original English version into Turkish, with established validity and reliability in the Turkish population (Supplementary Appendix 1) (22-25). To minimize inter-examiner variability, a standardized administration guideline has been developed (Supplementary Appendix 2) (23). The test consists of 11 items covering five cognitive domains: orientation (10 points), registration (3 points), attention and calculation (5 points), recall (3 points), and language (9 points), yielding a total possible score ranging from 0 to 30. Higher scores indicate better cognitive functioning (23). In standard administration, the total score is calculated as the sum of correct responses. However, this method may be misleading when participants refuse to answer specific items. In epidemiological studies, refusal to respond is often considered to reflect an inability to provide the correct answer (26). Therefore, in cases where participants did not complete all items, only the scores obtained from the completed items were included in the analysis.

In the Turkish population, the MMSE has demonstrated high sensitivity (0.91) and specificity (0.95) for the diagnosis of mild dementia using a cut-off value of 24, with positive and negative predictive values of 0.90 and 0.95, respectively (23,25). In the present study, consistent with the literature, a threshold value of 24 was adopted. MMSE scores were not used for diagnostic purposes; rather, they were evaluated as a screening indicator reflecting the participants' overall level of cognitive performance.

Statistical analysis

Statistical analyses were performed using SPSS software (version 26.0; IBM Corp., Armonk, NY, USA). The normality of data distribution was assessed using the Shapiro–Wilk test. Continuous variables were presented as mean ± standard deviation (SD)

when normally distributed and as median (minimum–maximum) when not normally distributed. Categorical variables were expressed as numbers and percentages. For comparisons between independent groups, parametric variables were analyzed using the Student's *t*-test (for two groups) or one-way analysis of variance (ANOVA) (for ≥3 groups), whereas non-parametric variables were analyzed using the Mann–Whitney *U* test (for two groups) or the Kruskal–Wallis test (for ≥2 groups). In post hoc analyses following ANOVA, Tukey's test was used when homogeneity of variances was confirmed, and the Games–Howell test was applied when the assumption of homogeneity was violated. For pairwise comparisons following the Kruskal–Wallis test, Bonferroni correction was employed. Associations between variables were evaluated using Pearson correlation analysis for parametric variables and Spearman correlation analysis for non-parametric variables. Independent risk factors predicting an MMSE score <24 were identified using logistic regression analysis. A *p-value* <0.05 was considered statistically significant. For post hoc analyses, a threshold of *p* <0.017 was applied.

Sample size calculation

The sample size was calculated using G*Power software (version 3.1.9.4). Based on a power of 90%, an effect size of 0.4, and a significance level (α) of 0.05, the minimum required sample size was determined to be 84 participants. To increase the statistical power of the study, a total of 112 participants were ultimately included: 56 in the Po-hypoPT (+) group, 28 in the Po-hypoPT (–) group, and 28 in the control group.

RESULTS

The demographic and clinical characteristics of the study population are presented in Table 1. No statistically significant differences were observed among the Po-hypoPT (+), Po-hypoPT (–), and control groups in terms of age, sex distribution, educational level, smoking status, or body mass index (BMI) (p > 0.05 for all comparisons). The laboratory findings of the study population are summarized in Table 2. Albumin-corrected serum calcium levels were significantly lower in the Po-hypoPT (+) group compared with both

Table 1. Demographic and clinical characteristics of the study population

Demographic and Clinical Characteristics	Po-hypoPT (+) (n=56)	Po-hypoPT (–) (n=28)	Control (n=28)	p value
Age, years	48.62 ± 7.12	51.32 ± 5.65	47.75 ± 5.68	0.093
Sex, Female, n (%)	53 (94.6%)	25 (89.3%)	25 (89.3%)	0.482
Education level, n (%)				0.873
Primary/secondary school	43 (76.8%)	21 (75%)	23 (82.1%)	
High school or higher	13 (23.2%)	7 (25%)	5 (17.9%)	
Smoking, n (%)	12 (21.4%)	6 (21.4%)	10 (35.7%)	0.155
BMI, kg/m²	28.62 ± 4.40	28.39 ± 4.83	30.61 ± 6.12	0.166
Time since thyroid surgery, years	14 (3–32)	14.5 (1–30)	–	N/A
Symptomatic hypocalcemia, n (%)	32 (57.1%)	–	–	N/A
Mean LT4 dose, µg/day	112.5 (50–200)	100 (75–175)	–	N/A
Mean calcitriol dose, µg/day	0.5 (0.25–2)	–	–	N/A
Mean calcium dose, g/day	2 (0–4)	–	–	N/A

Data are presented as n (%), mean ± standard deviation, or median (minimum–maximum). For comparisons between independent groups, Student's t-test (for two groups) and ANOVA (for ≥2 groups) were used for normally distributed variables. Categorical variables were compared using the chi-square (χ^2) test or Fisher's exact test. A p-value <0.05 was considered statistically significant. Statistically significant values ($p < 0.05$) were marked with an asterisk (*) and indicated in bold. p_a denotes comparisons between the Po-hypoPT (+) and Po-hypoPT (–) groups. For the time since thyroid surgery, p_a=0.902; for the mean daily LT4 dose, p_a=0.312.

Abbreviations: LT4, levothyroxine sodium; Po-hypoPT, postoperative hypoparathyroidism; BMI, body mass index.

the Po-hypoPT (–) and control groups (both $p < 0.001$). No significant difference was observed between the Po-hypoPT (–) and control groups ($p = 0.994$). Serum phosphorus levels were significantly higher in the Po-hypoPT (+) group compared with the other two groups ($p < 0.001$), whereas no significant difference was detected between the Po-hypoPT (–) and control groups ($p = 0.546$). Intact parathyroid hormone (iPTH) levels were markedly lower in the Po-hypoPT (+) group, while no significant difference was found between the Po-hypoPT (–) and control groups ($p = 0.851$). A comparative box-plot graph of cognitive assessment scores is shown in Figure 1. No statistically significant difference was observed among the Po-hypoPT (+), Po-hypoPT (–), and healthy control groups with respect to total Mini-Mental State Examination (MMSE) scores ($p = 0.115$). However, among the cognitive subdomains, significant differences were identified in the registration and language subscales ($p = 0.002$ and $p = 0.031$, respectively).

Correlation analysis

Correlation analyses were performed in the overall study population to examine the associations between registration and language subscale scores and participants' demographic, clinical, and laboratory parameters. A positive correlation was observed between the registration score and intact parathyroid hormone (iPTH) levels ($r = 0.257$, $p = 0.006$). The language score showed a negative correlation with age ($r = -0.230$, $p = 0.015$) and a positive correlation with albumin-corrected serum calcium (Alb-cCa) levels ($r = 0.204$, $p = 0.031$). The distribution of the study population according to total cognitive assessment scores is illustrated in Figure 2. When groups were compared based on the total MMSE score, the proportion of participants scoring <24 was 58.9% in the Po-hypoPT (+) group, 42.9% in the Po-hypoPT (–) group, and 25% in the control group. A statistically significant difference was found among the groups with respect to MMSE total score distribution ($p = 0.015$). In pairwise comparisons, the proportion of participants with MMSE scores <24 was significantly higher in the Po-hypoPT (+) group compared with the control group ($p = 0.003$). In contrast, no significant difference was observed between the Po-hypoPT (–) and control groups regarding MMSE total score distribution ($p > 0.05$). The comparison of participants' characteristics after categorization according to the MMSE total score cut-off value of 24 is presented in Table 3.

Table 2. Laboratory findings of the study population

Laboratory parameters	Po-hypoPT (+) (n=56)	Po-hypoPT (-) (n=28)	Control (n=28)	p value
FPG, mg/dl	91.75±9.62	96.89±8.76	94.57±8.99	0.053
Creatinine, mg/dl	0.74±0.12	0.73±0.13	0.71±0.13	0.723
TSH, µIU/L	1.49 (0.52-4.4)	1.56 (0.53-3.91)	1.71 (0.56-3.86)	0.623
Alb-cCa, mg/dl	8±0.62	9.08±0.47	9.07±0.27	**<0.001***
Phosphorus, mg/dl	4.45±0.76	3.73±0.47	3.61±0.39	**<0.001***
25 (OH) Vitamin D, µg/L	34.95 (21-56)	32.9 (22.6-71.8)	29.7 (22.5-46)	**0.006***
iPTH, ng/L	11.56±6.1	43.05±16.18	44.96±9.23	**<0.001***
24 h urinary Ca, mg/day	167.5 (11-687)	-	-	N/A

Data are presented as n (%), mean ± standard deviation, or median (minimum–maximum). For comparisons among independent groups, one-way ANOVA was used for normally distributed variables and the Kruskal–Wallis test for non-normally distributed variables. Post hoc analyses were performed using the Tukey test when variances were homogeneous and the Games–Howell test when variances were not homogeneous; Bonferroni correction was applied after the Kruskal–Wallis test. Categorical variables were compared using the chi-square (χ^2) test or Fisher's exact test, as appropriate. A p-value <0.05 was considered statistically significant; for post hoc comparisons, p <0.017 was considered significant. Statistically significant values are indicated with an asterisk (*).pa denotes comparisons between the Po-hypoPT (+) and Po-hypoPT (–) groups, pb denotes comparisons between the Po-hypoPT (+) and Control groups, and pc denotes comparisons between the Po-hypoPT (–) and Control groups. For Alb-cCa, pa<0.001, pb<0.001, and pc=0.994; for phosphorus, pa<0.001, pb<0.001, and pc=0.546; for 25(OH) vitamin D, pa=0.377, pb=0.001, and pc=0.073; for iPTH, pa<0.001, pb<0.001, and pc=0.851.

Alb-cCa: Albumin-corrected serum calcium, ALP: Alkaline phosphatase, FPG: Fasting plasma glucose, Ca: Calcium; h: Hour; iPTH: Intact parathyroid hormone, Po-hypoPT: Postoperative hypoparathyroidism, TSH: Thyroid-stimulating hormone.

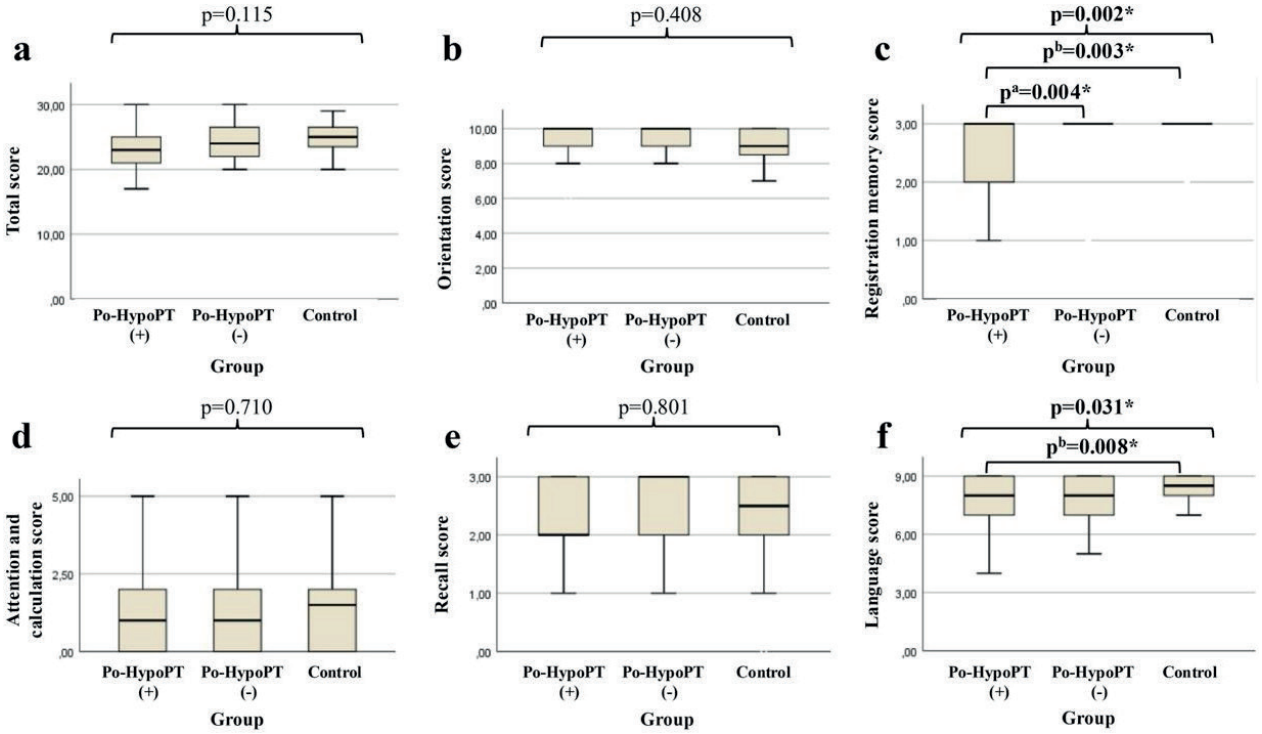


Figure 1. Comparative box plots of cognitive assessment scores among the study groups. The figure presents the following subdomains: (a) Total score, (b) Orientation, (c) Registration memory, (d) Attention and calculation, (e) Recall, and (f) Language. Comparisons among the independent groups were performed using one-way analysis of variance (ANOVA) for normally distributed variables and the Kruskal–Wallis test for non-normally distributed variables. For post hoc analyses, Tukey's test was applied following ANOVA when the assumption of homogeneity of variances was met, whereas the Games–Howell test was used when variances were unequal. Following the Kruskal–Wallis test, pairwise comparisons were performed using the Bonferroni correction. A two-sided p value <0.05 was considered statistically significant; however, for post hoc pairwise comparisons, p<0.017 was regarded as statistically significant. Statistically significant differences (p<0.05) are indicated by an asterisk (*) and shown in bold. . pa indicates the comparison between the Po-HypoPT (+) and Po-HypoPT (–) groups, whereas pb indicates the comparison between the Po-HypoPT (+) and control groups.

DISCUSSION

In the present study, although no significant difference was observed among the groups in terms of total MMSE scores, selective cognitive involvement was noted in patients with hypoparathyroidism, particularly in the domains of registration and language. These findings suggest that domain-specific cognitive alterations may be present even when global cognitive performance appears to be preserved.

Early detection of cognitive dysfunction is important for identifying patients at risk and planning appropriate intervention strategies. However, comprehensive neuropsychological assessments may be time-consuming and costly in routine clinical practice. Therefore, easily applicable screening tools may be useful in detecting early cognitive changes, particularly in high-risk patient populations (14). The Mini-Mental State Examination can be considered a practical, time-efficient, and cost-effective screening instrument for the preliminary evaluation of cognitive function (23).

In recent years, the relationship between PTH and cognitive function has attracted increasing attention. Several studies have reported that elevated PTH levels may be associated with cognitive decline and an increased risk of dementia (27-29). However,

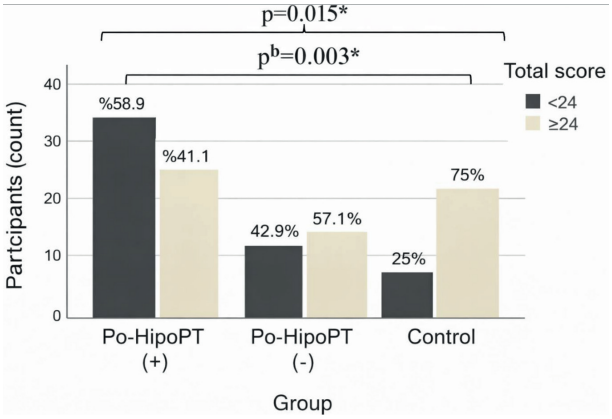


Figure 2. Distribution of the study population according to total cognitive assessment scores. Participants were categorized into two groups based on the Mini-Mental State Examination (MMSE) cut-off score of 24 as <24 and ≥24. Comparisons of categorical variables were performed using the chi-square (χ^2) test or Fisher's exact test, as appropriate. A two-sided p value <0.05 was considered statistically significant. Statistically significant differences ($p<0.05$) are indicated by an asterisk (*) and shown in bold. pb indicates the comparison between the Po-HypoPT (+) and control groups.

evidence also suggests that the low PTH levels and hypocalcemia observed in hypoparathyroidism may likewise be linked to adverse cognitive outcomes (15,30). These findings imply that the effects of PTH and calcium homeostasis on cognitive processes may be bidirectional and complex. Hypoparathyroidism is a rare disorder, and the available evidence regarding its association with cognitive dysfunction remains limited. In the majority of published studies, cognitive status has been assessed using self-reported measures of quality of life rather than objective, performance-based cognitive tests. Instruments such as the Short Form Health Survey (SF-36) have been widely used for this purpose. This reliance on subjective assessments creates uncertainty regarding the extent to which the reported cognitive impairment in hypoPT is objectively measurable and clinically meaningful (31). Abnormal PTH levels have been hypothesized to affect cognitive function through mechanisms such as neuronal calcium imbalance, cerebral hypoperfusion, and impaired neuronal signaling. However, a systematic review conducted within this framework did not demonstrate a significant association between PTH levels and cognitive dysfunction, suggesting that cognitive impairment may be more closely related to disturbances in calcium metabolism rather than PTH itself (17). In a web-based survey including 374 patients with hypoPT, the majority of participants reported fatigue, emotional disturbances, and cognitive complaints (12). In contrast, another study evaluating 19 patients—18 of whom had postoperative hypoPT—reported possible cognitive impairment in 68% of patients based on global cognitive scores and in 32% when premorbid-adjusted fluid cognition scores were used. Processing speed was identified as the most frequently affected cognitive domain, and impairment in this domain was associated with lower serum calcium and higher serum phosphate levels (31). In a study conducted by Aggarwal et al. involving 62 patients with idiopathic hypoPT and employing comprehensive neuropsychological testing, neuropsychological dysfunction characterized by reduced inhibitory control, visuospatial impairment, and psychomotor slowing was identified in 32% of patients. Poor cognitive performance was associated with longer disease duration, increased serum calcium–phosphorus product, and lower total serum calcium levels. An important limitation of this study was the exclusion of patients with postoperative hypoPT (15). Similarly, in a study by Saponaro et al., which evaluated 33 patients with postoperative hypoPT receiving conventional therapy using standardized neuropsychological tests, no overt global cognitive impairment was detected. However, visuospatial attention, executive function, and semantic memory were found to be modulated by (Alb-cCa) levels, with lower Alb-cCa levels associated with poorer performance in these domains (20). The variability in the reported prevalence of cognitive impairment in the literature may be related to heterogeneity in the etiopathogenesis of hypoPT (idiopathic versus postoperative), differences in disease duration, variability in neurocognitive assessment methods, and patient age.

In our study, 58.9% of patients with chronic postoperative hypoparathyroidism had an MMSE total score <24, and this proportion was significantly higher compared with the control group. These findings

Table 3. Comparison of participants' characteristics after categorization according to the total score cut-off value of 24

Demographic and Clinical Characteristics	Total score <24 (n=52)	Total score ≥24 (n=60)	p value
Group, n (%)			0.015*
Po-hypoPT (+)	33 (63.5%)	23 (38.3%)	
Po-hypoPT (–)	12 (23.1%)	16 (26.7%)	
Control	7 (13.5%)	21 (35%)	
Age, years	50.09 ± 6.38	48.26 ± 6.57	0.126
Sex, Female, n (%)	50 (48.5%)	53 (51.5%)	0.172
Education level, n (%)			0.042*
Primary/secondary school	45 (86.5%)	42 (70%)	
High school or higher	7 (13.5%)	18 (30%)	
Smoking, n (%)	16 (26.7%)	10 (19.2%)	0.379
BMI, kg/m²	29.04 ± 4.22	29.08 ± 5.66	0.969
Laboratory parameters			
FPG, mg/dL	91.26 ± 9.21	95.88 ± 9.15	0.009*
Creatinine, mg/dL	0.72 ± 0.11	0.74 ± 0.13	0.522
TSH, IU/L	1.21 (0.52–4.31)	1.71 (0.56–4.4)	0.208
Ca, mg/dL	8.9 ± 0.73	9.08 ± 0.72	0.197
Alb-cCa, mg/dL	8.39 ± 0.73	8.66 ± 0.74	0.052
Phosphorus, mg/dL	4.16 ± 0.74	3.98 ± 0.71	0.208
25(OH) Vitamin D, ng/mL	33.25 (22.2–50)	33.95 (21–71.8)	0.809
iPTH, ng/L	24.11 ± 18.58	30.97 ± 19.32	0.059

Data are presented as n (%), mean ± standard deviation, or median (minimum–maximum). For comparisons between independent groups, Student's *t*-test was used for normally distributed variables and the Mann–Whitney *U* test for non-normally distributed variables. Categorical variables were compared using the chi-square (χ^2) test or Fisher's exact test. A p-value <0.05 was considered statistically significant. Statistically significant values ($p < 0.05$) are indicated with an asterisk (*) and shown in bold. p_a denotes comparisons between the Po-hypoPT (+) and Po-hypoPT (–) groups; p_β denotes comparisons between the Po-hypoPT (+) and Control groups; *pc* denotes comparisons between the Po-hypoPT (–) and Control groups. For Group: p_a=0.164, p_β=0.003, *pc*=0.158 For Education level: p_a=0.015, p_β=0.057, *pc*=0.518

Abbreviations: FPG, fasting plasma glucose; Alb-cCa, albumin-corrected serum calcium; Ca, calcium; TSH, thyroid-stimulating hormone; ALP, alkaline phosphatase; iPTH, intact parathyroid hormone; Po-hypoPT, postoperative hypoparathyroidism.

suggest that cognitive impairment in this patient population should not be overlooked.

Statistically significant between-group differences were observed only in the registration and language subdomains of the cognitive assessment. The positive correlation identified between registration scores and iPTH levels may be related to the effects of PTH on cerebral blood flow and its regulatory role in calcium–phosphate homeostasis. This finding suggests potential neurophysiological effects of PTH on the central nervous system. Sikjaer et al. reported reduced long-term memory performance in patients with chronic hypoparathyroidism and demonstrated that disease duration may influence thalamic volume, highlighting a possible link between cognitive symptoms and structural brain alterations (32). In contrast, Berisha et al., in a study comparing 30 patients with chronic hypoparathyroidism to healthy controls, found no significant differences in verbal learning, memory, or executive function domains (5). Language emerged as another cognitive domain exhibiting significant

between-group differences. Language scores showed a negative correlation with age and a positive correlation with albumin-corrected serum calcium levels. Age-related cognitive change is a well-established physiological process, and language function is among the domains known to be sensitive to aging. Language represents a multidimensional and complex cognitive construct; visual confrontation naming performance has been reported to remain relatively stable until approximately 70 years of age, followed by a gradual decline. Similarly, verbal fluency—the ability to generate words within a specific category over a defined time interval—tends to decrease with advancing age (33). The observed negative association between language scores and age therefore appears consistent with expected physiological cognitive aging. Cognitive impairment in hypoparathyroidism has most frequently been described in the domains of attention/processing speed and executive function. Dysfunction in these areas may indirectly influence language-related tasks such as verbal fluency and naming (5). Given the central role of calcium in synaptic transmission and neuronal excitability, transient slowing of processing speed during hypocalcemia has been reported (20,31). This mechanism may partly explain the association observed between lower serum calcium levels and reduced language performance.

Neurological manifestations in hypoparathyroidism have been suggested to be associated with intracranial calcifications. These calcifications may impair cognitive function by altering cerebral blood flow and disrupting neurotransmission (34). Kowdley et al. reported cognitive impairment in 65% of individuals with hypoparathyroidism and attributed this finding to the presence of intracranial calcifications (30). In the present study, the absence of basal ganglia calcifications may be explained by the inclusion of patients with postoperative hypoparathyroidism who were treated at an early stage of the disease.

No significant sex-related differences were observed in cognitive function scores in the present study. In contrast, Aggarwal et al. reported a higher prevalence of neurocognitive dysfunction in women compared with men among patients with idiopathic hypoparathyroidism, suggesting that socioeconomic and cultural factors might contribute to this disparity (15). Thyroid dysfunction, particularly hypothyroidism, is considered an important confounding factor in cognitive assessment and has been associated with adverse effects on attention, memory, and psychomotor speed (35). In our cohort, no significant differences in MMSE scores were observed between patients who developed hypoparathyroidism after thyroidectomy and those who did not. Because all post-thyroidectomy patients had adequate thyroid hormone replacement and were biochemically euthyroid, the observed findings are unlikely to be attributable to differences in thyroid function. This finding suggests that maintenance of euthyroidism through appropriate thyroid hormone replacement may largely mitigate cognitive effects attributable to hypothyroidism. Furthermore, the absence of significant differences in MMSE scores between groups supports the notion that the observed cognitive changes may be more closely related to hypoparathyroidism-specific mechanisms rather than thyroid function per se. Nevertheless, given that the MMSE is a screening instrument designed primarily to assess global cognitive function, its sensitivity in detecting subtle, domain-specific cognitive differences may be limited.

In the present study, cognitive assessment in patients with hypoparathyroidism was conducted using a screening and classification tool intended to reflect overall cognitive performance rather than for diagnostic purposes. The absence of comprehensive, performance-based neuropsychological test batteries represents an important limitation of the study. In addition, the relatively small sample size and the cross-sectional design restrict the ability to draw causal inferences. The lack of longitudinal follow-up data examining the relationship between improvements in biochemical parameters and changes in MMSE scores constitutes another limitation. Furthermore, basal ganglia calcifications were evaluated only in patients who had previously undergone cranial computed tomography, and systematic neuroimaging was not performed in all participants.

CONCLUSION

The findings of this study suggest that cognitive function may be adversely affected in patients with hypoparathyroidism. The independent association of

albumin-corrected serum calcium levels and age with language performance supports the potential impact of chronic hypocalcemia and aging on language-related cognitive domains. These results indicate that postoperative hypoparathyroidism should be considered not merely as a biochemical disorder but also as a clinical condition potentially associated with domain-specific cognitive alterations. Accordingly, a comprehensive approach incorporating the assessment of cognitive function alongside biochemical parameters may be warranted during the long-term follow-up of patients with hypoparathyroidism.

Ethical approval

The study was approved by the Ethics Committee of Ankara Training and Research Hospital (Approved date and number: October 27, 2021; No. E-21-784). It was conducted in accordance with the principles of the Declaration of Helsinki.

Author contribution

Conceptualised and supervised the study: PG, TO; Concept: PG, CÇ, TO; Design: ŞED, ÇEÖ, IT; Data Collection or Processing: PG, ŞED; Analysis or Interpretation: PG, ÇEO, CÇ; Literature Search: PG, TO; Writing: PG, IT. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Shoback DM, Bilezikian JP, Costa AG, et al. Presentation of hypoparathyroidism: etiologies and clinical features. J Clin Endocrinol Metab. 2016; 101(6): 2300-12. [Crossref]
2. Palermo A, Santonati A, Tabacco G, et al. PTH(1-34) for surgical hypoparathyroidism: a 2-year prospective, open-label investigation of efficacy and quality of life. J Clin Endocrinol Metab. 2018; 103(1): 271-80. [Crossref]
3. Underbjerg L, Sikjaer T, Mosekilde L, Rejnmark L. Postsurgical hypoparathyroidism--risk of fractures, psychiatric diseases, cancer, cataract, and infections. J Bone Miner Res. 2014; 29(11): 2504-10. [Crossref]
4. Bilezikian JP, Khan A, Potts JT, et al. Hypoparathyroidism in the adult: epidemiology, diagnosis, pathophysiology, target-organ involvement, treatment, and challenges for future research. J Bone Miner Res. 2011; 26(10): 2317-37. [Crossref]
5. Tmava-Berisha A, Fahrleitner-Pammer A, Stross T, et al. Cognitive function in individuals with chronic hypoparathyroidism-a prospective observational study. J Clin Endocrinol Metab. 2025; 110(8): 2157-63. [Crossref]
6. Jørgensen CU, Homøe P, Dahl M, Hitz MF. Postoperative chronic hypoparathyroidism and quality of life after total thyroidectomy. JBMR Plus. 2021; 5(4): e10479. [Crossref]
7. Nicolai A, Lazzarino LG. Dementia syndrome in patients with postsurgical hypoparathyroidism and extensive brain calcifications. Eur Neurol. 1994; 34(4): 230-5. [Crossref]
8. Bohrer T, Krannich JH. Depression as a manifestation of latent chronic hypoparathyroidism. World J Biol Psychiatry. 2007; 8(1): 56-9. [Crossref]
9. Vokes T. Quality of life in hypoparathyroidism. Bone. 2019; 120: 542-7. [Crossref]
10. Abate EG, Clarke BL. Review of hypoparathyroidism. Front Endocrinol (Lausanne). 2017; 7: 172. [Crossref]
11. Fisher GG, Chaffee DS, Tetrick LE, Davalos DB, Potter GG. Cognitive functioning, aging, and work: a review and recommendations for research and practice. J Occup Health Psychol. 2017; 22(3): 314-36. [Crossref]
12. Hadker N, Egan J, Sanders J, Lagast H, Clarke BL. Understanding the burden of illness associated with hypoparathyroidism reported among patients in the PARADOX study. Endocr Pract. 2014; 20(7): 671-9. [Crossref]
13. Goswami R, Sharma R, Sreenivas V, Gupta N, Ganapathy A, Das S. Prevalence and progression of basal ganglia calcification and its pathogenic mechanism in patients with idiopathic hypoparathyroidism. Clin Endocrinol (Oxf). 2012; 77(2): 200-6. [Crossref]
14. Kumar G, Kaur D, Aggarwal P, Khurana T. Hypoparathyroidism presenting as cognitive dysfunction. BMJ Case Rep. 2013; 2013: bcr2013009220. [Crossref]
15. Aggarwal S, Kailash S, Sagar R, et al. Neuropsychological dysfunction in idiopathic hypoparathyroidism and its relationship with intracranial calcification and serum total calcium. Eur J Endocrinol. 2013; 168(6): 895-903. [Crossref]
16. Terada T, Kakimoto A, Yoshikawa E, et al. The possible link between gabaergic dysfunction and cognitive decline in a patient with idiopathic hypoparathyroidism. Intern Med. 2015; 54(17): 2245-50. [Crossref]
17. Lourida I, Thompson-Coon J, Dickens CM, et al. Parathyroid hormone, cognitive function and dementia: a systematic review. PLoS One. 2015; 10(5): e0127574. [Crossref]
18. Arlt W, Fremerey C, Callies F, et al. Well-being, mood and calcium homeostasis in patients with hypoparathyroidism receiving standard treatment with calcium and vitamin D. Eur J Endocrinol. 2002; 146(2): 215-22. [Crossref]

19. Meola A, Vignali E, Matrone A, Cetani F, Marcocci C. Efficacy and safety of long-term management of patients with chronic post-surgical hypoparathyroidism. J Endocrinol Invest. 2018; 41(10): 1221-6. **[Crossref]**

20. Saponaro F, Alfi G, Cetani F, et al. Serum calcium levels are associated with cognitive function in hypoparathyroidism: a neuropsychological and biochemical study in an Italian cohort of patients with chronic post-surgical hypoparathyroidism. J Endocrinol Invest. 2022; 45(10): 1909-18. **[Crossref]**

21. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12(3): 189-98. **[Crossref]**

22. Tombaugh TN, McIntyre NJ. The Mini-Mental State Examination: a comprehensive review. J Am Geriatr Soc. 1992; 40(9): 922-35. **[Crossref]**

23. Güngen C, Ertan T, Eker E, Yaşar R, Engin F. Reliability and validity of the standardized Mini Mental State Examination in the diagnosis of mild dementia in Turkish population. Turk Psikiyatri Derg. 2002; 13(4): 273-81.

24. Keskinoglu P, Ucku R, Yener G. Pretest results of the revised standardised Mini Mental Examination test in community dwelling elderly. J Neurol Sci. 2008; 25(1): 18-24.

25. Keskinoglu P, Ucku R, Yener G, Yaka E, Kurt P, Tunca Z. Reliability and validity of revised Turkish version of Mini Mental State Examination (rMMSE-T) in community-dwelling educated and uneducated elderly. Int J Geriatr Psychiatry. 2009; 24(11): 1242-50. **[Crossref]**

26. Fillenbaum GG, George LK, Blazer DG. Scoring nonresponse on the Mini-Mental State Examination. Psychol Med. 1988; 18(4): 1021-5. **[Crossref]**

27. Björkman MP, Sorva AJ, Tilvis RS. Does elevated parathyroid hormone concentration predict cognitive decline in older people? Aging Clin Exp Res. 2010; 22(2): 164-9. **[Crossref]**

28. Walker MD, McMahon DJ, Inabnet WB, et al. Neuropsychological features in primary hyperparathyroidism: a prospective study. J Clin Endocrinol Metab. 2009; 94(6): 1951-8. **[Crossref]**

29. Roman SA, Sosa JA, Mayes L, et al. Parathyroidectomy improves neurocognitive deficits in patients with primary hyperparathyroidism. Surgery. 2005; 138(6): 1121-8; discussion 1128-9. **[Crossref]**

30. Kowdley KV, Coull BM, Orwoll ES. Cognitive impairment and intracranial calcification in chronic hypoparathyroidism. Am J Med Sci. 1999; 317(5): 273-7. **[Crossref]**

31. Rubin MR, Tabacco G, Omeragic B, Majeed R, Hale C, Brickman AM. A pilot study of cognition among hypoparathyroid adults. J Endocr Soc. 2022; 6(3): bvac002. **[Crossref]**

32. Sikjaer T, Eskildsen SF, Underbjerg L, Østergaard L, Rejnmark L, Evald L. Hypoparathyroidism: changes in brain structure, cognitive impairment, and reduced quality of life. J Bone Miner Res. 2024; 39(7): 855-66. **[Crossref]**

33. Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. Clin Geriatr Med. 2013; 29(4): 737-52. **[Crossref]**

34. Sardella A, Bellone F, Morabito N, et al. The association between hypoparathyroidism and cognitive impairment: a systematic review. J Endocrinol Invest. 2021; 44(5): 905-19. **[Crossref]**

35. Samuels MH. Psychiatric and cognitive manifestations of hypothyroidism. Curr Opin Endocrinol Diabetes Obes. 2014; 21(5): 377-83. **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):236-241
DOI: 10.54307/NWMJ.2026.249

RESEARCH ARTICLE

Ryan tumor regression scoring in rectal cancers operated after neoadjuvant therapy; single-center experience

Songül Peltek Özer[1], **Bahri Özer**[2]

[1]Department of Pathology, Abant İzzet Baysal University Hospital, Bolu, Türkiye
[2]Department of General Surgery, Abant İzzet Baysal University Hospital, Bolu, Türkiye

Cite as: Peltek Özer S, Özer B. Ryan tumor regression scoring in rectal cancers operated after neoadjuvant therapy; single-center experience. Northwestern Med J. 2026;6(3):236-241.

ABSTRACT

Aim: This study aimed to evaluate pathological response rates using the Ryan tumor regression grading system in rectal cancer patients who underwent surgery after neoadjuvant therapy and to investigate the relationship between response levels and clinicopathological parameters.

Materials and Methods: In this retrospective single-center study, 58 patients diagnosed with rectal cancer between January 2020 and December 2025 and treated with neoadjuvant therapy followed by surgery were included. Patients who did not receive neoadjuvant therapy were excluded. Demographic, clinical, and pathological data were recorded. Tumor response was graded according to the Ryan tumor regression system, and associations with clinicopathological variables were analyzed statistically.

Results: Ryan score distribution was 10.3% (n=6) for Score 0, 20.7% (n=12) for Score 1, 25.9% (n=15) for Score 2, and 43.1% (n=25) for Score 3. Residual tumor size increased significantly as pathological response decreased ($p<0.001$). No statistically significant associations were found between Ryan score and age, gender, total or metastatic lymph node count, T stage, lymphovascular invasion, perineural invasion, or tumor localization.

Conclusion: The Ryan tumor regression score is a practical and reliable method for evaluating treatment response in rectal cancer after neoadjuvant therapy. In this series, residual tumor size was significantly associated with treatment response, whereas other clinicopathological parameters were not.

Keywords: rectal cancer, neoadjuvant therapy, Ryan tumor regression score

INTRODUCTION

Rectal cancer constitutes a significant proportion of colorectal cancers and requires a multidisciplinary treatment approach, particularly in patients with locally advanced disease. In recent years, preoperative neoadjuvant chemoradiotherapy (CRT) has been adopted as the standard treatment protocol for

Corresponding author: Bahri Özer **E-mail:** bahriozer@hotmail.com
Received: 13.04.2026 **Accepted:** 22.06.2026 **Published:** 31.07.2026

locally advanced rectal cancer (1). The primary aims of neoadjuvant therapy are to reduce tumor size, increase the rate of surgical resectability, enhance the likelihood of sphincter-preserving surgery, and decrease the risk of local recurrence (2).

The histopathological evaluation of tumor response following neoadjuvant therapy is of great importance for both prognostic prediction and planning adjuvant treatment. For this purpose, several tumor regression grading systems have been developed. Among these, the Ryan tumor regression grading system (Tumor Regression Grade, TRG) has become the standard system recommended by the American Joint Committee on Cancer (AJCC) and the College of American Pathologists (CAP) (3,4). Numerous studies have shown that patients with low Ryan score (0–1) have lower local recurrence rates and significantly longer disease-free and overall survival (5-7).

In the present study, we aimed to evaluate the pathological response rates according to the Ryan tumor regression grading system in rectal cancer cases operated after neoadjuvant therapy and to investigate the relationship between response levels and clinical and pathological parameters.

MATERIALS AND METHODS

Patients diagnosed with rectal cancer between January 2020 and December 2025 were retrospectively reviewed in our study. Patients who did not receive neoadjuvant therapy were excluded from the study. Ethical approval was obtained from the Abant Izzet Baysal University Clinical Research Ethics Committee (2026/26).

Patient characteristics including gender, age, clinical tumor stage before neoadjuvant therapy, and tumor location in the rectum (upper, middle, lower) were recorded. Postoperative pathological parameters such as total number of excised lymph nodes, number of metastatic lymph nodes, tumor size, lymphovascular invasion, and perineural invasion were determined.

In the histopathological evaluation of the patients, the Ryan tumor regression grading system was used to assess the response to neoadjuvant therapy. Treatment response was classified as follows: Score 0 (no viable cancer cells, only fibrosis; complete response), Score 1 (single or small groups of cancer cells; near-complete response), Score 2 (residual cancer with evident tumor regression but more than single cells or rare small groups; partial response), and Score 3 (extensive residual cancer with no evident tumor regression; poor or no response) (Figure 1) (3,4). The tumor treatment response was evaluated using the Ryan score in accordance with the protocols of the College of American Pathologists. Acellular mucin pools observed in specimens following neoadjuvant therapy were interpreted as indicative of complete tumor regression and were excluded from pT staging and from the assessment of lymph node positivity.

In cases demonstrating limited treatment response, the largest measurable dimension of the tumor was recorded as the tumor size. In cases with a more pronounced treatment response, characterized by prominent central regression and a scattered distribution of residual tumor cell clusters, the greatest dimension of the largest residual tumor focus was recorded as the tumor size.

Statistical analysis of the data was performed using IBM SPSS (Statistical Package for the Social Sciences) for Windows, Version 20.0. The distribution of continuous variables was assessed using the Kolmogorov–Smirnov test. Data with normal distribution were expressed as mean ± standard deviation, while non-normally distributed data were expressed as median (minimum–maximum). For comparisons between two groups, Student's t-test was used for normally distributed variables, and for comparisons among three or more groups, one-way analysis of variance (ANOVA) was applied. For non-normally distributed data, the Kruskal–Wallis test was used for multiple group comparisons. Categorical variables were compared using the chi-square test and were expressed as number (n) and percentage (%). Relationships between variables were evaluated using Pearson correlation analysis, and Spearman rank correlation analysis was used when the assumption of normal distribution was not met. A p-value of <0.05 was considered statistically significant.

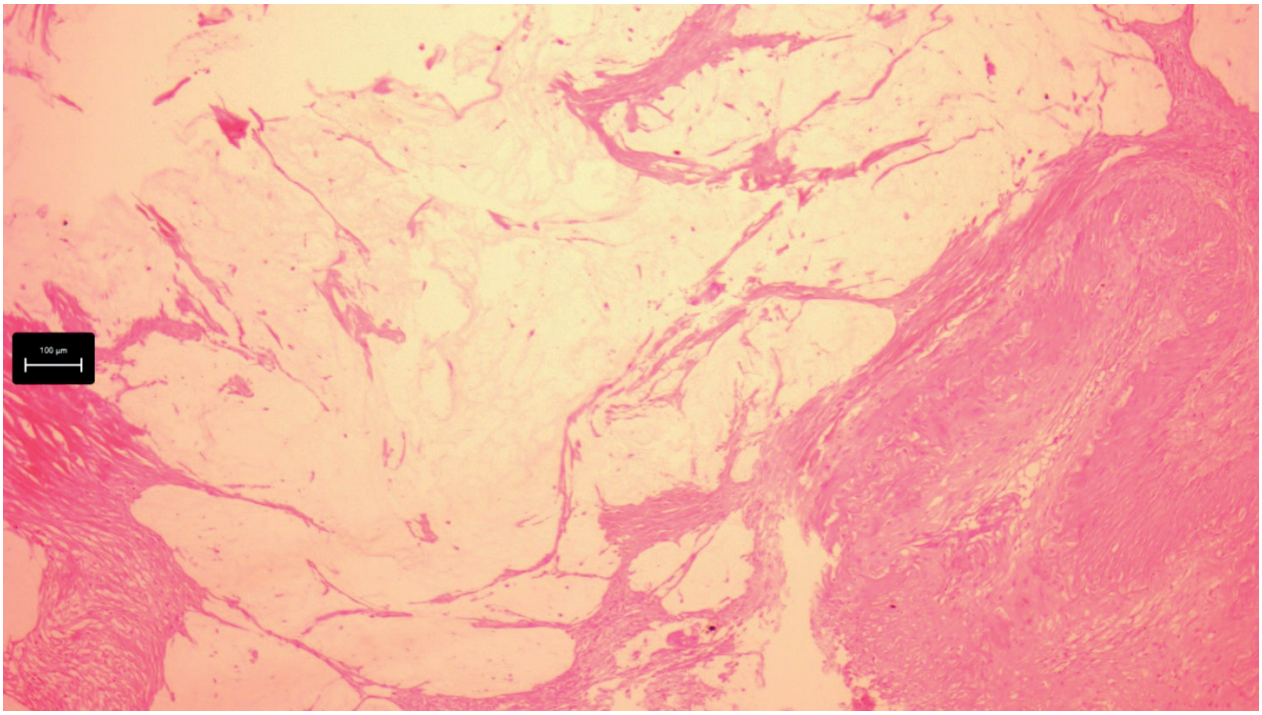


Figure 1. Ryan score 0 , full response with extensive hyalinization and mucin pools hematoxylin eosin x50.

RESULTS

A total of 58 patients were included in the study. Of these, 41 (71%) were male and 17 (29%) were female. No statistically significant difference was found between gender distribution and Ryan score groups (p=0.897). When gender distribution was evaluated according to Ryan score groups, there were 4 males and 2 females in the Score 0 group, 8 males and 4 females in the Score 1 group, 10 males and 5 females in the Score 2 group, and 19 males and 6 females in the Score 3 group.

The mean ages according to Ryan score groups were 64.3±11.6 years for Score 0, 67.5±10.7 years for Score 1, 63.7±10.9 years for Score 2, and 63.1±10.1 years for Score 3. No statistically significant difference was observed between the groups in terms of age (p=0.689).

The distribution of pathological response to neoadjuvant therapy was as follows: 10.3% (n=6) in Score 0, 20.7% (n=12) in Score 1, 25.9% (n=15) in Score 2, and 43.1% (n=25) in Score 3.

The distribution of residual tumor size according to Ryan score was as follows: 0 mm in Score 0, 7.5±3.6 mm in Score 1, 22.1±11.2 mm in Score 2, and 30.4±22.8 mm in Score 3. As the pathological response decreased, residual tumor size increased significantly, and this finding was statistically significant (p<0.001).

The mean total number of retrieved lymph nodes was 6.7±3.0 in Score 0, 10.6±6.3 in Score 1, 7.9±5.4 in Score 2, and 11.0±6.2 in Score 3. No statistically significant difference was found between the groups (p=0.228). Similarly, no statistically significant difference was found between Ryan score groups in terms of the number of metastatic lymph nodes (p=0.442).

Regarding T stage distribution, 5.2% of patients were T1, 32.8% were T2, 44.8% were T3, and 17.2% were T4. When T stage distribution was evaluated according to Ryan score groups, there were 0 T1, 2 T2, 3 T3, and 1 T4 patients in the Score 0 group; 0 T1, 5 T2, 6 T3, and 1 T4 patients in the Score 1 group; 1 T1, 4 T2, 6 T3, and 4 T4 patients in the Score 2 group; and 2 T1, 8 T2, 11 T3, and 4 T4 patients in the Score 3 group. No statistically

Table 1. Distribution of patients' demographic and clinical characteristics according to RYAN score

	Score 0 (n=6)	Score 1 (n=12)	Score 2 (n=15)	Score 3 (n=25)	P value
Age(years)	64.3 ± 11.6	67.5 ± 10.7	63.7 ± 10.9	63.1 ± 10.1	0.689
Gender (male/female)	4 / 2	8 / 4	10 / 5	19 / 6	0.897
Tumor size (mm)	0	7.5 ± 3.6	22.1 ± 11.2	30.4 ± 22.8	<0.001
Total lymph node count	6.7 ± 3.0	10.6 ± 6.3	7.9 ± 5.4	11.0 ± 6.2	0.228
Metastatic lymph node	0	0	0	0.24 ± 0.83	0.442
T Stage (T1/T2/T3/T4)	0 / 2 / 3 / 1	0 / 5 / 6 / 1	1 / 4 / 6 / 4	2 / 8 / 11 / 4	0.947
Lymphovascular invasion (n, %)	1 (%16.7)	0	0	1 (%4.0)	0.247
Perineural invasion (n, %)	0	1 (%8.3)	1 (%6.7)	1 (%4.0)	0.692
Tumor localization (Lower/Middle/Upper)	2 / 2 / 2	4 / 4 / 4	5 / 5 / 5	9 / 10 / 6	0.659

significant relationship was found between T stage and Ryan score (p=0.947).

Lymphovascular invasion was detected in 3.4% (n=2) of patients, one in the Score 0 group and one in the Score 3 group. No statistically significant relationship was found between lymphovascular invasion and Ryan score (p=0.247).

Perineural invasion was detected in 5.2% (n=3) of patients, with one patient in each of the Score 1, Score 2, and Score 3 groups. No statistically significant relationship was found between perineural invasion and Ryan score (p=0.692).

In terms of tumor localization, 34.5% of patients had tumors located in the lower rectum, 36.2% in the middle rectum, and 29.3% in the upper rectum. No statistically significant relationship was found between tumor localization and Ryan score (p=0.659) (Table 1).

DISCUSSION

In the present work, the relationship between the Ryan tumor regression score and clinical and pathological parameters was evaluated in patients with rectal cancer who underwent surgery after neoadjuvant chemoradiotherapy. The findings demonstrated a statistically significant association, particularly between tumor size and the Ryan score, while no significant associations were observed with other clinicopathological parameters.

The evaluation of pathological response after neoadjuvant therapy is considered an important prognostic indicator in rectal cancer. In the current literature, the Ryan tumor regression grading system has been shown to be a standardized and reliable method, with lower score (0–1) associated with better disease-free and overall survival (8,9). In our study, approximately 30% of patients had Ryan score between 0 and 1, supporting the notion that tumor regression is a direct indicator of treatment efficacy.

A significant relationship was observed between tumor size and the Ryan score. As the pathological response decreased, tumor size increased. Similarly, previous studies have reported that patients who respond well to neoadjuvant therapy show a marked reduction in tumor size, which is associated with improved prognosis (10-12). In the present study, tumor size was approximately 30 mm in 43% of patients, suggesting that tumor size is not only a morphological parameter but also an indicator of treatment response. On the other hand, no significant relationship was found between the Ryan score and age, gender, total number of lymph nodes, or number of metastatic lymph nodes. The literature presents conflicting results regarding the relationship between these parameters and tumor regression. While some studies have reported an association between lymph node involvement and pathological response, others have found no significant relationship (13,14). This discrepancy may be explained by the heterogeneous nature of tumor biology.

It is noteworthy that no significant relationship was found between T stage and the Ryan score. Although lower response rates are generally expected in tumors with advanced T stage, the response to neoadjuvant therapy may not be explained solely by anatomical staging. Recent studies emphasize that molecular characteristics and microscopic features of the tumor play a significant role in treatment response (15,16). A major challenge is that advanced T stage is often associated with a poorer response to neoadjuvant therapy and a more advanced clinical presentation. In such cases, surgeons may face more complex procedures, and patients may have a worse prognosis.

No significant relationship was found between the Ryan score and prognostic factors such as lymphovascular invasion and perineural invasion. The most likely explanation for this finding is the low incidence of these parameters and the limited sample size in our study. Larger series have demonstrated significant associations between these factors and poor prognosis (17,18). Similarly, no significant relationship was observed between tumor localization and pathological response. The literature indicates that tumors located in different segments of the rectum do not show significant differences in response to neoadjuvant therapy (19). This finding supports the notion that tumor location alone has a limited role in determining treatment response.

In the present study, a significant positive correlation was also identified between the Ryan score and tumor size in both Pearson and Spearman analyses, indicating that the results are statistically robust and consistent. Likewise, previous studies have reported significant correlations between tumor regression grade and tumor size (8,10,20).

This work has some limitations. First, its retrospective, single-center design limits the generalizability of the findings. Additionally, the relatively small sample size reduced the statistical power of some subgroup analyses. However, the homogeneity of the patient population and the use of standardized pathological evaluation are important strengths of the study.

In conclusion, a significant relationship was found between the Ryan tumor regression score and tumor size in patients with rectal cancer who underwent surgery after neoadjuvant therapy. The Ryan score is a practical and reliable method for evaluating treatment response and may be useful in prognostic assessment. These findings should be confirmed by larger, prospective studies.

Ethical approval

This study has been approved by the Bolu Abant Izzet Baysal University Non-Interventional Clinical Research Ethics Committee (approval date 20/01/2026, number 2026/26). Written informed consent was obtained from the participants.

Author contribution

Concept: SPO, BO; Design: SPO, BO; Data Collection or Processing: SPO, BO; Analysis or Interpretation: SPO, BO; Literature Search: SPO, BO; Writing: SPO, BO. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Ruo L, Tickoo S, Klimstra DS, et al. Long-term prognostic significance of extent of rectal cancer response to preoperative radiation and chemotherapy. Ann Surg. 2002; 236(1): 75-81. **[Crossref]**
2. Gavioli M, Luppi G, Losi L, et al. Incidence and clinical impact of sterilized disease and minimal residual disease after preoperative radiochemotherapy for rectal cancer. Dis Colon Rectum. 2005; 48(10): 1851-7. **[Crossref]**
3. Ryan R, Gibbons D, Hyland JMP, et al. Pathological response following long-course neoadjuvant chemoradiotherapy for locally advanced rectal cancer. Histopathology. 2005; 47(2): 141-6. **[Crossref]**

4. Lo DS, Pollett A, Siu LL, Gallinger S, Burkes RL. Prognostic significance of mesenteric tumor nodules in patients with stage III colorectal cancer. Cancer. 2008; 112(1): 50-4. **[Crossref]**
5. Birbeck KF, Macklin CP, Tiffin NJ, et al. Rates of circumferential resection margin involvement vary between surgeons and predict outcomes in rectal cancer surgery. Ann Surg. 2002; 235(4): 449-57. **[Crossref]**
6. Nagtegaal ID, Marijnen CAM, Kranenbarg EK, et al. Circumferential margin involvement is still an important predictor of local recurrence in rectal carcinoma: not one millimeter but two millimeters is the limit. Am J Surg Pathol. 2002; 26(3): 350-7. **[Crossref]**
7. Puppa G, Maisonneuve P, Sonzogni A, et al. Pathological assessment of pericolonic tumor deposits in advanced colonic carcinoma: relevance to prognosis and tumor staging. Mod Pathol. 2007; 20(8): 843-55. **[Crossref]**
8. Kim SH, Chang HJ, Kim DY, et al. What is the ideal tumor regression grading system in rectal cancer patients after preoperative chemoradiotherapy? Cancer Res Treat. 2016; 48(3): 998-1009. **[Crossref]**
9. Fokas E, Schlenska-Lange A, Polat B, Klautke G, Grabenbauer GG, Fietkau R, et al. Chemoradiotherapy plus induction or consolidation chemotherapy as total neoadjuvant therapy for patients with locally advanced rectal cancer: long-term results of the CAO/ARO/AIO-12 randomized clinical trial. JAMA Oncology. 2022; 8(1): e215445. **[Crossref]**
10. Seo N, Kim H, Cho MS, Lim JS. Response assessment with MRI after chemoradiotherapy in rectal cancer: current evidences. Korean J Radiol. 2019; 20(7): 1003-18. **[Crossref]**
11. Trakarnsanga A, Gönen M, Shia J, et al. Comparison of tumor regression grade systems for locally advanced rectal cancer after multimodality treatment. J Natl Cancer Inst. 2014; 106(10): dju248. **[Crossref]**
12. Petrelli F, Trevisan F, Cabiddu M, et al. total neoadjuvant therapy in rectal cancer: a systematic review and meta-analysis of treatment outcomes. Ann Surg. 2020; 271(3): 440-8. **[Crossref]**
13. Maas M, Nelemans PJ, Valentini V, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. Lancet Oncol. 2010; 11(9): 835-44. **[Crossref]**
14. Rödel C, Martus P, Papadoupolos T, et al. Prognostic significance of tumor regression after preoperative chemoradiotherapy for rectal cancer. J Clin Oncol. 2005; 23(34): 8688-96. **[Crossref]**
15. Bahadoer RR, Dijkstra EA, van Etten B, et al. Short-course radiotherapy followed by chemotherapy before Total Mesorectal Excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. Lancet Oncol. 2021; 22(1): 29-42. **[Crossref]**
16. Lee CI, Kanata D, Garcia-Aguilar J. Can we predict which patients with rectal cancer will have a complete response to neoadjuvant therapy? Clinics in Colon and Rectal Surgery. 2026. **[Crossref]**
17. Wang G, Wang JJ, Guan R, et al. Targeting strategies for glucose metabolic pathways and t cells in colorectal cancer. Curr Cancer Drug Targets. 2019; 19(7): 534-50. **[Crossref]**
18. Smith JD, Ruby JA, Goodman KA, et al. Nonoperative management of rectal cancer with complete clinical response after neoadjuvant therapy. Ann Surg. 2012; 256(6): 965-72. **[Crossref]**
19. Xie JW, Lu J, Xu BB, et al. Prognostic value of tumor regression grading in patients treated with neoadjuvant chemotherapy plus surgery for gastric cancer. Front Oncol. 2021; 11: 587856. **[Crossref]**
20. Luo D, Chen Y, Luo Z, Gong H, Li Q, Li X. Prognostic value of tumor regression grade combined with pathological lymph node status in initially node-positive rectal cancer treated with neoadjuvant chemoradiotherapy. J Invest Surg. 2025; 38(1): 2449669. **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):242-253
DOI: 10.54307/NWMJ.2026.252

RESEARCH ARTICLE

Evaluation of pentraxin-3 levels as a biomarker of large vessel occlusion in ischemic stroke

Hilmiye Tokmak[1], Ahmet Yabalak[2], Ayşegül Özyılmaz[2], Halime Şahan[2], Merve Alpay[3]

[1]Department of Neurology, Niğde Training and Research Hospital, Niğde, Türkiye
[2]Department of Neurology, Faculty of Medicine, Düzce University, Düzce, Türkiye
[3]Department of Biochemistry, Faculty of Medicine, Düzce University, Düzce, Türkiye

Cite as: Tokmak H, Yabalak A, Özyılmaz A, Şahan H, Alpay M. Evaluation of pentraxin-3 levels as a biomarker of large vessel occlusion in ıschemic stroke. Northwestern Med J. 2026;6(3):242-253.

ABSTRACT

Objective: Early diagnosis of acute ischemic stroke patients with large vessel occlusion (LVO) and the transfer of patients who require endovascular treatment to specialized stroke centers are critical for patients. Pentraxin-3 (PTX-3) is a biomarker of vascular inflammation, and its role in stroke, including its relationship with clinical features, is being investigated. In this study, we evaluated the value of PTX-3 levels in the diagnosis of LVO and their relationship with clinical features.

Methods: This study included retrospectively and prospectively enrolled patients diagnosed with acute ischemic stroke. Patients were divided into two groups according to imaging findings: those with and those without LVO. Serum PTX-3 levels were measured and compared between the groups. Stroke severity was evaluated at admission using the National Institutes of Health Stroke Scale (NIHSS), and clinical outcome was evaluated at the 90th day using the modified Rankin Scale (mRS). Correlation, ROC analysis, and logistic regression methods were used in statistical analyses.

Results: Data from 81 patients were analyzed. Although PTX-3 levels were higher in patients with LVO, the difference was not statistically significant. The discriminative power of PTX-3 for LVO was limited (AUC = 0.58). A positive correlation was found between PTX-3 levels and admission NIHSS. Although an association between PTX-3 and poor clinical outcome was observed in univariate analyses, PTX-3 was not an independent marker in multivariable analyses. The addition of PTX-3 to clinical models did not increase predictive power.

Conclusion: In patients with acute ischemic stroke with LVO, PTX-3 was not a sufficient biomarker for confirming the diagnosis. Although PTX-3 was associated with stroke severity, it could not predict clinical outcome independently. Based on the current findings, PTX-3 does not appear to have a clear contribution to the clinical decision-making process.

Keywords: acute ischemic stroke, clinical outcome, large vessel occlusion, pentraxin-3

Corresponding author: Hilmiye Tokmak **E-mail:** hilmiyetokmak@gmail.com
Received: 20.01.2026 **Accepted:** 24.04.2025 **Published:** 24.04.2026

INTRODUCTION

Ischemic stroke, which poses a significant burden to public health, has entered a new era in treatment since the effectiveness of endovascular thrombectomy was demonstrated in 2015. The demonstrated effectiveness of stroke treatments, together with the recognition that this effect is time-dependent, has revealed that neuronal loss continues with every minute lost (1,2). Therefore, many countries worldwide are establishing centers dedicated to stroke treatment according to existing healthcare resources and are developing organizational schemes aimed at shortening the time to reach these centers and inter-center transfer durations.

Stroke centers are comprehensive units where both intravenous thrombolytic (IVT) and endovascular treatment (EVT) can be performed. These treatments are applied to patients with large vessel occlusion (LVO), and the success of the treatment is time-dependent. Therefore, it is important that patients with suspected ischemic stroke are rapidly evaluated and transferred to comprehensive centers. Especially in rural areas where advanced imaging methods such as CT and MR are not available, supporting the diagnosis with a measurable blood biomarker is essential. However, to date, a reliable biomarker that can predict LVO has not been identified (3).

Inflammatory markers begin to be produced in the occluded or hypoperfused vessel and brain tissue after acute ischemic stroke (4). Pentraxin-3 (PTX-3) is a member of the long pentraxin family (5). Although classical acute phase reactants are produced by the liver, the PTX-3 molecule is produced in various cells such as endothelial cells, macrophages, dendritic cells, and vascular smooth muscle cells (6). In an animal study, PTX-3 was demonstrated to support cerebral blood flow, angiogenesis, and neuronal viability (7). Although it has been identified as a neuroprotective molecule in animal experiments, whether it is associated with prognosis in patients with ischemic stroke has not been fully established. In a study conducted in stroke patients, Pentraxin-3 levels were reported to be associated with poor prognosis in minor strokes on the basis of large vessel atherosclerosis (8,9), whereas another study found no association with prognosis (10). In the literature, no study evaluating the relationship between the presence of LVO and PTX-3 levels in patients with ischemic stroke has been identified.

In patients with ischemic stroke and LVO, early recognition for EVT and rapid transfer to appropriate centers are necessary. In this context, the identification of a blood-measurable biomarker may offer new opportunities for clinical practice, especially in patient groups where advanced imaging modalities are not available or imaging cannot be performed. In this study, we aimed to evaluate the relationship between PTX-3 levels by comparing patients with and without LVO. As a secondary outcome, we planned to evaluate its association with the clinical outcomes of the patients.

MATERIALS AND METHODS

This study was designed with both retrospective and prospective components. Patients who had blood samples collected within the scope of the thesis study titled "Comparison of Matrix Metalloproteinase-28 Levels According to Etiology in Patients with Ischemic Stroke," conducted between June 2024 and October 2025, and who met the inclusion and exclusion criteria of the present study were included. If the target sample size was not reached, patients who presented to our clinic until January 2026 and met the study criteria were prospectively enrolled to complete the sample.

We obtained general medical histories of patients admitted to our hospital with acute ischemic stroke and performed neurological examinations. We recorded age, sex, admission examination findings, vascular risk factors, and post-stroke echocardiography (ECHO) and electrocardiography (ECG) findings of the cases. The presence of LVO in patients was recorded. The treatments applied to the patients were recorded. Blood tests were obtained before administration if reperfusion therapy (EVT or IVT) was to be given. Stroke severity was assessed at admission using the National Institutes of Health Stroke Scale (NIHSS). Clinical outcome was evaluated at the third month using the modified Rankin Scale (mRS).

Patients who met the study criteria were divided into two groups according to imaging findings: patients

with occlusion of the internal carotid artery (ICA) or the M1 segment of the middle cerebral artery (MCA) were classified as the large vessel occlusion group; patients without occlusion in these segments were classified as the control group.

Hemogram, ferritin, C-reactive protein (CRP), troponin I or T, urea, creatinine, low-density lipoprotein (LDL), and triglyceride levels obtained at emergency department admission were evaluated. Blood samples for PTX-3 were obtained at the time of admission to the emergency department. The samples were kept at room temperature for 2 hours and then centrifuged at 1000 rpm for 10 minutes to obtain serum, which was stored at −80°C. At the end of the study period, the samples were analyzed in the biochemistry laboratory using the A.B.T. Human PTX-3/TGS-14 (Pentraxin 3) ELISA kit (Cat. No# ABT3185Hu, Atlas Biotechnology, Ankara, Turkey).

Statistical analysis

Distribution of continuous variables was assessed with the Shapiro–Wilk test and graphical inspection. Normally distributed continuous variables were reported as mean ± standard deviation and compared with the Student's t-test (or Welch's t-test when variance homogeneity, assessed by Levene's test, was doubtful); non-normally distributed continuous and ordinal variables were reported as median with interquartile range and compared with the Mann–Whitney U test. Categorical variables were compared with the Pearson χ^2 test or Fisher's exact test, as appropriate. Spearman's rank correlation coefficient was used to examine the association between serum PTX-3 and clinical and laboratory variables; 95% confidence intervals were derived by the Fisher z transformation. Discriminative performance of PTX-3 for each endpoint was quantified by the receiver-operating-characteristic area under the curve with bootstrap percentile 95% confidence intervals (2,000 resamples), and the Youden index was used to identify an optimal cut-off. Independent predictors of each endpoint were modeled by binary logistic regression. A parsimonious multivariable model was pre-specified for each endpoint, with predictors chosen for clinical relevance (primary exposure and core risk factors) rather than data-driven selection; variance-inflation factors, the Hosmer–Lemeshow goodness-of-fit test, and Nagelkerke's R^2 are reported. Sensitivity analyses (log transformation, Winsorization, outlier exclusion, median imputation, and sex- and age-stratified re-analysis) were pre-specified to probe the robustness of the primary inference. Two-sided $\alpha = .05$ was used throughout. Analyses were conducted in Python 3.11 with pandas 2.2, scipy 1.13, statsmodels 0.14, and scikit-learn 1.5.

RESULTS

Sample characteristics and between-group comparisons

The analysis sample comprised 81 patients with acute anterior-circulation ischemic stroke presenting within 24 hours of symptom onset: 40 patients with large-vessel occlusion (LVO+; internal carotid artery and/or middle-cerebral-artery M1 segment) and 41 without (LVO−). Baseline demographic, clinical, and laboratory characteristics are presented in Table 1. The groups were comparable in age (M = 68.6, SD = 13.0 years in LVO+ vs. M = 68.2, SD = 12.7 years in LVO−; p = .647) and in sex distribution (47.5% vs. 39.0% female; p = .585). Patients with LVO presented with substantially greater neurological deficit than those without, as indicated by a median NIHSS score on admission of 14 (IQR = 10–20) versus 4 (IQR = 2–6; p < .001; rank-biserial r = −0.79) and by a median baseline mRS of 5 (IQR = 4–5) versus 1 (IQR = 1–3; p < .001). Atrial fibrillation (AF) was markedly more prevalent in LVO+ patients than in LVO− patients, both on admission electrocardiography (60.0% vs. 7.3%; p < .001, $\phi = .53$) and on the composite ECG-or-Holter indicator (65.0% vs. 22.0%; p < .001, $\phi = .41$). Routine hematological and biochemical parameters — including hemoglobin, C-reactive protein, ferritin, troponin I, urea, and creatinine — did not differ between the groups (all p ≥ .34). Unfavorable 90-day functional outcome (mRS > 2) occurred in 62.5% of LVO+ patients compared with 7.3% of LVO− patients (p < .001, $\phi = .55$).

Serum pentraxin-3 concentrations by LVO status

Admission serum PTX-3 concentrations were compatible with a normal distribution in both groups (Shapiro–Wilk W = 0.98, p = .747 in LVO+; W = 0.95,

Table 1. Baseline demographic, clinical, and laboratory characteristics of the study groups

Variable	Total (N = 81)	LVO+ (n = 40)	LVO− (n = 41)	Test statistic	p
Age, years, Mdn (IQR)	69 (61–78)	69 (63–79)	68 (59–78)	U = 869.0	.647
Female sex, n (%)	35 (43.2)	19 (47.5)	16 (39.0)	$\chi^2(1) = 0.30$	.585
AF on admission ECG, n (%)	27 (33.3)	24 (60.0)	3 (7.3)	$\chi^2(1) = 22.97$	< .001
AF (ECG or Holter), n (%)	35 (43.2)	26 (65.0)	9 (22.0)	$\chi^2(1) = 13.59$	< .001
Holter-detected AF, n (%)	8 (9.9)	2 (5.0)	6 (14.6)	†	.264
Positive echo finding, n (%)	75 (96.2)	36 (92.3)	39 (100.0)	†	.240
NIHSS on admission, Mdn (IQR)	7 (4–16)	14 (10–20)	4 (2–6)	U = 1471.0	< .001
Baseline mRS, Mdn (IQR)	3 (1–5)	5 (4–5)	1 (1–3)	U = 1456.0	< .001
Hemoglobin, g/dL, M (SD)	12.97 (1.72)	12.80 (1.99)	13.13 (1.42)	t(79) = −0.87	.388
CRP, mg/dL, Mdn (IQR)	0.31 (0.17–0.64)	0.29 (0.15–0.75)	0.37 (0.21–0.57)	U = 748.0	.499
Ferritin, ng/mL, Mdn (IQR)	73.7 (48.2–155.0)	72.6 (39.4–155.0)	82.8 (54.2–208.0)	U = 524.5	.361
Troponin I, ng/L, Mdn (IQR)	8.4 (4.0–14.9)	9.0 (5.3–15.5)	7.6 (3.3–13.9)	U = 880.0	.340
Urea, mg/dL, Mdn (IQR)	35.3 (26.5–51.0)	36.1 (26.2–55.0)	36.4 (27.0–50.0)	U = 895.0	.402
Creatinine, mg/dL, Mdn (IQR)	0.94 (0.74–1.08)	0.89 (0.70–1.12)	0.90 (0.77–1.05)	U = 843.5	.608
Serum PTX-3, pg/mL, M (SD)	337.2 (141.7)	355.6 (137.1)	319.2 (145.5)	t(79) = 1.16	.249
90-day mRS, Mdn (IQR)	1 (1–4)	4 (1.75–6)	1 (1–1)	U = 1335.0	< .001
Unfavourable outcome, n (%)	28 (34.6)	25 (62.5)	3 (7.3)	$\chi^2(1) = 24.87$	< .001

LVO = large vessel occlusion; M = mean; SD = standard deviation; Mdn = median; IQR = interquartile range (Q1–Q3); AF = atrial fibrillation; ECG = electrocardiogram; NIHSS = National Institutes of Health Stroke Scale; mRS = modified Rankin Scale; PTX-3 = pentraxin 3; CRP = C-reactive protein; U = Mann–Whitney U statistic; t = Student t statistic; χ^2 = Pearson chi-square statistic. Normally distributed continuous variables (hemoglobin, serum PTX-3; assessed by Shapiro–Wilk test within each group) were compared with the Student t test. Non-normally distributed continuous and ordinal variables were compared with the Mann–Whitney U test. Categorical variables were compared with the Pearson χ^2 test. † Fisher's exact test (expected cell count < 5). Denominators for percentages use the non-missing count per variable (range 78–81).

p = .054 in LVO−), and variance homogeneity was supported by Levene's test (F(1,79) = 0.12, p = .735). Mean serum PTX-3 was numerically higher in the LVO+ group (M = 355.64, SD = 137.09 pg/mL) than in the LVO− group (M = 319.16, SD = 145.46 pg/mL), but the between-group difference did not reach statistical significance (t(79)= 1.16, p = .249, Cohen's d = 0.26; 95% CI of mean difference [−26.00, 98.97] pg/mL). The distribution of PTX-3 values by group is illustrated in Figure 1. Welch's t-test produced an essentially identical result (t(78.91) = 1.16, p = .249), and a Mann–Whitney U test confirmed the non-significant pattern (U = 914.5, p = .378, rank-biserial r = 0.12; Hodges–Lehmann location shift = 25.2 pg/mL, bootstrap 95% CI [−45.1, 94.3]). The post-hoc power achieved for the observed effect size (d = 0.26) was approximately .21. Additional robustness checks (log transformation, Winsorization, outlier exclusion, and stratified re-analysis) uniformly confirmed the null finding.

Correlations of serum pentraxin-3 with clinical variables

Spearman rank-order correlations among serum PTX-3 and the principal clinical and laboratory variables are presented in Table 2. Serum PTX-3 was significantly and positively correlated with admission NIHSS (ρ = .29, 95% CI [.08, .48], p = .009), indicating that higher PTX-3 concentrations accompanied greater neurological deficit at presentation; the bivariate relationship is illustrated in Figure 2. Correlations of PTX-3 with age (ρ = .20, p = .077), baseline mRS (ρ = .20, p = .077), serum creatinine (ρ = .22, p = .055), and 90-day mRS (ρ = .18, p = .115) were of borderline significance and small in magnitude. Serum PTX-3 was

not significantly correlated with CRP (ρ = .17, p = .132), consistent with the view that the two acute-phase reactants reflect distinct biological compartments. As expected, admission NIHSS was strongly correlated with baseline mRS (ρ = .91, p < .001) and with 90-day mRS (ρ = .72, p < .001).

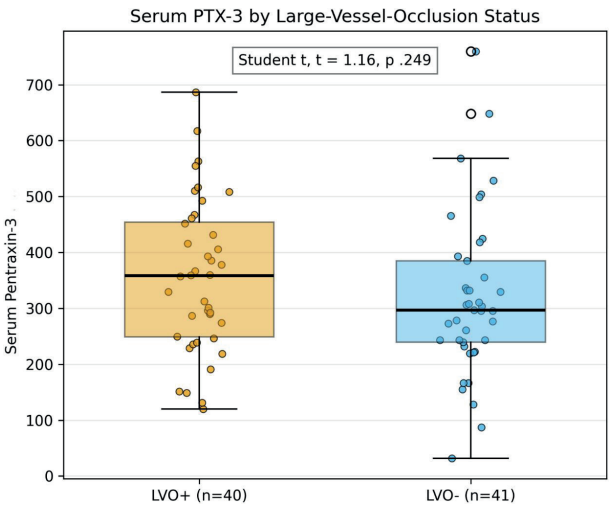


Figure 1. Serum pentraxin-3 concentrations in patients with and without large-vessel occlusion.

Box-and-whisker plot of admission serum PTX-3 in the LVO+ (n = 40) and LVO− (n = 41) groups. Boxes = interquartile range; horizontal line = median; whiskers = 1.5 × IQR; individual observations overlaid with lateral jitter. t(79) = 1.16, p = .249, Cohen's d = 0.26, 95% CI of mean difference [−26.00, 98.97] pg/mL.

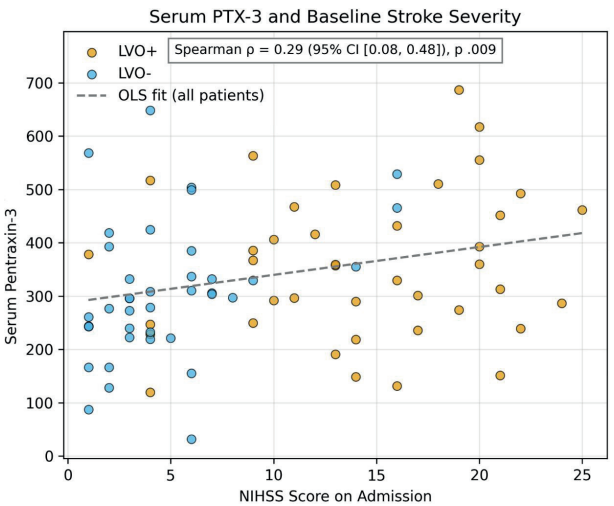


Figure 2. Scatter plot of serum pentraxin-3 against admission NIHSS, by LVO status.

Observations are colour-coded by LVO status. The dashed grey line is an ordinary-least-squares fit (visual reference only); the reported statistic is Spearman's ρ = .29, 95% CI [.08, .48], p = .009.

Table 2. Means, standard deviations, and spearman correlations among study variables

Variable	M	SD	1	2	3	4	5	6	7	8
1. Serum PTX-3 (pg/mL)	337.18	141.70	—							
2. NIHSS on admission	9.53	6.95	.29**	—						
3. Baseline mRS	2.98	1.60	.20	.91***	—					
4. 90-day mRS	2.49	1.94	.18	.72***	.81***	—				
5. Age (years)	68.40	12.76	.20	.17	.14	.27*	—			
6. CRP (mg/dL)	0.86	1.63	.17	.04	-.03	.06	.03	—		
7. Hemoglobin (g/dL)	12.97	1.72	-.08	-.13	-.12	-.18	-.43***	-.16	—	
8. Creatinine (mg/dL)	1.03	0.41	.22	.12	.09	.18	.07	.07	.03	—

N = 69–81 per correlation (pairwise deletion of missing values). M and SD are shown for all variables to facilitate comparison; note that CRP, creatinine, and NIHSS are non-normally distributed. Values in the lower triangle are Spearman's rank-order correlation coefficients (ρ). PTX-3 = pentraxin 3; NIHSS = National Institutes of Health Stroke Scale; mRS = modified Rankin Scale; CRP = C-reactive protein.

* p < .05. ** p < .01. *** p < .001.

Discriminative performance of serum pentraxin-3

The discriminative performance of serum PTX-3 and benchmark comparators is summarized in Table 3 and illustrated in Figure 3. For discrimination of large-vessel occlusion (Figure 3A), serum PTX-3 yielded an AUC of .58 (95% CI [.45, .71]); the Youden-optimal cut-off of 357 ng/mL provided a sensitivity of 52.5% and a specificity of 73.2%, with positive and negative predictive values of 65.6% and 61.2%. The 95% CI of the AUC crossed the chance-level value of 0.50. By contrast, admission NIHSS yielded an AUC of .90 (95% CI [.82, .96]). Adding PTX-3 to the full clinical model (age, sex, NIHSS, AF) did not improve discrimination (ΔAUC = +.001; LR $\chi^2(1) = 0.51$, p = .474). For 90-day unfavorable outcome (Figure 3B), PTX-3 produced an AUC of .66 (95% CI [.51, .79]), a value only marginally above the chance level of .50 at the lower confidence bound, indicating limited and statistically uncertain discriminative ability for this endpoint. Adding PTX-3 to the full clinical model (age, sex, NIHSS, LVO) similarly provided negligible incremental value (ΔAUC = +.003; LR $\chi^2(1) = 1.13$, p = .289).

Univariate and multivariable logistic regression for large-vessel occlusion

Univariate and prespecified hierarchical multivariable logistic-regression results for the LVO endpoint are presented in Table 4. In univariate analyses, the strongest predictors of LVO were admission NIHSS (OR = 1.39, 95% CI [1.22, 1.60], p < .001), baseline mRS (OR = 3.47, 95% CI [2.15, 5.59], p < .001), and atrial fibrillation (OR = 6.81, 95% CI [2.55, 18.19], p < .001). Serum PTX-3 was not significantly associated with LVO at the univariate level (OR = 1.002, 95% CI [0.999, 1.005], p = .248). Age, sex, and all laboratory markers were non-significant (all p ≥ .34).

Table 3. Receiver-operating-characteristic analysis: discriminative performance of single markers and clinical models for the two endpoints

Endpoint / Model	n	AUC [95% CI]	Sensitivity, %	Specificity, %	ΔAUC	LR $\chi^2(1)$, p
Large-vessel occlusion						
Serum PTX-3	81	.58 [.45, .71]	52.5	73.2	—	—
CRP	81	.46 [.33, .59]	—	—	—	—
NIHSS	81	.90 [.82, .96]	—	—	—	—
Clinical model [a]	81	.92 [.85, .97]	—	—	ref.	—
Clinical model + PTX-3	81	.92 [.85, .97]	—	—	+.001	0.51, .474
90-day unfavourable outcome						
Serum PTX-3	81	.66 [.51, .79]	60.7	71.7	—	—
NIHSS	81	.91 [.83, .97]	—	—	—	—
LVO status	81	.81 [.72, .89]	—	—	—	—
Clinical model [b]	81	.94 [.88, .99]	—	—	ref.	—
Clinical model + PTX-3	81	.95 [.89, .99]	—	—	+.003	1.13, .289

AUC = area under the receiver-operating-characteristic curve; CI = confidence interval; ΔAUC = change in AUC when PTX-3 is added to the clinical model; LR = likelihood-ratio test comparing the clinical model with vs. without PTX-3; PTX-3 = pentraxin 3; CRP = C-reactive protein; NIHSS = National Institutes of Health Stroke Scale; AF = atrial fibrillation; LVO = large-vessel occlusion. Sensitivity and specificity are reported at the Youden-optimal cut-off of 357 pg/mL for PTX-3 only (PPV = 65.6%, NPV = 61.2% for LVO; PPV = 53.1%, NPV = 77.6% for 90-day outcome). Composite-model AUCs were derived from the predicted probabilities of the corresponding logistic-regression equations. Bootstrap percentile CIs (2,000 replicates). [a] Clinical model for LVO: age, sex, NIHSS, and AF. [b] Clinical model for 90-day outcome: age, sex, NIHSS, and LVO status.

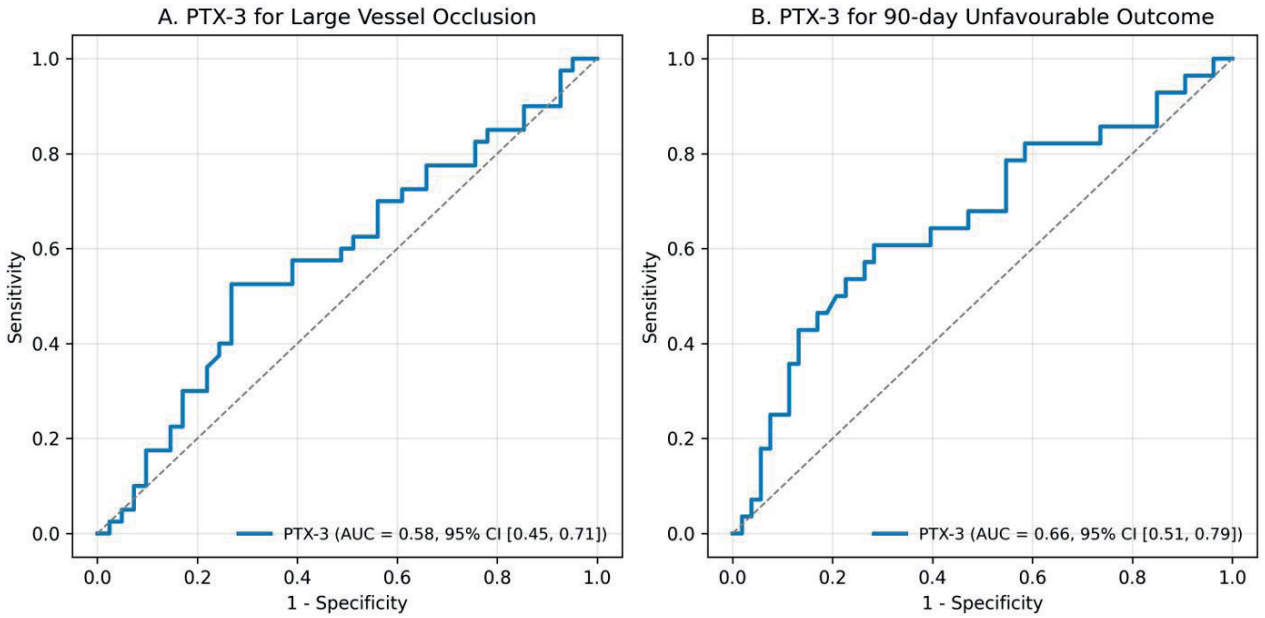


Figure 3. Receiver-operating-characteristic curves for serum pPentraxin-3.

Panel A: serum PTX-3 for discrimination of large-vessel occlusion (AUC = .58, 95% CI [.45, .71]). Panel B: serum PTX-3 for 90-day unfavourable outcome (AUC = .66, 95% CI [.51, .79]). Dashed 45° line = chance-level discrimination. Bootstrap percentile CIs (2,000 replicates).

The hierarchical multivariable sequence (Table 4, Models 1–4) demonstrates that the addition of admission NIHSS to the demographic block (Model 2) increased Nagelkerke's R^2 from .010 to .625 and produced an adjusted OR for NIHSS of 1.44 (95% CI [1.24, 1.67], p < .001); further addition of atrial fibrillation (Model 3) and then serum PTX-3 (Model 4) yielded only marginal gains in R^2 (to .643 and .647, respectively). In the fully adjusted model (Model 4), only NIHSS retained independent significance (adjusted OR = 1.42, 95% CI [1.21, 1.65], p < .001). Although atrial fibrillation was a strong univariate predictor of LVO (OR = 6.60), its effect attenuated substantially after adjustment for admission NIHSS (Model 3: OR = 2.73, p = .157), consistent with shared variance between stroke severity and cardioembolic etiology rather than independent predictive value. PTX-3 was non-significant in the fully adjusted model (OR = 0.998, 95% CI [0.993, 1.003], p = .486). Variance-inflation factors ranged from 1.86 to 8.06 (all < 10). The model showed good discrimination (C-statistic = .92; optimism-corrected .89) and acceptable calibration (Hosmer–Lemeshow p = .722). A likelihood-ratio test confirmed that adding PTX-3 did not improve model fit (LR $\chi^2(1) = 0.51$, p = .474).

Univariate and multivariable logistic regression for 90-day unfavorable outcome

Twenty-eight of the 81 patients with complete 90-day follow-up (34.6%) experienced an unfavorable outcome. Univariate and prespecified hierarchical multivariable logistic-regression estimates are reported in Table 5. In univariate analyses, serum PTX-3 was weakly but significantly associated with unfavorable outcome (OR = 1.004, 95% CI [1.000, 1.008], p = .041), as were admission NIHSS (OR = 1.33, 95% CI [1.18, 1.49], p < .001), baseline mRS (OR = 2.64, 95% CI [1.78, 3.93], p < .001), LVO status (OR = 19.00, 95% CI [5.00, 72.19], p < .001), and atrial fibrillation (OR = 4.15, 95% CI [1.60, 10.78], p = .003). Age and renal indices showed non-significant trend-level associations (age p = .094; urea p = .069; creatinine p = .083).

Table 4. Univariate and hierarchical multivariable logistic regression for large-vessel occlusion

Predictor	Univariate	p	Model 1 (Demographics)	Model 2 (+ NIHSS)	Model 3 (+ AF)	Model 4 (+ PTX-3)
	OR [95% CI]		OR [95% CI]	OR [95% CI]	OR [95% CI]	OR [95% CI]
Age (per year)	1.003 [0.969, 1.038]	.886	1.000 [0.965, 1.036]	0.975 [0.928, 1.025]	0.968 [0.916, 1.023]	0.969 [0.918, 1.024]
Sex (female = 1)	1.414 [0.585, 3.417]	.442	1.416 [0.575, 3.488]	2.958 [0.742, 11.793]	2.795 [0.674, 11.588]	2.850 [0.681, 11.922]
NIHSS (per point)	1.393 [1.217, 1.595]	< .001	—	1.436 [1.238, 1.666]***	1.398 [1.204, 1.624]***	1.415 [1.211, 1.653]***
AF (present = 1)	6.603 [2.466, 17.672]	< .001	—	—	2.733 [0.676, 11.052]	2.860 [0.699, 11.705]
PTX-3 (per 1 pg/mL)	1.002 [0.999, 1.005]	.248	—	—	—	0.998 [0.993, 1.003]
Baseline mRS (per point)	3.466 [2.151, 5.585]	< .001	—	—	—	—
Hemoglobin (per g/dL)	0.892 [0.689, 1.154]	.384	—	—	—	—
CRP (per mg/dL)	1.043 [0.795, 1.368]	.762	—	—	—	—
Creatinine (per mg/dL)	1.329 [0.448, 3.942]	.608	—	—	—	—
Troponin I (per ng/L)	0.991 [0.974, 1.009]	.340	—	—	—	—
Urea (per mg/dL)	1.009 [0.988, 1.031]	.402	—	—	—	—
Ferritin (per ng/mL) [c]	0.999 [0.995, 1.002]	.436	—	—	—	—
Model summary						
Nagelkerke R^2			.010	.625	.643	.647
Hosmer–Lemeshow p			.221	.819	.839	.722
C-statistic			.58	.92	.92	.92

N = 81 (40 LVO+, 41 LVO−). Variables were entered in four prespecified hierarchical blocks: Model 1 = age, sex (demographics); Model 2 = + NIHSS (clinical stroke severity); Model 3 = + AF (etiological indicator); Model 4 = + serum PTX-3 (primary biomarker). The univariate column reports the crude OR per one-unit increase for each candidate predictor. Optimism-corrected C-statistic for Model 4 (1,000 bootstrap replicates) = .89. Variance-inflation factors in Model 4: age 8.06, sex 1.86, NIHSS 3.88, AF 2.36, PTX-3 6.96 (all < 10). LR test, Model 3 vs. Model 4: $\chi^2(1) = 0.51$, p = .474. OR = odds ratio; CI = confidence interval; AF = atrial fibrillation (ECG or Holter); NIHSS = National Institutes of Health Stroke Scale; mRS = modified Rankin Scale; CRP = C-reactive protein. [c] n = 69 for ferritin (13 missing). *** p < .001.

Table 5. Univariate and hierarchical multivariable logistic regression for 90-day unfavourable outcome (mRS > 2)

Predictor	Univariate	p	Model 1 (Demographics)	Model 2 (+ NIHSS)	Model 3 (+ LVO)	Model 4 (+ PTX-3)
	OR [95% CI]		OR [95% CI]	OR [95% CI]	OR [95% CI]	OR [95% CI]
Age (per year)	1.035 [0.994, 1.077]	.094	1.036 [0.994, 1.079]	1.047 [0.984, 1.114]	1.049 [0.985, 1.118]	1.046 [0.981, 1.115]
Sex (female = 1)	1.348 [0.543, 3.345]	.521	1.648 [0.636, 4.274]	4.617 [0.849, 25.120]	4.037 [0.692, 23.530]	3.617 [0.627, 20.850]
NIHSS (per point)	1.326 [1.179, 1.491]	< .001	—	1.451 [1.223, 1.721]***	1.361 [1.134, 1.634]**	1.340 [1.118, 1.605]**
LVO status (LVO+ = 1)	19.00 [5.00, 72.19]	< .001	—	—	3.550 [0.639, 19.724]	4.005 [0.691, 23.202]
PTX-3 (per 1 pg/mL)	1.004 [1.000, 1.008]	.041	—	—	—	1.003 [0.997, 1.008]
AF (present = 1)	4.148 [1.596, 10.779]	.003	—	—	—	—
Baseline mRS (per point)	2.645 [1.781, 3.930]	< .001	—	—	—	—
Hemoglobin (per g/dL)	0.866 [0.653, 1.150]	.322	—	—	—	—
CRP (per mg/dL)	1.147 [0.901, 1.461]	.265	—	—	—	—
Creatinine (per mg/dL)	2.927 [0.869, 9.864]	.083	—	—	—	—
Urea (per mg/dL)	1.023 [0.998, 1.047]	.069	—	—	—	—
Troponin I (per ng/L)	0.996 [0.981, 1.011]	.580	—	—	—	—
Ferritin (per ng/mL) [c]	1.001 [0.998, 1.004]	.510	—	—	—	—
Model summary						
Nagelkerke R^2			.080	.652	.671	.681
Hosmer–Lemeshow p			.719	.119	.513	.185
C-statistic			.64	.93	.94	.95

N = 81 (28 unfavourable-outcome events / 53 favourable). Variables were entered in four prespecified hierarchical blocks: Model 1 = age, sex (demographics); Model 2 = + NIHSS (clinical stroke severity); Model 3 = + LVO status (vessel-occlusion indicator); Model 4 = + serum PTX-3 (primary biomarker). Optimism-corrected C-statistic for Model 4 (1,000 bootstrap replicates) = .87. Events-per-variable ratio = 28/5 = 5.6. LR test, Model 3 vs. Model 4: $\chi^2(1) = 1.13$, p = .289. OR = odds ratio; CI = confidence interval; AF = atrial fibrillation (ECG or Holter); NIHSS = National Institutes of Health Stroke Scale; LVO = large-vessel occlusion; mRS = modified Rankin Scale; CRP = C-reactive protein. [c] n = 69 for ferritin (13 missing). ** p < .01. *** p < .001.

The hierarchical multivariable sequence (Table 5, Models 1–4) showed that admission NIHSS was the dominant predictor: adding NIHSS to the demographic block (Model 2) raised Nagelkerke's R^2 from .080 to .652 (adjusted OR for NIHSS = 1.45, 95% CI [1.22, 1.72], p < .001). Adding LVO status (Model 3) increased R^2 to .671, and adding PTX-3 (Model 4) to .681. In the fully adjusted model (Model 4), NIHSS remained the only statistically significant independent predictor (adjusted OR = 1.34, 95% CI [1.12, 1.61], p = .002); the effect of LVO status attenuated (adjusted OR = 4.01, 95% CI [0.69, 23.20], p = .122); and the unadjusted PTX-3 association became non-significant (adjusted OR = 1.003, 95% CI [0.997, 1.008], p = .292). A likelihood-ratio test confirmed the absence of incremental value for PTX-3 (LR $\chi^2(1) = 1.13$, p = .289). The events-per-variable ratio (28/5 = 5.6) is below the conventional threshold of 10 and indicates that multivariable estimates should be interpreted cautiously.

DISCUSSION

In our study, we evaluated the ability of serum PTX-3 levels to confirm the diagnosis in patients with acute ischemic stroke with large vessel occlusion. The findings demonstrated that PTX-3 levels alone are not a sufficient biomarker to confirm the diagnosis of LVO. However, PTX-3 levels correlated with NIHSS at admission, suggesting that it may indirectly reflect stroke severity. When the relationship between PTX-3 and clinical outcome was examined, PTX-3 was not an independent prognostic marker.

PTX-3 is an acute phase reactant consisting of high molecular weight multimers connected to each other by interchain disulfide bonds, which is produced by the immune system as a response, particularly in atherosclerotic lesions, as well as in various cardiovascular and cerebrovascular diseases (11,12). In the pathophysiology of acute ischemic stroke, activation of the immune response against the resulting vascular inflammation leads to an increase in PTX-3 levels (13,14). High levels of CRP, one of the classical short pentraxins, have been associated with poor prognosis in ischemic stroke (15).

In the literature, PTX-3 has been reported to be associated with inflammation and atherosclerosis and to be a prognostic marker, especially in cardiovascular diseases (16,17). In a study conducted in patients with ischemic stroke, high PTX-3 levels were also significantly associated with the extent and severity of carotid artery stenosis (18). In a study examining patients with acute myocardial infarction, the prognostic value of PTX-3 was higher than that of troponin, creatine kinase, CRP, and N-terminal pro-brain natriuretic peptide (19). The findings regarding the relationship between PTX-3 levels and disease severity and clinical outcome in stroke patients are not consistent. While some studies have reported that PTX-3 levels are associated with increased disease severity and poor clinical outcome, other studies have stated that this relationship could not be demonstrated (8,10,20). Our study likewise found no significant relationship between clinical outcome and PTX-3.

In our study, the lack of a significant discriminative power of PTX-3 levels in predicting the presence of LVO may be because this biomarker mainly reflects the systemic inflammatory response. Large vessel occlusion, on the other hand, is primarily a local vascular event and involves complex pathophysiological processes that cannot be explained solely by the degree of inflammation. This situation may explain why PTX-3 is insufficient for discriminating the presence of LVO.

On the other hand, the positive correlation found between PTX-3 levels and NIHSS in our study suggests that this biomarker may be associated with disease severity. This finding is consistent with previous studies showing that PTX-3 is a marker reflecting vascular inflammatory burden and may indicate the degree of neurological damage. However, the finding that PTX-3 was not an independent marker of clinical outcome in multivariable analyses suggests that the prognostic value of this molecule is limited.

The relatively limited sample size of this study and the low number of events in multivariable analyses may limit the generalizability of the obtained results. In addition, the single-center design and the specific characteristics of the patient population may further limit the generalizability of the results to different populations.

CONCLUSION

Our study has demonstrated that PTX-3 levels are not a sufficient biomarker for the diagnosis of LVO in acute ischemic stroke. Although PTX-3 levels are associated with stroke severity, they are not an independent marker of clinical outcome. These findings suggest that PTX-3 has limited value in clinical decision-making and does not contribute to existing clinical parameters.

Ethical approval

This study has been approved by the Düzce University Non-Interventional Health Research Ethics Committee (approval date 27/07/2025, number 2025/165). Written informed consent was obtained from the participants.

Author contribution

Concept: AY; Design: AY, AÖ; Data Collection or Processing: HŞ; Analysis or Interpretation: AY; Writing: HT, AY, AÖ; Supervision: MA. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Musuka TD, Wilton SB, Traboulsi M, Hill MD. Diagnosis and management of acute ischemic stroke: speed is critical. CMAJ. 2015; 187(12): 887-93. [Crossref]
2. Dixit A, Roy U. Efficacy and safety of endovascular thrombectomy in large-core infarcts: a systematic review and meta-analysis. Neurol Clin Neurosci. 2026;14(1):12-21. [Crossref]
3. Baez SDLC, García Del Barco D, Hardy-Sosa A, et al. Scalable bio marker combinations for early stroke diagnosis: a systematic review. Front Neurol. 2021; 12: 638693. [Crossref]
4. Anrather J, Iadecola C. Inflammation and stroke: an overview. neurotherapeutics. 2016; 13(4): 661-70. [Crossref]
5. Garlanda C, Bottazzi B, Bastone A, Mantovani A. Pentraxins at the crossroads between innate immunity, inflammation, matrix deposition, and female fertility. Annu Rev Immunol. 2005; 23: 337-66. [Crossref]
6. Wu Q, Cao F, Tao J, Li X, Zheng SG, Pan HF. Pentraxin 3: a promising therapeutic target for autoimmune diseases. Autoimmun Rev. 2020; 19(12): 102584. [Crossref]
7. Rajkovic I, Wong R, Lemarchand E, et al. Pentraxin 3 promotes long-term cerebral blood flow recovery, angiogenesis, and neuronal survival after stroke. J Mol Med (Berl). 2018; 96(12): 1319-32. [Crossref]
8. Zhang Y, Hu H, Liu C, Wu J, Zhou S, Zhao T. Serum pentraxin 3 as a biomarker for prognosis of acute minor stroke due to large artery atherosclerosis. Brain Behav. 2021; 11(1): e01956. [Crossref]
9. Sezer S, Uçar F, Ulusoy EK, et al. Serum amyloid A, fetuin-A, and pentraxin-3 levels in patients with ischemic stroke: novel prognostic biomarkers? Turk J Med Sci. 2014; 44(1): 16-23. [Crossref]
10. Zhang CY, Han HD, Wang SY, Huang SR, Deng BQ. Pentraxin-3 in thrombolytic therapy for acute ischemic stroke: no relation with curative effect and prognosis. Med Sci Monit. 2018; 24: 4427-32. [Crossref]
11. Shindo A, Maki T, Mandeville ET, et al. Astrocyte-derived pentraxin 3 supports blood-brain barrier integrity under acute phase of stroke. Stroke. 2016; 47(4): 1094-100. [Crossref]
12. Shindo A, Tanemura H, Yata K, et al. Inflammatory biomarkers in atherosclerosis: pentraxin 3 can become a novel marker of plaque vulnerability. PLoS One. 2014; 9(6): e100045. [Crossref]
13. Adams HP, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke. 1993; 24(1): 35-41. [Crossref]
14. Katakami N, Kaneto H, Sakamoto F, et al. Plasma pentraxin 3 levels are associated with carotid IMT in type 1 diabetic patients. Diabetes Res Clin Pract. 2013; 99(2): 185-91. [Crossref]
15. Huang Y, Jing J, Zhao XQ, et al. High-sensitivity C-reactive protein is a strong risk factor for death after acute ischemic stroke among Chinese. CNS Neurosci Ther. 2012; 18(3): 261-6. [Crossref]
16. Ristagno G, Fumagalli F, Bottazzi B, et al. Pentraxin 3 in cardiovascular disease. Front Immunol. 2019; 10: 823. [Crossref]
17. Adam CA, Șalaru DL, Prisacariu C, Marcu DTM, Sascău RA, Stătescu C. Novel biomarkers of atherosclerotic vascular disease-latest insights in the research field. Int J Mol Sci. 2022; 23(9): 4998. [Crossref]

18. Yi L, Tang J, Shi C, et al. Pentraxin 3, TNF-α, and LDL-C are associated with carotid artery stenosis in patients with ischemic stroke. Front Neurol. 2020; 10: 1365. **[Crossref]**

19. Latini R, Maggioni AP, Peri G, et al. Prognostic significance of the long pentraxin PTX3 in acute myocardial infarction. Circulation. 2004; 110(16): 2349-54. **[Crossref]**

20. Ryu WS, Kim CK, Kim BJ, Kim C, Lee SH, Yoon BW. Pentraxin 3: a novel and independent prognostic marker in ischemic stroke. Atherosclerosis. 2012; 220(2): 581-6. **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):254-262
DOI: 10.54307/NWMJ.2026.258

RESEARCH ARTICLE

Evaluation of range of motion after fasciectomy in Dupuytren's contracture and the effect of rehabilitation*

Ertuğrul Karanfil[1], Edanur Gezer Özçelik[1], Berkalp Berker[1], Mehmet Furkan Tuncer[1], Şükrü Burak Tönük[2], Metin Görgü[1]

[1]Department of Plastic, Reconstructive and Aesthetic Surgery, Abant İzzet Baysal University, Bolu, Türkiye
[2]Department of Physical Therapy and Rehabilitation, İzzet Baysal Training and Research Hospital, Bolu, Türkiye

* A preliminary cohort of 21 patients was presented at the TPRECD Congress (15–19 October 2025), and the study was subsequently expanded to 41 patients.

Cite as: Karanfil E, Gezer Özçelik E, Berker B, Tuncer MF, Tönük ŞB, Görgü M. Evaluation of range of motion after fasciectomy in Dupuytren's contracture and the effect of rehabilitation. Northwestern Med J. 2026;6(3):254-262.

ABSTRACT

Aims: Dupuytren's disease affects the palmar fascia and causes a movement disorder. Treatment varies considerably, and many treatment options exist. This study aimed to determine the extent to which joint range of motion (ROM) increases after surgery and how postoperative rehabilitation affects functional outcomes.

Materials and Methods: Forty-one patients who underwent fasciectomy for Dupuytren's contracture were analyzed, with a mean follow-up of 12 months. We collected demographic details, affected fingers, surgical techniques, physical therapy status after surgery, and both preoperative and late postoperative joint ROM values. We measured ROM by calculating the total flexion-extension arc of the metacarpophalangeal and proximal interphalangeal joints with a goniometer. We used paired-samples and independent-samples t-tests for statistical analysis.

Results: The study included 36 male and 5 female patients with a mean age of 65.7 ± 7.8 years. The fourth and fifth fingers were most often affected. The mean increase in postoperative ROM was 37.12 ± 6.45° for patients receiving physical therapy and 29.18 ± 15.55° for those who did not, showing a significant difference between the two groups (p value = 0.029). In both groups, the ROM gains achieved after surgery were mostly maintained at late follow-up.

Conclusion: Functional ROM achieved after fasciectomy for Dupuytren's contracture is preserved over the long term. Postoperative physical therapy offers an additional significant improvement in ROM and seems to help maintain functional recovery.

Keywords: Dupuytren's contracture, range of motion, rehabilitation

Corresponding author: Ertuğrul Karanfil **E-mail:** ertugrulkaranfill@gmail.com
Received: 12.05.2026 **Accepted:** 24.07.2026 **Published:** 31.07.2026

INTRODUCTION

Dupuytren's contracture is a long-term, progressive fibroproliferative disease of the palmar fascia. It causes progressive flexion contractures at the metacarpophalangeal (MCP) and proximal interphalangeal (PIP) joints, characterized by nodules and pathological cords (Figure 1). It most often affects the fourth and fifth fingers and occurs more frequently in men than in women (1,2). Histologically, it involves a fibroproliferative process where fibroblasts change into myofibroblasts in the palmar fascia (2). Dupuytren's disease is a common fibroproliferative disorder of the palmar fascia that predominantly affects individuals of Northern European descent and shows a marked male predominance. Progressive digital flexion contractures may impair hand function and interfere with daily activities (3).

Surgery is the most effective treatment for Dupuytren's contracture (4). The main goal is to straighten the fingers and improve hand function by removing the pathological tissue. Although there may be recurrence in some cases , these surgeries can provide good long-term outcomes when done correctly (5).

Currently, limited fasciectomy is the most commonly performed surgical procedure. This technique consists of excision of affected palmar and digital fascia while preserving the digital norovascular bundle. Research shows that this approach is effective for correcting contractures, especially in moderate to severe cases involving the PIP joint. It may have better long-term results with low recurrence rates than percutaneous methods (2,3,6).

Alternatively, percutaneous needle fasciotomy is a quick and minimally invasive option, especially for MCP joint contractures. However, it has higher recurrence rates because of not removing all the pathological tissue (6-8). Dermofasciectomy is a more aggressive procedure than other methods; it removes all affected fascia and skin. This approach is primarily used for recurrent or severe cases with significant skin involvement. This aproach has a low recurrence rate but is only done in selected cases because of longer recovery times and a higher risk of complications (9). Limited surgical techniques like segmental fasciectomy or minimal aponeurotomy may be suggested for mild or localized disease (10).

Selection of the suitable surgical treatment for Dupuytren's contracture should be individualized for each patient. Factors to evaluate included the severity of the contracture, the affected joints, the patient's age, job requirements, expected function, and the chance of recurrence. Current literature supports limited fasciectomy as the best option for most patients.

In this study, we aimed to evaluate the long-term postoperative joint range of motion in patients who underwent limited or total fasciectomy for Dupuytren's contracture at our clinic. We also wanted to learn how postoperative physical therapy affected these surgical results. The main goal of surgical treatment is to increase finger extension and enhance hand function by removing the tissue that causes the contracture.

MATERIALS AND METHODS

Approval for this study was received from our hospital's Local Ethics Committee (Decision No. 2024/110). We included 41 patients diagnosed with Dupuytren's contracture who underwent surgical treatment.

19 patients received local anesthesia, 5 received a regional block, and 17 received general anesthesia. A tourniquet is used on the upper arm. We preferred Bruner or zigzag incisions for surgery because they follow the hand's natural skin lines and provide better exposure of the affected fascia. If skin contracture was present, we modified these incisions using Z-plasties. In areas where the tissue had adhered, we carefully dissected skin flaps to separate them from the palmar fascia. Eight patients had limited fasciectomy, while 33 underwent total fasciectomy. The surgery involved removing the palmar aponeurosis and the pathological fascial bands that connect to the tendons and fingers. In 25 patients with PIP joint involvement, we traced the faulty cords to the joint and removed them completely and precisely. If the PIP joint capsule was contracted, a limited capsulotomy was performed.

After confirming that there was no bleeding, we closed the incisions in one layer using 5/0 or 6/0 polypropylene sutures while keeping the hand

extended. For skin defects, we used a full-thickness skin graft from the thigh. Four patients (9.8%) required FTSG reconstruction due to skin defects that could not be closed directly. All four healed well, and their postoperative ranges of motion were similar to those of the other patients. We placed a 5-mm Penrose drain at the proximal end of the incision.

After surgery, we maintained the hand in a neutral position using a splint. The drain was removed the day after surgery, and patients wore the splint for a week. Everyone was encouraged to raise the affected hand and start finger exercises early in the recovery process.

Joint range of motion measurement

We measured range of motion (ROM) at the MCP and PIP joints using a standard goniometer before surgery and at least six months after surgery. We recorded the total flexion-extension range.

Physical therapy protocol

Starting in the third week after surgery, patients participated in a personalized rehabilitation program that included active and passive joint mobilization, stretching exercises, and the use of a night splint. The average duration for physical therapy was 6 ± 2 weeks. In patients who underwent FTSG reconstruction, we delayed initiation of physical therapy to allow for proper graft integration.

Statistical analysis

We conducted statistical analyses using SPSS version 25.0 (IBM Corp., Armonk, NY, USA). We expressed continuous variables as mean ± standard deviation (SD) and categorical variables as frequency and percentage (%). We assessed the normality of the data distribution using the Shapiro-Wilk test.

We used paired-samples t-tests for intra-group comparisons of preoperative and postoperative ROM. To compare the changes in ROM between groups (those receiving physical therapy vs. those not receiving it), we performed independent samples t-tests. If the data did not meet the assumption of normality, we used the Mann-Whitney U test as a non-parametric alternative. A p-value of <0.05 was considered statistically significant.

Inclusion and exclusion criteria

The study included patients with primary Dupuytren's contracture who had a flexion contracture of ≥20° at the MCP and/or PIP joint, had undergone limited or total fasciectomy, and could be followed for at least 6 months. We excluded patients with secondary contractures resulting from trauma, infection, burns, or neurological diseases, as well as those with congenital deformities, rheumatological diseases, or incomplete data records.

RESULTS

In all 41 patients included in the study(36 men, 5 women; mean age 65.7 ± 7.8 years), significant improvement in MCP and/or PIP joint range of motion was achieved following limited or total fasciectomy. The majority of the affected fingers were the 4th and 5th fingers (Table 1). The mean follow-up period was 12 months.

Twenty-four patients (58.5%) followed the postoperative physical therapy program. while 17(41.5%) did not participate. In both groups, a statistically significant increase in postoperative ROM values was observed compared to the preoperative period ($p < 0.001$). The mean ROM increase was 37.12 ± 6.45° in the group receiving physical therapy, while it was 29.18 ± 15.55° in the group not receiving physical therapy. A statistically significant difference was found

Table 1. Distribution of patients by gender and location

Gender	**n**	**%**
Female	5	12.2
Male	36	87.8
Affected finger	**n**	**%**
Thumb	2	4.9
2nd and 3rd fingers	10	24.4
4th and 5th fingers	29	70.7

Table 2. Comparison of ROM values between patients who received and did not receive postoperative physical therapy (n=41)

Group	n	Preoperative ROM (Mean ± SD)	Postoperative ROM (Mean ± SD)	ROM increase (Mean ± SD)	Intra-group p-value	p-value between-groups
Received physical therapy	24	139.17 ± 6.16	176.29 ± 2.71	+37.12± 6.45	< 0,001	0.029
No physical therapy	17	141.82 ± 15.65	171 ± 4.11	+29.18 ± 15.55	< 0,001	0.029

ROM: Range of Motion of the joint

SD: Standard Deviation

Table 3. Distribution of involved joints

Joint involvement	Total Patients	Limited Fasciectomy	Total Fasciectomy
MCP + PIP (combined)	30	1 (%3,3)	29 (%96,7)
MCP only	9	5 (%55,6)	4 (%44,4)
PIP only	2	2 (%100)	0
Total	41	8 (%19,5)	33 (%80,5)

MCP: Metacarpophalengeal

PIP: Proximal Interphalengeal

between the groups in terms of ROM gain (p = 0.029) (Table 2). While total fasciectomy was performed on the vast majority (96.7%) of patients with combined MCP + PIP contractures, limited fasciectomy was preferred for 55.6% of patients with MCP contractures alone. Limited fasciectomy was performed on both of the 2 patients with PIP contractures alone (Table 3).

DISCUSSION

The incision techniques used in Dupuytren's contracture surgery aim to ensure adequate exposure of the pathological fascia while preserving skin vascularity and minimizing postoperative complications. Various skin incision techniques, including zigzag (Bruner), longitudinal, and transverse palmar incisions, have been described in the literature for Dupuytren contracture surgery (1,2,10). In cases of Dupuytren's contracture where digital extension is present, converting longitudinal digital incisions into Z-plasty is a commonly used approach. This technique reduces skin tension, thereby decreasing the incidence of postoperative complications (11,12). Consequently, the use of Z-plastymodified incisions is recommended (13). In addition, multiple transverse palmar incisions may also be performed (14); however, long-term recurrence rates have been found to be higher (8). In recent years, there has been a growing consensus that a single standardized incision is not universally applicable to all patients. Consequently, for cases presenting with palmar and digital involvement, incisions that follow the natural skin creases, are tailored to the individual, and are modified with Z-plasty when necessary, are highly recommended (20).

In the current study, a standard incision technique was not used; instead, individualized incisions that align with the hand's natural lines, safely expose the entire pathological fascia, and are supported by Z-plasty in cases with skin contraction were preferred (Figure 2). This approach is consistent with the functional and combined incision principles described in the literature (13). The low rates of skin complications and the level of functional improvement suggest that individualized incision planning may contribute to surgical success.

Open limited fasciectomy, the most commonly performed surgical method for Dupuytren's contracture, has been shown to provide good to excellent clinical outcomes in short- and medium-term follow-ups; success rates reported in the literature range from 60% to 90% (15). While early-stage surgical improvement is generally reported to be good in cases undergoing limited or total fasciotomy, it has

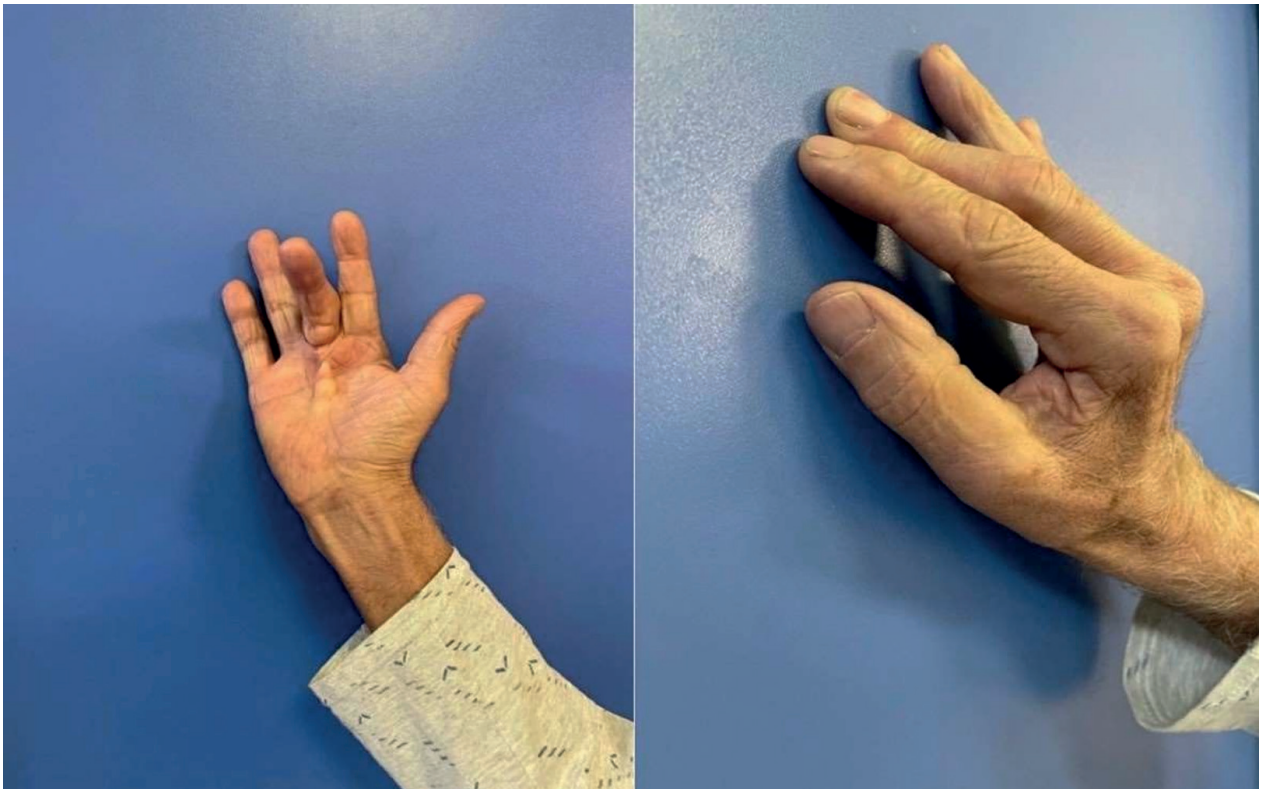


Figure 1. Preoperative view of the Dupuytren's contracture, 90-degree flexion at the MP (metacarpophalengeal) joint.

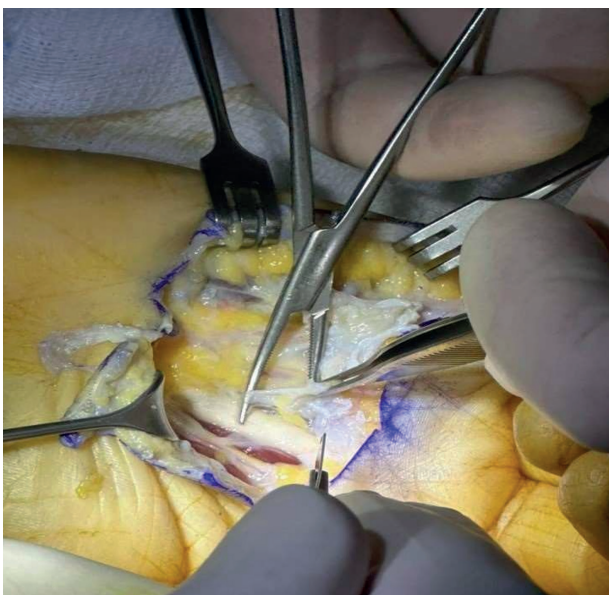


Figure 2. Postoperative view of the Dupuytren's contracture.

been demonstrated that recurrence rates can reach significant levels in long-term follow-ups (1,2,8). While rapid early functional recovery has been reported with minimally invasive methods such as percutaneous needle fasciotomy, high recurrence rates may be encountered in the long term, particularly in cases with PIP joint involvement (8). Dermofasciectomy, on the other hand, has been associated with lower recurrence rates in the literature; however, due to a longer recovery period and increased risk of complications, it is generally recommended for selected patients (1,2). In the current study, total fasciectomy was performed in the vast majority of cases and an average increase of 29–37° in total postoperative joint range of motion was achieved (Figure 3). These values are consistent with the ranges of functional improvement reported in the literature following fasciectomy (15) (Table 2). Furthermore, the absence of clinically significant loss of motion during the follow-up period supports the efficacy of the surgical technique employed. In Dupuytren's contracture surgery, the use of full-

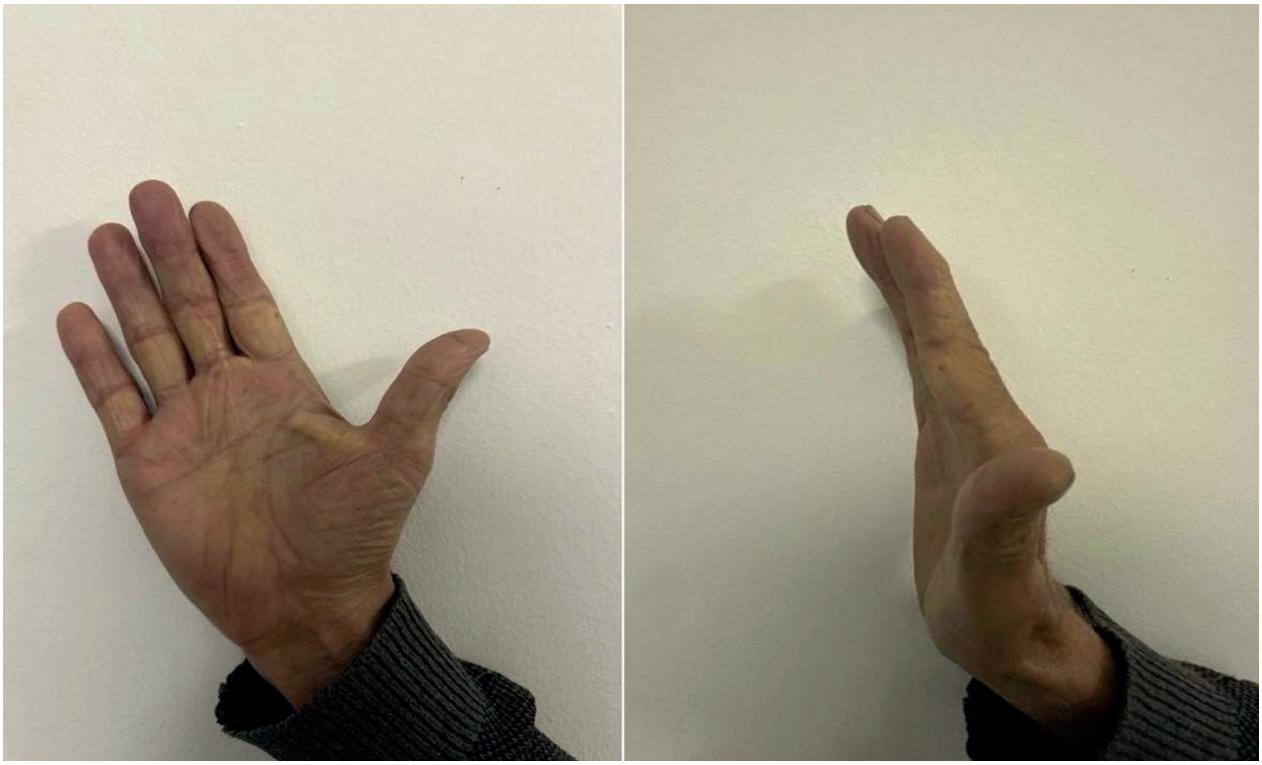


Figure 3. Intraoperative view of the Dupuytren's contracture.

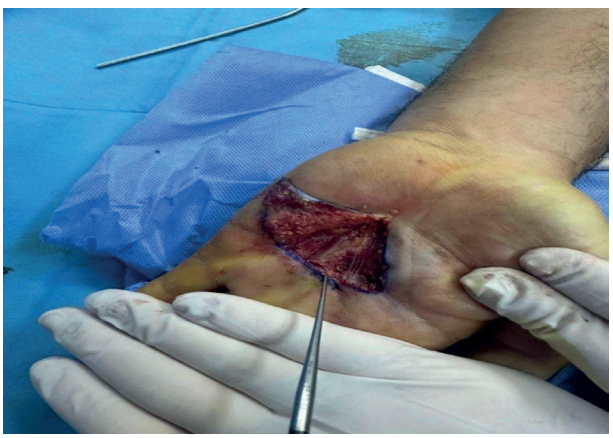


Figure 4. Intraoperative view of the Dupuytren's contracture.

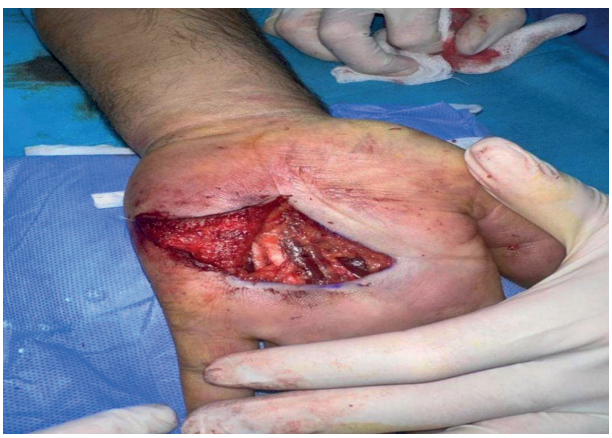


Figure 5. Intraoperative view of the Dupuytren's contracture, after the removal of the fascia.

thickness skin grafts (FTSG) is recommended in cases of advanced disease, significant skin contracture, and situations where primary skin closure is not possible following the excision of pathological fascia (Figure 6).

Recent studies emphasize that reduced skin elasticity and increased dermal adhesions in the presence of long-standing contractures are key factors determining the need for a graft. Clinical series and reviews published in recent years report that full-thickness

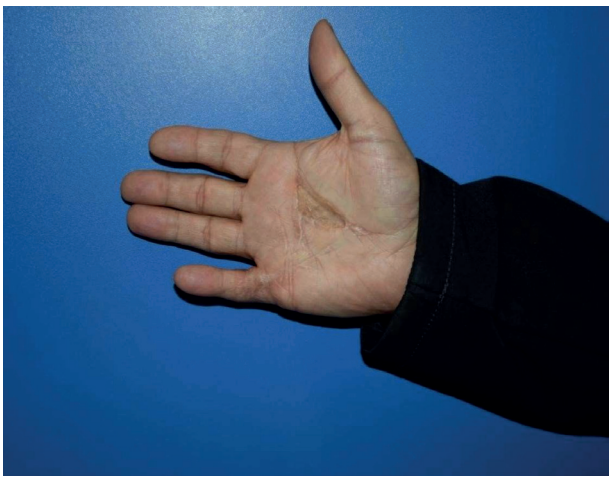


Figure 6. Repair with a graft, post-op 12 months.

skin grafts, particularly when performed in conjunction with dermofasciectomy, significantly reduce local recurrence rates. In a study published by Ansari et al., it was reported that surgical success under the graft was high in patients who underwent dermofasciectomy and FTSG, and that recurrence tended to occur more frequently in areas where grafts were not applied (16). Additionally, they reported that full-thickness skin grafts applied in the early postoperative period following dermofasciectomy are reliable and do not adversely affect functional outcomes. Current literature indicates that the rates of full-thickness skin graft application in series involving dermofasciectomy range between 20% and 40% (17). In contrast, it is stated that the need for grafts is significantly lower in fasciectomy-based surgeries and is generally limited to cases with advanced-stage, severe skin contractures (17). Although we did not perform dermofasciectomy in our study, full-thickness skin grafts were applied in 4 out of 41 cases (9.8%) to close the defect formed after releasing the contracture. This rate is lower than graft rates reported in dermofasciectomy series.

This can be explained by the fact that fasciectomy was performed without skin excision in the vast majority of cases and that incisions were planned in accordance with skin-sparing principles. Indeed, designing the incisions to align with the hand's natural creases and modifying them with Z-plasty in cases accompanied by skin contracture contributed to reducing skin tension and enabled primary closure (12,13). Since full-thickness skin grafts are known to reduce recurrence in the long term, their use is recommended in appropriate patients for Dupuytren's contracture surgery (17,18). In conclusion, the use of full-thickness skin grafts in our study was limited to cases selected in accordance with current indications described in the literature; we believe that the need for grafts can be reduced when combined with appropriate surgical selection for the patient and skin-sparing surgical techniques.

In the current study, a significant increase in total joint ROM was observed in the late postoperative period in patients who underwent surgery for Dupuytren contracture. The increase in ROM was significant in both the group receiving physical therapy and the group not receiving it; an average increase of 37.12 ± 6.45° in the group receiving physical therapy and 29.18 ± 15.55° in the group not receiving physical therapy was observed; the difference The difference between the groups was statistically significant ($p = 0.029$) (Table 2). This finding suggests that surgical treatment is the primary determinant of functional improvement in Dupuytren's contracture. It has been reported in some studies that nighttime splinting following Dupuytren's contracture surgery may negatively affect surgical success and lead to a decrease in joint range of motion (12). Additionally, a multicenter study demonstrated that postoperative physical therapy interventions did not provide an additional gain in range of motion (12). However, a different study noted that hand rehabilitation may contribute to maintaining the joint range of motion achieved through surgery (19). In the current study statistical analysis revealed that postoperative physical therapy for Dupuytren's contracture provided an additional contribution to ROM increase. This finding suggests that physical Therapy yields a significant advantage ROM.

Although the observed difference reached statistical significance, its clinical impact was relatively small. This may be because all participants were taught specific massage techniques and home exercises, which constituted a key part of their physical therapy. As a result, the patients' strict adherence to these home programs likely compensated for the missed in-person hospital sessions. In our study, focusing only on the MCP joint level led to better outcomes than cases with contractures that extended to the

PIP joint level, which matches previous reports (4). These observations are based on clinical impressions rather than on group analyses and should therefore be interpreted with caution. Higher recurrence rates and limited functional improvements have been reported, especially in cases involving the PIP or DIP (Distal Interphalengeal) joints (4). These findings show that the level of the affected anatomical structure is critical for predicting postoperative outcomes.

The vast majority of patients achieved a good functional range of motion after surgery, with significant improvements observed. While postoperative physical therapy can significantly improve the range of motion, differences in treatment protocols and patient compliance must be considered. Therefore, careful selection of the timing of surgery, treatment methods, and rehabilitation strategies remains crucial.

This study has some limitations. Its retrospective nature and reliance on patient records made it difficult to objectively monitor adherence to postoperative physical therapy. Furthermore, the lack of randomized testing may lead to bias. We were unable to standardize the duration and type of physical therapy because they were based on patient reports of home exercises. Additionally, our assessment focused solely on range-of-motion measurements because we did not aggregate patient-reported functional outcomes.

While an average follow-up period of 12 months is suitable for short- and medium-term assessments, it is insufficient to definitively evaluate long-term recurrence rates. Combining patients who underwent limited and total fasciectomies reduced our ability to assess potential differences between these two techniques. Furthermore, the timelines for initiating physical therapy for patients who underwent skin grafting were inconsistent, which may have affected the outcomes. In particular, the observation that patients with isolated MCP joint contractures benefited more from treatment stemmed from clinical impressions rather than formal subgroup analysis, which affects the generalizability of this finding.

CONCLUSION

Both limited and total fasciectomies can significantly improve joint function in patients with Dupuytren's contracture. Gains in mobility are generally sustained over time.

Postoperative physical therapy was associated with an additional significant improvement in range of motion, but the overall effect was modest. This demonstrates that although surgery is the primary factor in recovery, rehabilitation plays a crucial role in optimizing and maintaining outcomes. Personalized surgical plans and well-chosen rehabilitation programs can enhance recovery after Dupuytren's contracture surgery.

Ethical approval

This study has been approved by the Bolu Abant Izzet Baysal University Ethics Committee (approval date 15.05.2024, number 2024/110). Written informed consent was obtained from the participants.

Author contribution

Concept: EK, MG, ŞBT; Design: EK, ŞBT; Data Collection or Processing: EGÖ, BB, MFT; Analysis or Interpretation: BB, MFT; Literature Search: BB, EGÖ; Writing: EK, MG, EGÖ. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Townley WA, Baker R, Sheppard N, Grobbelaar AO. Dupuytren's contracture unfolded. BMJ. 2006; 332(7538): 397-400. **[Crossref]**
2. Khashan M, Smitham PJ, Khan WS, Goddard NJ. Dupuytren's disease: review of the current literature. Open Orthop J. 2011; 5(Suppl 2): 283-8. **[Crossref]**

3. Wilbrand S, Ekbom A, Gerdin B. The sex ratio and rate of reoperation for Dupuytren's contracture in men and women. J Hand Surg Br. 1999; 24(4): 456-9. **[Crossref]**
4. Dias JJ, Braybrooke J. Dupuytren's contracture: an audit of the outcomes of surgery. J Hand Surg Br. 2006; 31(5): 514-21. **[Crossref]**
5. Hueston JT. Regression of Dupuytren's contracture. J Hand Surg Br. 1992; 17(4): 453-7. **[Crossref]**
6. Oppermann J, Unglaub F, Müller LP, Löw S, Hahn P, Spies CK. Percutaneous needle aponeurotomy for Dupuytren's contracture. Orthopade. 2017; 46(4): 315-20. **[Crossref]**
7. Beaudreuil J, Lermusiaux JL, Teyssedou JP, et al. Multi-needle aponeurotomy for advanced Dupuytren's disease: preliminary results of safety and efficacy (MNA 1 study). Joint Bone Spine. 2011; 78(6): 625-8. **[Crossref]**
8. van Rijssen AL, Gerbrandy FSJ, ter Linden H, Klip H, Werker PMN. A comparison of the direct outcomes of percutaneous needle fasciotomy and limited fasciectomy for Dupuytren's disease: a 6-week follow-up study. J Hand Surg Am. 2006;31(5):717-25. **[Crossref]**
9. Raval P, Kulkarni K, Johnson N, et al. Frequency and reporting of complications after dupuytren contracture interventions: a systematic review and meta-analysis. Plast Reconstr Surg. 2025; 155(1): 115-25. **[Crossref]**
10. Moermans JP. Segmental aponeurectomy in Dupuytren's disease. J Hand Surg Br. 1991; 16(3): 243-54. **[Crossref]**
11. Bulstrode NW, Jemec B, Smith PJ. The complications of Dupuytren's contracture surgery. J Hand Surg Am. 2005; 30(5): 1021-5. **[Crossref]**
12. Jerosch-Herold C, Shepstone L, Chojnowski AJ, Larson D, Barrett E, Vaughan SP. Night-time splinting after fasciectomy or dermo-fasciectomy for Dupuytren's contracture: a pragmatic, multicenter, randomized controlled trial. BMC Musculoskelet Disord. 2011; 12: 136. **[Crossref]**
13. Gelberman RH, Panagis JS, Hergenroeder PT, Zakaib GS. Wound complications in the surgical management of Dupuytren's contracture: a comparison of operative incisions. Hand. 1982; 14(3): 248-54. **[Crossref]**
14. Lee H, Eo S, Cho S, Jones NF. The surgical release of dupuytren's contracture using multiple transverse incisions. Arch Plast Surg. 2012; 39(4): 426-30. **[Crossref]**
15. Crean SM, Gerber RA, Le Graverand MPH, Boyd DM, Cappelleri JC. The efficacy and safety of fasciectomy and fasciotomy for Dupuytren's contracture in European patients: a structured review of published studies. J Hand Surg Eur Vol. 2011; 36(5): 396-407. **[Crossref]**
16. Ansari SA, Kennedy JA, Younis F. Skin graft review in dermofasciectomy for dupuytren's contracture within 48 hours: challenging historical practice. J Hand Microsurg. 2021; 14(1): 71-8. **[Crossref]**
17. De Ketele A, Degreef I. Full-thickness skin grafting in preventing recurrence of Dupuytren's disease: A systematic review. Hand Surg Rehabil. 2023; 42(4): 273-83. **[Crossref]**
18. Rodrigues JN, Becker GW, Ball C, et al. Surgery for Dupuytren's contracture of the fingers. Cochrane Database Syst Rev. 2015; 2015(12): CD010143. **[Crossref]**
19. Turesson C. The role of hand therapy in dupuytren disease. Hand Clin. 2018; 34(3): 395-401. **[Crossref]**
20. Stoneham A, Fetouh S, Kazzam M, Warwick D. The 'Universal Approach' for Dupuytren's disease: A safe and reproducible sequence for planning fasciectomy incisions. JPRAS Open. 2024;41:372-75. **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):263-267
DOI: 10.54307/NWMJ.2026.195

CASE REPORT

Case report of congenital myasthenic syndrome due to CHRNE mutation diagnosed in adulthood*

Fatma Bilgili[1], **Fatma Nur Elyıldırım**[2], **Serpil Yıldız**[2]

[1]Department of Neurology, Ardahan Public Hospital, Ardahan, Türkiye
[2]Department of Neurology, Bolu Abant İzzet Baysal University, Bolu, Türkiye

* This article was presented as an oral presentation at the 7th TND Neuromuscular Diseases Congress held in Çeşme, İzmir, Türkiye between September 20, 2024 and September 22, 2024.

Cite as: Bilgili F, Elyıldırım FN, Yıldız S. Case report of congenital myasthenic syndrome due to CHRNE mutation diagnosed in adulthood. Northwestern Med J. 2026;6(3):263-267.

ABSTRACT

Congenital myasthenic syndromes are a rare subgroup of neuromuscular diseases caused by genetic defects in proteins involved in the structure, function and repair of the neuromuscular junction. Although these syndromes usually start in the neonatal and childhood periods, they may rarely start in adulthood. Because the clinical and laboratory findings of congenital myasthenic syndrome that starts in adulthood closely resemble those of seronegative myasthenia gravis, a differential diagnosis between the two is essential. This differential diagnosis is clinically important because treatment approaches and drug responses differ between the two diseases. Here, we aimed to draw attention to this disease and increase clinical awareness by presenting a case of congenital myasthenic syndrome diagnosed in adulthood.

Keywords: CHRNE, congenital nyasthenic syndrome, neuromuscular diseases

INTRODUCTION

Myasthenia, first described by Friedrich Jolly in 1895, is a clinical syndrome characterized by diffuse muscle weakness that can be acquired or inherited and affects transmission at the neuromuscular junction (1). Myasthenia gravis (MG) is a classic example of an antibody-mediated autoimmune disease that occurs frequently in adults. Autoantibodies targeting acetylcholine receptors (AChR), muscle-specific kinase (MuSK), low-density lipoprotein receptor-related protein 4 (Lrp4) and agrin have been identified in its pathogenesis (2). If no antibodies are detected, it is defined as seronegative myasthenia gravis (2). Congenital myasthenic syndrome (CMS) is a form of myasthenia that usually occurs in newborns and early childhood due to various genetic mutations (3). Clinical symptoms in CMS can be heterogeneous, depending on the underlying genetic mutation and the affected site at the neuromuscular junction (presynaptic, synaptic or postsynaptic) (4). Although

Corresponding author: Fatma Bilgili **E-mail:** fatmabilgili19@gmail.com
Received: 11.05.2025 **Accepted:** 09.06.2026 **Published:** 31.07.2026

rare, CMS can initially present in adulthood and closely mimic seronegative myasthenia gravis, both clinically and on electromyographic (EMG) evaluation (3). It is important to distinguish between CMS and myasthenia gravis due to differences in treatment and drug responses. We present a case of congenital myasthenic syndrome diagnosed in adulthood to highlight this underrecognized condition and raise awareness among clinicians.

CASE REPORT

A 22-year-old male patient presented with fatigue for the last 3 years and ptosis of the right eyelid in the evening for the last 5-6 months. On detailed history-taking for myasthenic syndromes, the patient reported no additional complaints. The patient had visited an ophthalmologist once during childhood because of difficulty fully opening his eyes, but no diagnosis was established at that time. There were no other remarkable features in the patient's medical history. There was no family history of similar disorders; however, the patient's parents were first cousins. Neurologic examination was normal. Detailed blood tests were normal and acetylcholine receptor antibodies were negative. Other antibodies associated with myasthenia gravis could not be analyzed because they were not available for testing at our institution. Cranial and cervical magnetic resonance imaging (MRI) showed no pathologic findings. EMG examination showed a more than 10% decremental response in the nasalis muscle with repetitive nerve stimulation at a frequency of 3-5 Hz. We requested genetic testing for congenital myasthenic syndromes because of a history of suspected childhood ptosis. Genetic analysis revealed a pathogenic homozygous mutation in the CHRNE gene.

DISCUSSION

CMS is a rare, clinically and genetically heterogeneous disorder of juvenile and adult onset caused by abnormalities in the structure and function of the neuromuscular junction (1). Data on the prevalence of CMS are limited and are derived primarily from isolated case reports. To date, only a limited number of CMS patients have been identified in Turkey, and the genotypic and phenotypic characteristics of CMS in this population remain unclear (5). According to a recent study, the prevalence of CMS is estimated to be 1/10 of that of myasthenia gravis, corresponding to 25-125 per million (4). The prevalence of CMS has been reported to vary between ethnic groups and regions. This variation may be due to the mild nature of CMS symptoms and diagnostic delays caused by its differential diagnosis with many other conditions (4,6,7). CMS is a rare disorder, with even rarer forms in which symptoms begin in adulthood, as observed in the patient presented above. Differentiating these patients from those with seronegative MG is challenging but critical, given the decisive implications for treatment. Weakness, particularly affecting the eye, bulbar, and proximal extremity muscles; symptoms extending from birth to early childhood; a positive family history; a negative antibody test profile; clinical and neurophysiological myasthenic findings; EMG findings consistent with neuromuscular junction dysfunction; and an inadequate response to immunotherapy and acetylcholinesterase inhibitors all suggest a diagnosis of CMS (3). Myopathy is another important diagnosis to consider, particularly in patients with late-onset limb-girdle muscle weakness who lack the more typical findings of CMS, such as ptosis and marked restriction of eye movement. However, even when elevated serum creatine kinase (CK) levels and myopathic changes on EMG are present, these patients should still be evaluated for CMS (5).

CMS is clinically characterized by abnormal fatigue or temporary or permanent weakness of extraocular, facial, bulbar, trunk, respiratory or limb muscles. Onset may be intrauterine or congenital, or may occur in infancy or childhood, and rarely in adolescence or adulthood. Severity ranges from mild muscle weakness to respiratory failure and early death (4). Although all CMS subtypes share common clinical features such as fatigue and muscle weakness, the age of onset, specific symptoms, and treatment efficacy may differ depending on the underlying genetic defect and associated molecular mechanisms (8). Treatment options and evaluations based on genetic subtype are given in Table 1 (8).

The structures affected within the neuromuscular junction due to genetic mutation are shown in

Table 1. Congenital myasthenic syndromes: A clinical and treatment approach

Syndrome	Treatment	Notes
AChR deficiency (*CHRNE*)	3,4-DAP, pyridostigmine	Albuterol is an important add on agent
AChR kinetic defect: fast channel syndrome	Pyridostigmine, 3,4-DAP	Avoid quinidine and fluoxetine; beta-adrenergic agonists may be an important adjunct
ALG2 and ALG14 myasthenia	Pyridostigmine	3,4-DAP might confer added benefit
ChAT deficiency	Pyridostigmine; parenteral neostigmine for apneic episodes	Even asymptomatic patients should be treated due to risk of fatal apnea. Severe phenotype may not respond to cholinergic drugs and may require permanent ventilatory support and apnea monitor; beta-adrenergic agonists may be an important adjunct
DPAGT1 myasthenia	Pyridostigmine	May respond to add on 3,4-DAP and albuterol
GFPT1 myasthenia	Pyridostigmine, 3,4-DAP	May see additional improvement with an add on adrenergic agents such as albuterol or ephedrine
GMPPB myasthenia	Pyridostigmine	May see additional improvement with added 3,4-DAP, albuterol
MUNC13-1 myasthenia	3,4-DAP	Perhaps improvement with add on pyridostigmine
MYO9A deficiency	Pyridostigmine	Respiratory crises have occurred following intake of 3,4-DAP and fluoxetine
PREPL deficiency	May respond to pyridostigmine early in life	
Rapsyn deficiency	Pyridostigmine, 3,4-DAP	Albuterol may be helpful as an add on treatment; worsening with fluoxetine reported
SCN4A sodium channel myasthenia	Pyridostigmine	Acetazolamide also can mitigate periodic paralysis due to mutations in the same gene
SLC5A7 deficiency	Pyridostigmine	Treats episodic apnea
SNAP25 deficiency	3,4-DAP	Improves weakness, not ataxia Not pyridostigmine
Synaptobrevin-1 myasthenia	Pyridostigmine may be helpful	
Synaptotagmin-2 myasthenia	3,4-DAP	Pyridostigmine ineffective in a small study
VAChT or (*SLC18A3*) myasthenia	Pyridostigmine	Moderate improvement
AChE deficiency (COLQ)	Albuterol or ephedrine	Avoid pyridostigmine and 3,4-DAP
Agrin deficiency	Ephedrine	No response to cholinergic agents
DOK7 deficiency	Albuterol or ephedrine	Avoid pyridostigmine
LRP4 deficiency	Albuterol	Avoid cholinergic agents
MuSK defect	Albuterol Variable response to 3,4-DAP	Pyridostigmine conventional doses may worsen symptoms
AChR kinetic defect: slow-channel syndrome	Fluoxetine, quinidine	Avoid cholinergic agents

*This table is adapted from the study ' Congenital Myasthenic Syndromes: a Clinical and Treatment Approach' by Farmakidis C. et al.

Table 2 (1). Of the 32 genes identified to date to be associated with CMS, mutations in six of them (CHRNE, RAPSN, COLQ, DOK7, GFPT1, and CHAT) account for approximately 90% of cases (3,4). The CHRNE gene encodes the epsilon subunit of the acetylcholine receptor (AChR) and its mutation leads to a deficiency of acetylcholine receptors or abnormal channel kinetics. To date, the most frequently reported mutation in CMS has been in this gene (9). The most common mutations in Turkey are those that cause AChR deficiency (particularly the c.1219+2T>G mutation in the CHRNE gene) (5).

Most patients with CHRNE mutations present with mild, progressive bulbar, respiratory or diffuse limb weakness, often accompanied by ptosis or ophthalmoplegia at birth. However, the type and severity of clinical manifestations can vary considerably between affected families (4,10). Repetitive nerve stimulation (RNS) may show a decremental response or appear normal, while single fiber EMG (SF-EMG) may show increased jitter (4,11). Most patients respond well to acetylcholinesterase inhibitors (AChEIs), but in some, pyridostigmine and 3,4-Diaminopyridine (3,4-DAP) may be ineffective or may worsen their condition. In some cases, albuterol or salbutamol alone can be highly effective. Fluoxetine alone may be ineffective, but its combination with salbutamol may give better results (4,9,11).

Early diagnosis of this condition, which cannot be reliably distinguished from myasthenia gravis due to overlapping clinical features and findings, is crucial to ensure appropriate treatment. In suspected cases, a definitive diagnosis can only be established through genetic testing. In our case, the childhood onset of symptoms and the suspicious clinical presentation led us to request genetic testing. The genetic test results identified the most common CHRNE mutation, consistent with data from Turkey. After receiving the genetic test results, we started our patient on fluoxetine treatment and observed an improvement in his symptoms.

Through this case, we aimed to contribute to the limited body of literature on congenital myasthenic

Table 2. Classification of congenital myasthenic syndromes (CMS) related to pattern of inheritance and molecular targets at the neuromuscular junction

I. Pattern of inheritance	
Autosomal dominant (gain-of-function)	Slow-channel syndrome, SNAP25B*, SYT2*
Autosomal recessive (loss-of-function)	All other subtypes
II. Site of defect and molecular targets at the neuromuscular junction	
Presynaptic defects	ChAT deficiency, SNAP25B deficiency, synaptotagmin-2 deficiency
Acetylcholine receptor defects	Primary deficiency, slow-channel syndrome (CHRNA1, CHRNB, CHRND, CHRNE, CHRNG), fast-channel syndrome (CHRNA, CHRND, CHRNE)
Synaptic basal lamina defects	Acetylcholinesterase deficiency (ColQ), β2-laminin deficiency
Endplate development and maintenance congenital defects	Agrin deficiency, MuSK deficiency, LRP4 deficiency, Dok-7 deficiency, rapsyn deficiency, COL13A1 mutations
Metabolic and mitochondrial disorders	Congenital disorders of glycosylation, SLC25A1 gene mutations
Others	Congenital myopathies with secondary neuromuscular transmission compromise (MTM1, RYR1, DNM2, TPM3, BIN1); PREPL deletion; plectin deficiency

*Extremely rare presentations. ** This table is adapted from the study 'Clinical and Genetic Basis of Congenital Myasthenic Syndromes' by Souza P. et al.

syndrome, a condition whose incidence is primarily inferred from case reports.

Ethical approval

Since this is a case report, ethics committee approval was not required. Consent was obtained from the patient.

Author contribution

Surgical and Medical Practices: FB, SY; Concept: FB, SY; Design: FB, FE; Data Collection or Processing: FB, FE; Analysis or Interpretation: SY, FB, FE; Literature Search: SY, FB, FE; Writing: FB, FE. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Souza PVSD, Batistella GNDR, Lino VC, Pinto WBVDR, Annes M, Oliveira ASB. Clinical and genetic basis of congenital myasthenic syndromes. Arq Neuropsiquiatr. 2016; 74(9): 750-60. **[Crossref]**
2. Dresser L, Wlodarski R, Rezania K, Soliven B. Myasthenia gravis: epidemiology, pathophysiology and clinical manifestations. J Clin Med. 2021; 10(11): 2235. **[Crossref]**
3. Gül Mert G, Özcan N, Hergüner Ö, et al. Congenital myasthenic syndrome in Turkey: clinical and genetic features in the long-term follow-up of patients. Acta Neurol Belg. 2021; 121(2): 529-34. **[Crossref]**
4. Finsterer J. Congenital myasthenic syndromes. Orphanet J Rare Dis. 2019; 14(1): 57. **[Crossref]**
5. Durmus H, Shen XM, Serdaroglu-Oflazer P, et al. Congenital myasthenic syndromes in Turkey: Clinical clues and prognosis with long term follow-up. Neuromuscul Disord. 2018; 28(4): 315-22. **[Crossref]**
6. Mihaylova V, Scola RH, Gervini B, et al. Molecular characterisation of congenital myasthenic syndromes in Southern Brazil. J Neurol Neurosurg Psychiatry. 2010; 81(9): 973-7. **[Crossref]**
7. Parr JR, Andrew MJ, Finnis M, Beeson D, Vincent A, Jayawant S. How common is childhood myasthenia? The UK incidence and prevalence of autoimmune and congenital myasthenia. Arch Dis Child. 2014; 99(6): 539-42. **[Crossref]**
8. Farmakidis C, Pasnoor M, Barohn RJ, Dimachkie MM. congenital myasthenic syndromes: a clinical and treatment approach. Curr Treat Options Neurol. 2018; 20(9): 36. **[Crossref]**
9. Ardissone A, Moroni I, Bernasconi P, Brugnoni R. Congenital myasthenic syndrome: phenotypic variability in patients harbouring p. T159P mutation in CHRNE gene. Acta Myologica. 2017; 36(1): 28-32.
10. Webster R, Maxwell S, Spearman H, et al. A novel congenital myasthenic syndrome due to decreased acetylcholine receptor ion-channel conductance. Brain. 2012; 135(Pt 4): 1070-80. **[Crossref]**
11. Tan JS, Ambang T, Ahmad-Annuar A, Rajahram GS, Wong KT, Goh KJ. Congenital myasthenic syndrome due to novel CHAT mutations in an ethnic kadazandusun family. Muscle Nerve. 2016; 53(5): 822-6. **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

Northwestern Med J. 2026;6(3):268-273
DOI: 10.54307/NWMJ.2026.201

CASE REPORT

Mucinous carcinoma of the breast detected incidentally during screening mammography: a case report

Nur Tunç Karabekiroğlu[1], **Gülşah Ünsal Kayar**[2], **Selma Erdoğan Düzcü**[1]

[1]Department of Pathology, Faculty of Medicine, Bolu Abant İzzet Baysal University, Bolu, Türkiye
[2]Department of Pathology, Bolu İzzet Baysal State Hospital, Bolu, Türkiye

Cite as: Tunç Karabekiroğlu N, Ünsal Kayar G, Erdoğan Düzcü S. Mucinous carcinoma of the breast detected incidentally during screening mammography: a case report. Northwestern Med J. 2026;6(3):268-273.

ABSTRACT

Mucinous carcinoma of the breast is a rare histological subtype of invasive breast carcinoma characterized by abundant extracellular mucin production and generally favorable biological behavior. However, because it may present as a circumscribed or partially lobulated lesion on imaging, it can mimic benign breast masses and create a diagnostic challenge. We report the case of a 52-year-old asymptomatic postmenopausal woman in whom a left retroareolar breast lesion was detected incidentally during routine screening mammography. Initial mammography revealed a 15 mm partially irregular lobulated mass categorized as BI-RADS 4. Subsequent ultrasonography and breast magnetic resonance imaging confirmed the lesion, and interval enlargement resulted in a BI-RADS 5 assessment. Excisional biopsy demonstrated a well-circumscribed 2.3-cm gelatinous tumor. Microscopically, more than 90% of the lesion consisted of extracellular mucin pools containing neoplastic epithelial cells arranged in adenoid and cribriform-like structures, consistent with pure mucinous carcinoma. Immunohistochemically, the tumor showed strong, diffuse estrogen receptor and progesterone receptor positivity (95% each), HER2/c-erbB-2 negativity (score 0), MUC2 positivity, and a low Ki-67 proliferation index (<5%), supporting a luminal A-like immunophenotype. Based on the available pathological tumor size, the tumor was categorized as pT2. Regional nodal status could not be pathologically assessed because a sentinel lymph node biopsy was not performed (pNX). The patient underwent surgical excision followed by adjuvant chemotherapy and radiotherapy, and was scheduled for endocrine therapy. No evidence of recurrence or disease progression was observed during 12 months of follow-up. This case highlights the importance of screening mammography in the detection of clinically occult mucinous carcinoma and underscores the value of radiologic-pathologic correlation in lesions with deceptively benign imaging features.

Keywords: mucinous carcinoma, breast, screening mammography, histopathology, immunohistochemistry

Corresponding author: Nur Tunç Karabekiroğlu **E-mail:** tncnur@gmail.com
Received: 03.05.2025 **Accepted:** 12.05.2026 **Published:** 31.07.2026

INTRODUCTION

Mucinous carcinoma of the breast is a rare histological subtype of invasive breast carcinoma characterized by prominent extracellular mucin production and accounts for approximately 1% to 6% of all breast malignancies (1,2). It occurs predominantly in postmenopausal women and is generally associated with a more favorable prognosis than invasive breast carcinoma of no special type (3,4). Histologically, mucinous carcinoma is classified as pure or mixed, with pure mucinous carcinoma defined by a mucinous component constituting at least 90% of the tumor (5). This distinction is clinically important because pure mucinous carcinoma is associated with a lower rate of axillary lymph node involvement and less aggressive biological behavior than mixed mucinous carcinoma (6,7). Radiologically, however, these tumors may appear as circumscribed, oval, or lobulated masses and may therefore mimic benign lesions, leading to potential diagnostic underestimation (1,8,9).

In this report, we present a case of pure mucinous carcinoma detected incidentally during routine screening mammography in a 52-year-old asymptomatic woman and discuss its clinicopathological and radiologic features in light of the current literature.

CASE REPORT

A 52-year-old woman with no active complaints presented to the Cancer Early Diagnosis, Screening and Training Center (KETEM) for routine breast cancer screening. Bilateral mammography revealed a suspicious lesion. A partially irregular lobulated mass measuring 15 mm was detected in the retroareolar region of the left breast and categorized as BI-RADS 4. The patient was referred to our hospital for further evaluation. Ultrasonography and repeat mammography confirmed the lesion, and additional breast magnetic resonance imaging (MRI) was requested (Figure 1). Breast MRI demonstrated a solid lesion in the left

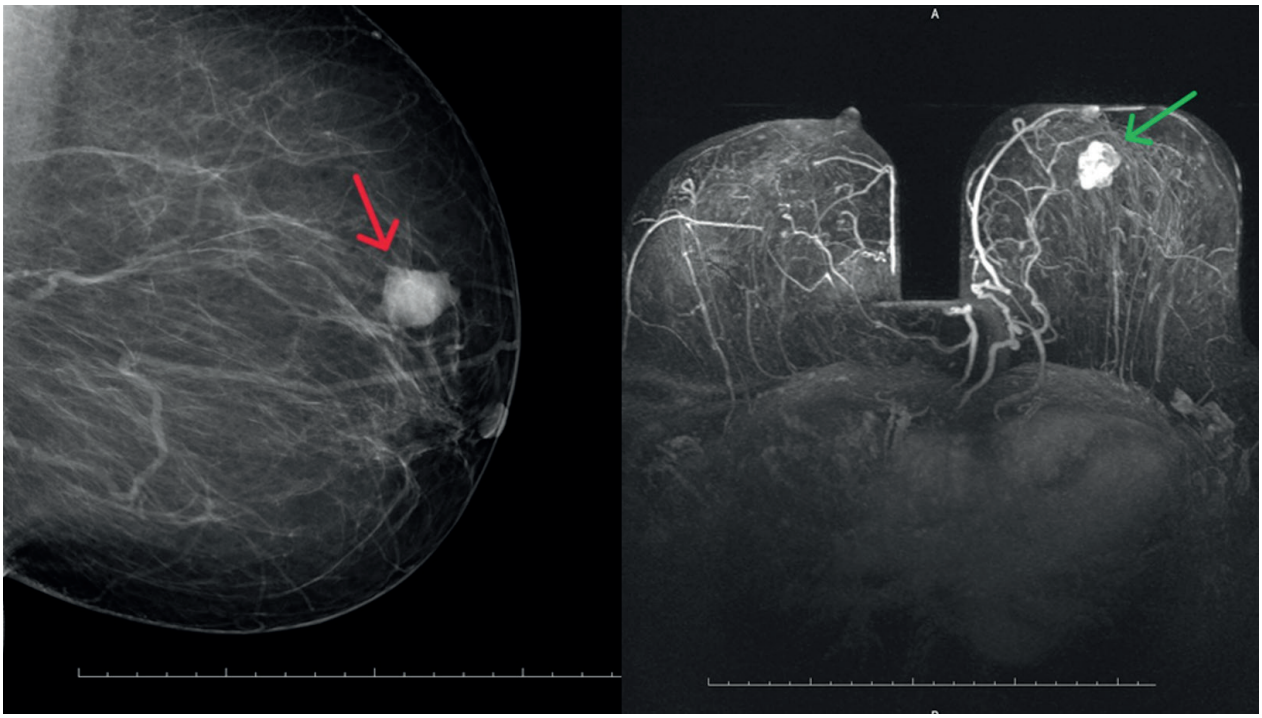


Figure 1. (a) Mammography showing a mass in the left retroareolar region (red arrow). **(b)** Dynamic contrast-enhanced breast MRI showing an irregular enhancing lesion in the left breast (green arrow).

breast with interval enlargement during follow-up, and the lesion was categorized as BI-RADS 5. After informed consent was obtained, an excisional biopsy was performed.

On gross examination, the excised breast tissue measured 8 × 6 × 3.7 cm. Within the specimen, a well-circumscribed lesion measuring 2.3 × 2 × 2 cm was identified. The cut surface was gelatinous and showed focal hemorrhagic areas (Figure 2).

Microscopically, the lesion was composed of tumor cells with mild-to-moderate nuclear atypia arranged in epithelial clusters, adenoid and cribriform-like structures within abundant extracellular mucin pools (Figure 3). Occasional mitotic figures were noted. Peritumoral ductal dilatation and periductal lymphocytic inflammation were present. No lymphovascular invasion, necrosis, or microcalcification was identified.

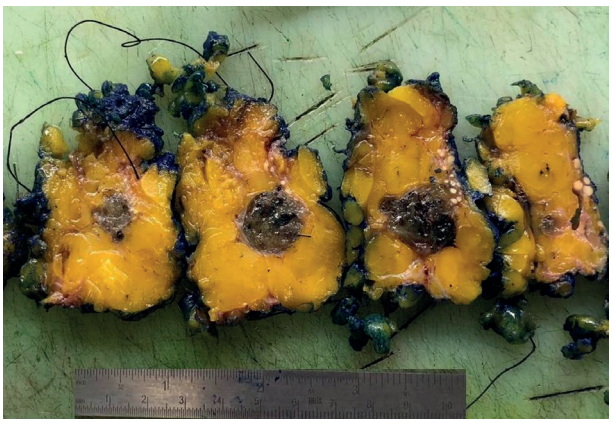


Figure 2. Gross appearance of the tumor showing a well-circumscribed gelatinous lesion with focal hemorrhagic areas.

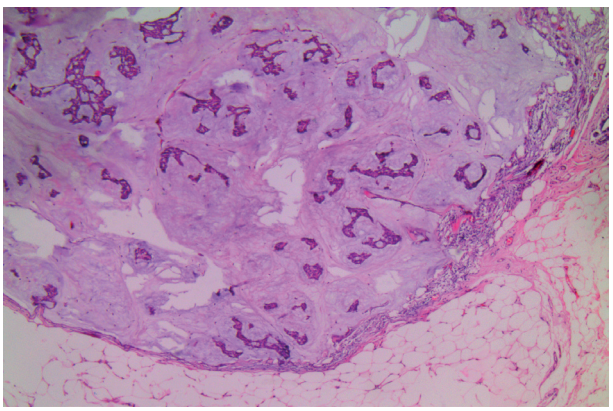


Figure 3. Tumor cells arranged in adenoid and cribriform-like structures floating within abundant extracellular mucin pools (H&E, ×40).

Immunohistochemically, the tumor cells showed strong and diffuse nuclear positivity for estrogen receptor (ER) and progesterone receptor (PR), each in 95% of tumor cells. MUC2 was positive, and E-cadherin expression was retained, although weakly. HER2/c-erbB-2 was negative (score 0). CDX2 and p63 were negative (Figure 4, Figure 5 and Figure 6). The Ki-67 proliferation index was below 5%. Histochemical staining with Alcian blue highlighted the mucinous areas (Figure 7).

Because more than 90% of the tumor consisted of extracellular mucin, the lesion was classified as pure mucinous carcinoma. The ER/PR-positive, HER2/c-erbB-2-negative, low-proliferative profile supported a luminal A-like immunophenotype. Based on the available pathological tumor size of 2.3 cm, the tumor was categorized as pT2. Because sentinel lymph node biopsy was not performed, the regional nodal status could not be pathologically assessed (pNX). The patient was referred to oncology, received four cycles of chemotherapy followed by radiotherapy, and was scheduled for aromatase inhibitor therapy. During 12 months of follow-up, no evidence of recurrence or disease progression was observed.

DISCUSSION

Mucinous carcinoma of the breast is an uncommon special subtype of invasive breast carcinoma that predominantly affects postmenopausal women and generally has a more favorable prognosis than invasive breast carcinoma of no special type (1,3,4). This favorable biological behavior is particularly associated with the pure subtype, which shows a lower rate of nodal metastasis and a lower proliferative index than mixed mucinous carcinoma (6,7).

One of the main diagnostic challenges in mucinous carcinoma is its imaging appearance. Unlike many invasive breast carcinomas, mucinous carcinomas may

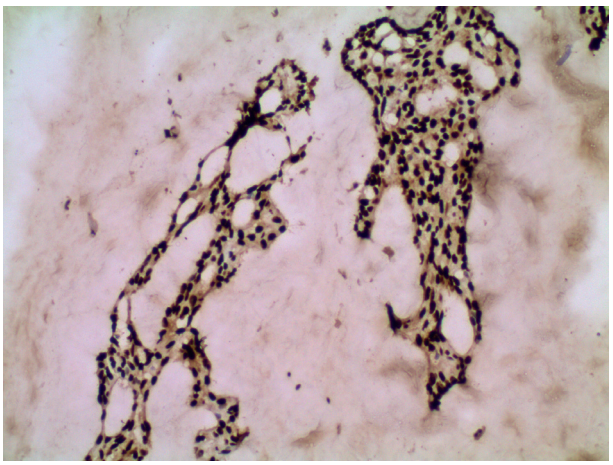


Figure 4. Diffuse nuclear estrogen receptor positivity in tumor cells (ER, ×200).

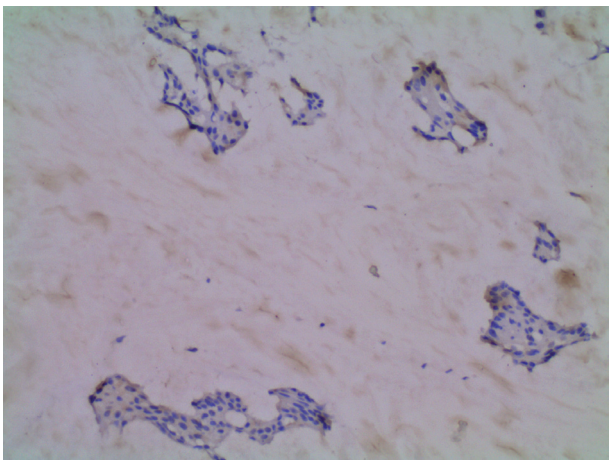


Figure 6. Absence of HER2/c-erbB-2 expression in tumor cells (HER2/c-erbB-2, score 0, ×200).

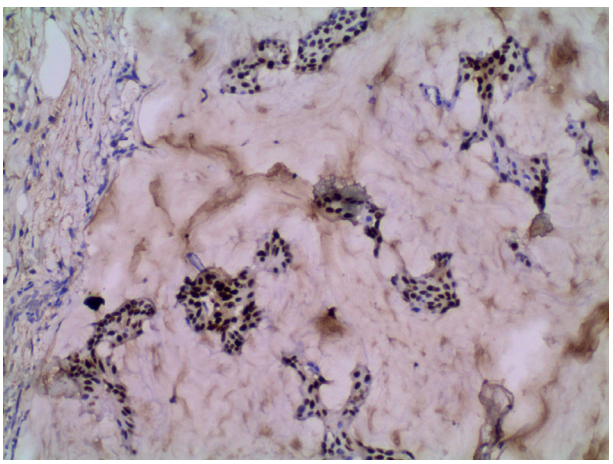


Figure 5. Diffuse nuclear progesterone receptor positivity in tumor cells (PR, ×200).

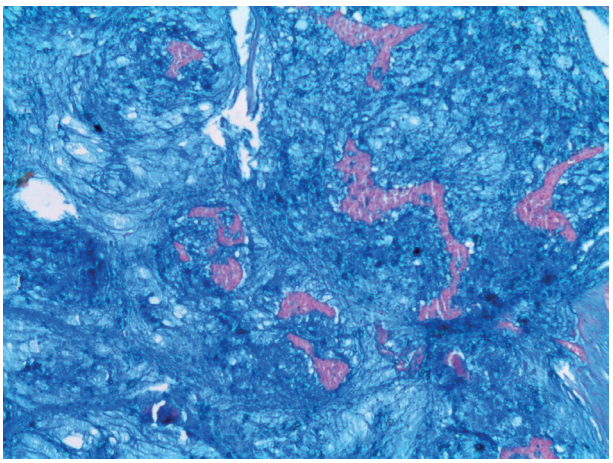


Figure 7. Extracellular mucin highlighted by Alcian blue staining (Alcian blue, ×100).

present as circumscribed, oval, or lobulated masses and may therefore mimic benign lesions (8,9). This issue is especially relevant in asymptomatic patients undergoing screening examinations. In the present case, the patient had no breast-related symptoms, and the lesion was detected incidentally during routine screening mammography, emphasizing the role of screening programs in identifying clinically occult breast malignancies at an early stage.

From a pathological perspective, classification into pure and mixed mucinous carcinoma is essential. Pure mucinous carcinoma is defined by a mucinous component comprising at least 90% of the tumor (1,5). In the current case, more than 90% of the lesion consisted of extracellular mucin pools containing floating epithelial structures, supporting a diagnosis of pure mucinous carcinoma. The gelatinous gross appearance and abundant extracellular mucin on microscopy were also characteristic. The differential diagnosis of mucin-rich breast lesions includes mucocele-like lesions, solid papillary carcinoma with mucin production, mucinous ductal carcinoma in situ, invasive lobular carcinoma with extracellular mucin, and metastatic mucinous carcinomas from extramammary sites; therefore, careful clinicoradiologic correlation

remains essential (5). Recent case reports have also emphasized that, despite variable clinical and radiological presentations, definitive diagnosis relies on the recognition of abundant extracellular mucin and confirmation by appropriate histopathological and immunohistochemical evaluation (10).

The immunophenotype observed in this patient was also typical. Pure mucinous carcinomas are usually hormone receptor-positive and HER2/c-erbB-2-negative and frequently show a luminal A-like profile with a low proliferative index (4,7). In our case, diffuse ER/PR positivity, HER2/c-erbB-2 negativity, and a Ki-67 index below 5% were fully concordant with this pattern and supported the relatively favorable biological behavior generally attributed to pure mucinous carcinoma.

With respect to staging, the tumor measured 2.3 cm on pathological examination, corresponding to pT2. Because sentinel lymph node biopsy was not performed, the regional nodal status could not be pathologically assessed and was therefore designated pNX. At 12 months of follow-up, there was no clinical or radiological evidence of recurrence or disease progression.

Treatment of mucinous carcinoma is individualized according to tumor size, nodal status, receptor profile, and multidisciplinary assessment (1,4,7). Surgery remains the cornerstone of treatment, while endocrine therapy is particularly relevant in hormone receptor-positive tumors. In the present case, the patient underwent excision followed by chemotherapy, radiotherapy, and planned aromatase inhibitor therapy. Overall, this case contributes to the literature by documenting an asymptomatic, screen-detected pure mucinous carcinoma and by emphasizing the diagnostic value of radiologic-pathologic correlation in lesions with deceptively benign imaging features.

CONCLUSION

Pure mucinous carcinoma of the breast is a rare invasive carcinoma that may remain clinically silent and may radiologically resemble a benign lesion despite harboring malignancy. This case underscores the importance of screening mammography in the early detection of asymptomatic breast cancer and highlights the central role of histopathological and immunohistochemical evaluation in establishing the correct diagnosis. Careful radiologic-pathologic correlation, accurate recognition of the pure subtype, and appropriate reporting of tumor extent and follow-up findings improve both patient management and the scientific value of case reports involving this uncommon breast carcinoma subtype. In the present case, the absence of recurrence or disease progression during 12 months of follow-up further supports the favorable short-term clinical course of pure mucinous carcinoma.

Ethical approval

This case report was conducted in accordance with institutional ethical guidelines. Written informed consent was obtained from the patient.

Author contribution

Concept: NTK, GÜK, SED; Design: NTK, GÜK, SED; Data Collection or Processing: NTK, GÜK, SED; Analysis or Interpretation: NTK, GÜK, SED; Literature Search: NTK, GÜK; Writing: NTK, GÜK. All authors reviewed the results and approved the final version of the article.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Kuzmiak CM, Calhoun BC. Pure mucinous carcinoma of the breast: radiologic-pathologic correlation. J Breast Imaging. 2023; 5(2): 180-7. [Crossref]
2. Komenaka IK, El-Tamer MB, Troxel A, et al. Pure mucinous carcinoma of the breast. Am J Surg. 2004; 187(4): 528-32. [Crossref]
3. Avisar E, Khan MA, Axelrod D, Oza K. Pure mucinous carcinoma of the breast: a clinicopathologic correlation study. Ann Surg Oncol. 1998; 5(5): 447-51. [Crossref]
4. Marrazzo E, Frusone F, Milana F, et al. Mucinous breast cancer: a narrative review of the literature and a retrospective tertiary single-centre analysis. Breast. 2020; 49: 87-92. [Crossref]

5. Joneja U, Palazzo J. The spectrum of mucinous lesions of the breast. Arch Pathol Lab Med. 2023; 147(1): 19-29. **[Crossref]**

6. Li N, Li JW, Qian Y, et al. Axillary lymph node metastasis in pure mucinous carcinoma of breast: clinicopathologic and ultrasonographic features. BMC Med Imaging. 2024; 24(1): 108. **[Crossref]**

7. Hamdy O, Shetiwy M, Saber MM, et al. Mucinous carcinoma of the breast: epidemiological, clinical, and prognostic characteristics; a single-center experience. Discov Oncol. 2025; 16(1): 1396. **[Crossref]**

8. Zhang L, Jia N, Han L, Yang L, Xu W, Chen W. Comparative analysis of imaging and pathology features of mucinous carcinoma of the breast. Clin Breast Cancer. 2015; 15(2): e147-54. **[Crossref]**

9. Bitencourt AGV, Graziano L, Osório CABT, et al. mri features of mucinous cancer of the breast: correlation with pathologic findings and other imaging methods. AJR Am J Roentgenol. 2016; 206(2): 238-46. **[Crossref]**

10. Rey J, Shukla S, Acharya S, Gadkari P, Sihag S. Mucinous carcinoma of the breast: a case report. Cureus. 2024; 16(3): e56515. **[Crossref]**

NORTHWESTERN MEDICAL JOURNAL

www.nwmedj.org

ISSN: 2979-9538