e-ISSN: 2149-3189

European Research Journal

Volume 12 Issue 3 March 2026

Available at https://dergipark.org.tr/en/pub/eurj

Published by Nicaea Medical Publishing

The European Research Journal

Aim and Scope

The European Research Journal (EuRJ) is an international, independent, double-blind peer reviewed, Open Access and online publishing journal, which aims to publish papers on all the related areas of basic and clinical medicine.

Editorial Board of the European Research Journal complies with the criteria of the International Council of Medical Journal Editors (ICMJE), the World Association of Medical Editors (WAME), and Committee on Publication Ethics (COPE).

The journal publishes a variety of manuscripts including original research, case reports, invited review articles, technical reports, how-to-do it, interesting images and letters to the editor. The European Research Journal has signed the declaration of the Budapest Open Access Initiative. All articles are detected for similarity or plagiarism. Publication language is English.

EuRJ recommends that all of our authors obtain their own ORCID identifier which will be included on their article.

The journal is published (January, February, March, April, May, June, July, September, October, November and December).

Abstracting and Indexing

The journal is abstracted and indexed with the following: ULAKBIM TR Index (ULAKBİM TR DİZİN), NLM Catalog (NLM ID: 101685727), Google Scholar (h-index: 15), EMBASE, ProQuest Central, EBSCO Academic Search Ultimate, J-Gate, EZB, TURK MEDLINE, Turkish Citation Index, ResearhGate, SOBIAD, ScienceGate, Publons, (Clarivate Web of Science)

Publisher

The European Research Journal (EuRJ)
Nicaea Medical Publishing
Konak Mh. Kudret Sk. Şenyurt İş Mrk. Blok No:6 İç kapı no: 3 Nilüfer/Bursa-Türkiye
info@nicaeamp.com

https://dergipark.org.tr/en/pub/eurj
https://www.nicaeamp.com
Note: The European Research Journal has changed publishers with its March 2026 issue

e-ISSN: 2149-3189

EDITORIAL BOARD

EDITOR-IN-CHIEF

EDITORS

OWNER ON BEHALF OF THE NICAEA MEDİCAL PUBLISHING

ASSISTANT EDITOR

SECTION EDITORS

Meliha KASAPOGLU AKSOY, MD.,
Associate Professor,
University of Health Sciences, Bursa Yuksek Ihtisas Training & Research Hospital,
Department of Physical Therapy and Rehabilitation,
Bursa, Türkiye

Arda ISIK, MD.,
Associate Professor,
Medeniyet University School of Medicine,
Department of General Surgery,
Istanbul, Türkiye

Kadir Kaan OZSIN, MD.,
Associate Professor,
University of Health Sciences, Bursa Yuksek Ihtisas Training & Research Hospital,
Department of Cardiovascular Surgery,
Bursa, Türkiye

Cihan AYDIN, MD.,
Associate Professor,
Tekirdağ Namık Kemal University, Faculty of Medicine,
Department of Cardiology,
Tekirdağ, Türkiye

Sayad KOCAHAN, PhD.,
Professor,
University of Health Sciences, Gülhane Medical Faculty,
Department of Physiology,
Ankara, Türkiye

Gokhan OCAKOGLU, PhD.,
Professor,
Uludag University School of Medicine,
Department of Biostatistics,
Bursa, Türkiye

Nurullah DOGAN, MD.,
Professor,
İstanbul Atlas University School of Medicine,
Department of Radiology,
Bursa, Türkiye

Ömer Faruk KARATAS, PhD.,
Professor,
Erzurum Technical University,
Department of Molecular Biology and Genetics,
Erzurum, Türkiye

Serhat YALÇINKAYA, PhD.,
Associate Professor,
Priavate Bursa NEV Health Group,
Department of Thoracic Surgery,
Bursa, Türkiye

Evren DİLEKTAŞLI, MD.,
Professor,
VM Medical Park Bursa Hospital
Department of General Surgery,
Bursa, Türkiye

Nurcan ÖZYAZICIOĞLU,
Professor,
Department Nursing and Health Sciences
Bursa Uludağ Unibersity
Bursa, Türkiye

Burcu DİNÇGEZ, MD.,
Professor,
University of Health Sciences, Bursa Yuksek Ihtisas Training & Research Hospital,
Department of Gynecology and Obstetrics,
Bursa, Türkiye

Yenal DUNDAR, MD.,
Consultant Psychiatrist
University of Liverpool,
Liverpool, UK

Başar CANDER, MD.,
Professor,
Bezmialem Vakif University,
Department of Emergency Medicine
İstanbul, Türkiye

Aylin COLPAN, MD.,
Associate Professor,
Jefferson University-Lehigh Valley Hospital
Department of Infectious Diseases
Allentown, ABD

Sanjiv RAMPAL, MD.,
Associate Professor,
International Medical University
Department of Orthopaedics, Sports Medicine
Kuala Lumpur, Malaysia

Table of Contents

Original Articles

The European Research Journal 2026;12(3):289-296 Original Article

DOI: https://doi.org/10.18621/eurj.1830453 *Perinatology*

Corpus Callosum Anomalies: Prenatal Diagnosis, Genetic Findings, and Perinatal Outcomes

Tuğçe Arslanoğlu[1], Sezin Uludağ[1], Deniz Kanber Açar[1], Alev Ateş Aydın[1]

[1]*Department of Perinatology, University of Health Sciences, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye*

ABSTRACT

Objectives: The aim of this study was to evaluate the subtype distribution of prenatally diagnosed corpus callosum (CC) anomalies. We also assessed associated central nervous system (CNS) and extracranial anomalies, genetic findings, perinatal outcomes, and the contribution of ultrasonography and fetal MRI to prenatal diagnosis.

Methods: This retrospective, single-center study was conducted in a tertiary perinatology clinic between 2016 and 2025. A total of 124 fetuses with prenatally diagnosed CC anomalies were included. Diagnosis was based on obstetric ultrasound, and fetal magnetic resonance imaging (MRI) was performed when ultrasound findings were inconclusive. Cases were classified by subtype, and associated CNS or extracranial anomalies, genetic test results, and pregnancy outcomes were recorded.

Results: The study included 124 fetuses diagnosed with CC anomalies during the prenatal period. The mean gestational age at diagnosis was 26.6±4.7 weeks. Pregnancy resulted in termination in 67.7% of cases and live birth in 32.3%. In cases that ended in termination, diagnosis and delivery occurred at earlier gestational weeks, and birth weights were lower. Additional anomalies were present in 42.7% of cases, most commonly involving the CNS and the heart. Complete CC agenesis was the most frequent subtype (54%). Genetic testing was more often performed in the termination group and identified chromosomal abnormalities such as trisomy 18, trisomy 13, and 22q11 deletion. Fetal MRI was performed in 45 cases, confirming the ultrasound diagnosis in 36 and leading to diagnostic revision in 9 cases.

Conclusions: This study summarizes our experience in the evaluation of pregnancies diagnosed with CC anomalies. In our cohort, prognosis was mainly influenced by whether the anomaly was isolated, along with fetal MRI and genetic test findings. Overall, our approach was similar to that reported in the literature. Nevertheless, each case required individual assessment, and counseling was adjusted according to the clinical findings.

Keywords: Corpus Callosum Anomalies, Fetal MRI, Prenatal Neurosonography, Genetic Anomalies, Perinatal Outcome

The Corpus Callosum (CC) is the largest bundle of white matter connecting the right and left hemispheres of the brain and plays an important role in the transfer of information between the hemispheres. CC agenesis is defined as the partial or complete failure of this structure to

Received: November 25, 2025 **Accepted:** December 28, 2025 **Available Online:** January 3, 2026 **Published:** March 1, 2026

How to cite this article: Arslanoğlu T, Uludağ S, Açar DK, Ateş Aydın A. Corpus Callosum Anomalies: Prenatal Diagnosis, Genetic Findings, and Perinatal Outcomes. Eur Res J. 2026;12(3):289-296. doi: 10.18621/eurj.1830453

Corresponding author: Tuğçe Arslanoğlu, MD., Phone: +90 212 404 15 00, E-mail: drtugcetunc@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

develop. The embryological development process is completed between approximately the 11th and 20th weeks of pregnancy [1]. The prevalence of CC agenesis in the population varies between 1.8 and 70 per 10,000 live births, depending on the diagnostic methods used and the population studied [2]. Prenatal diagnostic methods have become more widespread, facilitating the detection of these cases, especially since isolated cases can be asymptomatic after birth [3].

The corpus callosum regulates the transfer of cognitive, motor, and sensory information between the hemispheres. Postnatal outcomes of CC anomalies show a wide clinical spectrum. While isolated cases may be asymptomatic, some cases may present with developmental delay, learning difficulties, social communication problems, or epileptic seizures [4]. The prognosis is significantly worse when CC anomalies are associated with central nervous system (CNS) anomalies or genetic syndromes; severe motor sequelae, profound cognitive impairment, and high mortality rates have been reported [5].

The primary method in prenatal evaluation is ultrasound. The 2025 Delphi consensus defines the inclusion of the midsagittal section as part of routine examination as the standard approach [6]. When suspicious findings are observed, it is recommended that cases be referred to specialized centers and evaluated with fetal magnetic resonance imaging (MRI). MRI provides additional information regarding both the structural characteristics of the corpus callosum and associated CNS anomalies. Prenatal diagnosis is important not only for establishing the correct diagnosis but also for managing the pregnancy and providing counseling to the family.

The purpose of this study was to describe the subtype distribution, associated anomalies, genetic findings, and perinatal prognosis of fetuses diagnosed with CC anomalies during the prenatal period. We also aimed to explore the role of ultrasound and fetal MRI in prenatal diagnosis and clinical counseling.

METHODS

This retrospective, single-center study was performed in a tertiary perinatology clinic. A total of 124 fetuses diagnosed with CC anomalies during the prenatal period between January 2016 and January 2025 were included in the evaluation. All prenatally diagnosed cases were included, and the final study population consisted of 124 fetuses. Prenatal imaging findings, genetic results, and perinatal outcomes were evaluated in this cohort. The study was approved by the Clinical Research Ethics Committee of İstanbul S.B.Ü. Kanuni Sultan Süleyman Training and Research Hospital (Decision/Protocol No: 2025.07.177; date: 03.07.2025). The study was conducted in compliance with the principles of the Declaration of Helsinki, and all patient data were anonymized.

All cases were examined during obstetric ultrasonography performed in the second or third trimester. The evaluations were carried out by perinatology specialists using a Voluson E6 ultrasound system (GE Healthcare, Zipf, Austria). Fetuses in cephalic presentation were evaluated with the transvaginal probe RIC5-9 (5–9 MHz); for patients who declined a vaginal examination, the abdominal convex probe C2-9-D (2–5 MHz) was used.

The midsagittal sections were used as the basis for evaluating the presence and morphology of the CC; all segments of the CC (rostrum, genu, body, splenium) and the pericallosal artery were examined in detail. In the axial plane, the lateral ventricles, third ventricle, interhemispheric fissure, and cavum septi pellucidi (CSP) were assessed; the presence, width, and continuity of the CSP were recorded, and the cavum vergae was also examined when necessary. In the coronal plane, the symmetry of the cerebral hemispheres, the CSP, and the CC were observed together; in addition, the cavities and commissural structures were evaluated. Advanced fetal neurosonography was performed in cases where suspicious findings were detected. In our cohort, quantitative measurements of CC length and its subsegments were not performed systematically; instead, the diagnosis was primarily based on morphological assessment, complemented by confirmation through fetal MRI.

Fetal MRI was used for additional evaluation when ultrasound findings suggested agenesis, hypoplasia, or dysgenesis of the CC. Imaging was performed using a 1.5-Tesla Siemens Magnetom Aera device located in our hospital's Radiology Clinic. Sedation was not administered to pregnant women during MRI scanning. MRI examinations were scheduled shortly after diagnosis by ultrasound. Most

cases were evaluated between 24 and 32 weeks of gestation; imaging was also performed in earlier weeks in cases where the diagnosis remained uncertain or where family counseling could not be delayed. The primary aim of this approach was to confirm sonographic findings and rule out other associated CNS anomalies. Genetic counseling was provided to families in all cases where a diagnosis was made, and amniocentesis, chorionic villus sampling or cordocentesis was recommended when necessary. Families were informed that genetic testing evaluates chromosomal abnormalities and microdeletion/duplication syndromes. In the postnatal period, information was provided regarding the neurodevelopmental course, learning difficulties, risk of epilepsy, and follow-up of associated structural anomalies. Thus, comprehensive counseling covering both the prenatal and postnatal periods was provided. This approach is consistent with both our clinic's practices and current International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) recommendations [7].

The mother's age, gravida, parity, history of abortion, height, weight, and BMI values, as well as the week of diagnosis, week of delivery, and mode of delivery, were obtained from patient records. Neonatal data including sex, birth weight, and Apgar scores at 1 and 5 minutes, were recorded. CC anomalies were classified as complete, partial agenesis, hypoplasia, and dysgenesis. Associated findings included ventriculomegaly, cardiac defects, other CNS anomalies, and extracranial anomalies. All families received genetic counseling, and amniocentesis, chorionic villus sampling (CVS), or cordocentesis were recommended when indicated. Karyotype or chromosomal microarray analysis (CMA) results were evaluated in cases where testing was accepted. Fetal MRI findings were recorded in prenatal diagnosis and compared with ultrasound. Pregnancy outcomes were classified as live birth, intrauterine loss, or termination.

TABLE 1. Maternal Demographic and Obstetric Characteristics

	Mean±SD	Median (Min–Max)
Age (years)	34.98±6.59	34 (23-52)
Gravida	2.93±2.07	2 (1-10)
Parity	1.41±1.53	1 (0-7)
Abortions	0.52±1.01	0 (0-7)
Height (cm)	162.6±4.71	163 (150-175)
Weight (kg)	84.14±11.37	85 (70-120)
Body mass index (kg/m²)	31.94±5.07	31.2 (24.2-49.1)

SD, standard deviation; Min, minimum; Max, maximum.

Statistical Analysis

For statistical analyses, the Number Cruncher Statistical System (NCSS) 2007 software (Kaysville, Utah, USA) was used. While evaluating the study data, descriptive statistical methods (Mean, Standard Deviation, Median, Frequency, Ratio, Minimum, and Maximum) were employed, and the distribution of the data was assessed using the Shapiro-Wilk Test. The Mann-Whitney U test was used to compare quantitative data between two groups. Chi-square analysis was utilized to examine the relationship between qualitative data. Statistical significance was evaluated at $P<0.05$ levels.

RESULTS

A total of 124 fetuses diagnosed with CC anomalies in the prenatal period between 2016 and 2025 were included in this study. Maternal demographic and obstetric characteristics are summarized in Table 1. The mean gestational age at diagnosis was 26.6±4.7 weeks. The mean gestational age at delivery was 31.3±5.7 weeks and the mean birth weight was 2039.7±1136.4 g (Table 2).

Overall, 84 (67.7%) pregnancies resulted in termination and 40 (32.3%) in live birth. Compared with the live-birth group, pregnancies ending in termination were diagnosed at earlier gestational weeks, delivered at lower gestational ages, and had lower birth weights. Maternal age, gravida, parity, history of abortion, body mass index, and lateral ventricular width did not differ between the groups (Table 3).

Genetic testing was performed more frequently in

TABLE 2. Prenatal Diagnostic Characteristics and Perinatal Outcomes in Corpus Callosum Anomalies

	Mean±SD	Median (Min-Max)
Gestational age at diagnosis (weeks)	26.61±4.72	26.29 (18-36.86)
Birth weight (g)	2039.67±1136.35	1970 (115-3940)
Gestational age at delivery (weeks)	31.3±5.71	30.43 (22.43-40.86)
Lateral ventricular width (mm)	13.66±3.95	13 (5.5-25)
Apgar 1st min.	6.21±1.91	7 (1-9)
Apgar 5th min	8.17±1.7	9 (3-10)

SD, standard deviation; Min, minimum; Max, maximum; min, minute.

the termination group, and all three abnormal chromosomal results (trisomy 18, trisomy 13, and 22q11 deletion) occurred in this group. The distribution of associated anomalies was similar between the two groups. CNS anomalies were the most common in both groups, while gastrointestinal anomalies were more frequent among live-born infants. Detailed clinical and genetic data are presented in Table 4.

Complete CC agenesis was the most common subtype, accounting for 54% (n=67) of all cases, followed by partial CC agenesis (21%, n=26), CC

TABLE 3. Comparison of Clinical Characteristics Between Termination and Live Birth Groups

		n	Mean±SD	Median (Min-Max)	P-value*
Age (years)	Termination	84	35.3±6.58	34.5 (23-52)	0.457
	Live birth	40	34.33±6.65	34 (23-48)	
Gravida	Termination	84	2.78±1.92	2 (1-8)	0.178
	Live birth	40	3.26±2.34	3 (1-10)	
Parity	Termination	84	1.34±1.49	1 (0-6)	0.372
	Live birth	40	1.56±1.6	1 (0-7)	
Abortions	Termination	84	0.44±0.82	0 (0-4)	0.218
	Live birth	40	0.69±1.3	0 (0-7)	
BMI (kg/m²)	Termination	84	31.23±4.44	30.5 (24.2-46.7)	0.054
	Live birth	40	33.43±5.98	32.35 (25.3-49.1)	
Gestational age at diagnosis (weeks)	Termination	84	25.86±4.3	25.71 (18-35.57)	**0.024**
	Live birth	40	28.18±5.21	28.93 (20.43-36.86)	
Birth weight (g)	Termination	84	1082.79±736.12	732.5 (115-3285)	**0.001**
	Live birth	40	2873.87±665.31	3030 (1320-3940)	
Gestational age at delivery (weeks)	Termination	84	28.19±3.87	27.36 (22.43-38.71)	**0.001**
	Live birth	40	37.84±2.47	38.71 (31.43-40.86)	
Lateral ventricular width (mm)	Termination	84	12.99±3.49	12.75 (5.5-22.5)	0.089
	Live birth	40	15.07±4.54	14.5 (9-25)	

BMI, body mass index; SD, standard deviation; Min, minimum; Max, maximum.
*Mann-Whitney U test. Statistically significant P-values are shown in bold.

TABLE 4. Comparison of Clinical, Genetic, and Radiological Findings According to Termination and Live Birth Groups

		Grup			P-value*
		Termination	Live Birth	Total	
Fetal growth restriction (FGR)	Present	14 (16.7%)	5 (12.5%)	19 (15.3%)	0.737
	Absent	70 (83.3%)	35 (87.5%)	105 (84.7%)	
Karyotype	Absent	51 (60.7%)	33 (82.5%)	51 (60.7%)	**0.042**
	Normal	30 (35.7%)	7 (17.5%)	30 (35.7%)	
	Anormal	3 (3.6%)	0 (0%)	3 (3.6%)	
Chromosomal microarray - CMA	Absent	51 (60.7%)	33 (82.5%)	51 (60.7%)	**0.042**
	Normal	30 (35.7%)	7 (17.5%)	30 (35.7%)	
	Anormal	3 (3.6%)	0 (0%)	3 (3.6%)	
MRI	Confirmed	22 (78.6%)	14 (82.4%)	36 (80%)	0.538
	Revised	6 (21.4%)	3 (17.6%)	9 (20%)	
Fetal gender	Male	18 (52.9%)	23 (59%)	41 (56.2%)	0.778
	Female	16 (47.1%)	16 (41%)	32 (43.8%)	
Presence of associated anomalies	Absent	45 (53.6%)	26 (65%)	71 (57.3%)	0.313
	Present	39 (46.4%)	14 (35%)	53 (42.7%)	
Type of associated anomalies	Central nervous system	14 (35.9%)	4 (28.6%)	18 (34%)	0.181
	Cardiac	8 (20.5%)	2 (14.3%)	10 (18.9%)	
	Urinary system	4 (10.3%)	2 (14.3%)	6 (11.3%)	
	Facial anomaly	5 (12.8%)	0 (0%)	5 (9.4%)	
	Gastrointestinal anomaly	2 (5.1%)	4 (28.6%)	6 (11.3%)	
	Skeletal system	3 (7. 7%)	0 (0%)	3 (5.7%)	
	Multiple anomaly	3 (7.7%)	2 (14.3%)	5 (9.4%)	
Subtype of corpus callosum anomaly	CC Hypoplasia	6 (7.1%)	7 (17.5%)	13 (10.5%)	0.229
	CC Dysgenesis	11 (13.1%)	7 (17.5%)	18 (14.5%)	
	Complete CC agenesis	47 (56%)	20 (50%)	67 (54%)	
	Partial CC agenesis	20 (23.8%)	6 (15%)	26 (21%)	

Data are shown as n (%). MRI, magnetic resonance imaging.
*Chi-Square test. Statistically significant P-values are shown in bold.

dysgenesis (14.5%, n=18), and CC hypoplasia (10.5%, n=13). Subtype distribution did not differ significantly between the termination and live-birth groups (P=0.229). Fetal MRI was performed in 45 (36%) fetuses. The ultrasound-based diagnosis was confirmed in 36 (80%) cases, while MRI led to revision of the callosal subtype or detection of additional CNS anomalies in 9 (20%) cases.

DISCUSSION

In our study, most diagnoses of CC agenesis were made during the second trimester. The mean gestational age at diagnosis was approximately 26 weeks, with the earliest at 18 weeks. Similarly, Shakes *et al*. [7] reported that CC agenesis is most often diagnosed in the second or third trimester, while detection before 17

weeks remains particularly challenging. The relatively late gestational age at diagnosis in our series is thus in line with these embryological and imaging limitations and also reflects real-world referral patterns to our tertiary center.

The importance of the midsagittal ultrasound plane in prenatal screening has been increasingly emphasized in recent years. The 2025 Delphi consensus recommended incorporating midsagittal imaging of the fetal brain as part of routine obstetric ultrasound. In our series, when findings such as absence of the cavum septum pellucidum or abnormal ventricular configuration were observed, cases were referred to our perinatology unit for advanced neurosonography and, when indicated, fetal MRI. Most cases were diagnosed between 20 and 28 weeks. Many families reached our center at these relatively advanced weeks, which left limited time for MRI, genetic tests, and detailed counseling. This time pressure likely contributed to the higher termination rate. Earlier detection generally allows families more space to process information and consider their options.

Although ultrasound is our primary diagnostic tool for CC anomalies, definitive diagnosis is not always straightforward in cases of partial CC agenesis. Therefore, fetal MRI was additionally performed in 45 fetuses, and the diagnosis was confirmed in 36 (80%) cases. In 9 (20%) cases, MRI findings led to a revision of the diagnosis; some were reclassified as CC hypoplasia or CC dysgenesis, while in some cases, additional CNS anomalies not detected by ultrasound were identified.

Previous studies have also shown that the concordance between fetal ultrasound and MRI is generally high (approximately 85-95%) and that differences are mostly due to subtype differentiation or the detection of additional anomalies [8].

In our cohort, complete CC agenesis was the most common subtype (54%). Partial CC agenesis (21%), CC dysgenesis (14.5%), and CC hypoplasia (10.5%) followed. This general pattern was similar in both the termination and continuation groups. Complete CC agenesis remained the dominant form regardless of outcome. These findings are consistent with previous reports [9].

The lack of a clear difference between groups suggests that subtype alone does not guide families' decisions. Other issues seem to play a bigger role. These include the gestational week at diagnosis, the presence of additional anomalies, how the counseling unfolded, and the families' sociocultural or religious views.

Colpocephaly is a common finding in CC agenesis and is defined as enlargement of the posterior horns of the lateral ventricles. In our series, the mean atrial width was approximately 13.7 mm, and most fetuses had a ventricular width greater than 10 mm. The 2020 Society for Maternal-Fetal Medicine (SMFM) guidelines also state that atrial width may be normal in early pregnancy, with enlargement typically occurring later.

In our study, although ventricular enlargement was higher in the live birth group, the difference was not statistically significant. This finding suggests that ventricular size alone does not determine pregnancy outcome. Associated anomalies and genetic findings also appear to influence the decision-making process. Consistent with the literature, ventricular enlargement should be evaluated in conjunction with other pathologies rather than as an independent prognostic indicator [10-12].

In our series, 57.3% of CC anomalies were isolated, and 42.7% were non-isolated. This ratio is consistent with the 40–60% range reported in the literature. Most non-isolated cases had additional CNS anomalies (70%). The most common were interhemispheric cysts, vermis hypoplasia, and hydrocephalus. Among extracranial anomalies, the most common finding was cardiac anomalies (18.9%). The presence of at least one extracranial anomaly in approximately half of the non-isolated cases demonstrates why this distinction is clinically important. Genetic disorders were largely observed in the non-isolated group. The most common genetic disorders were trisomy 18 and 13. During the counseling process, families were counseled by evaluating MRI, neurosonography, and genetic results together. Discussions focused particularly on the uncertainty of the prognosis and what the distinction between isolated and non-isolated cases meant. We believe this counseling influenced the decision of some families who were undecided.

It was also evident in our series that the prognosis for isolated cases was generally better. However, as noted in the literature, neurodevelopmental problems can still occur in some cases that appear isolated. Therefore, although the term "isolated" is a positive

sign, it does not provide complete assurance; it is important to inform families about the need for careful counseling and long-term follow-up [13].

Genetic testing is essential in CC anomalies to understand both the cause and prognosis. One-third of the cases in our series underwent invasive testing; most were normal. However, trisomy 18, trisomy 13, and 22q11 deletion were detected in three fetuses, and all of these pregnancies were terminated. This situation shows that genetic disorders are more common in non-isolated cases and significantly affect the decision-making process [14]. The literature reports chromosomal anomaly rates within a wide range (3–40%), with trisomy 18 and 13 being the most common. Since additional findings may be detected by microarray or exome analysis even in cases that appear isolated, the SMFM and Delphi recommendations support performing amniocentesis with microarray analysis in all cases.

Neurodevelopmental course is an important determinant of the clinical significance of CC anomalies. In isolated cases, the majority of children (70–75%) show normal development or experience only mild difficulties. In contrast, approximately one-quarter may develop learning difficulties, language delays, or problems with social communication. More severe neurodevelopmental presentations - such as epilepsy or significant developmental delay - are less common and are often associated with accompanying white matter abnormalities or underlying genetic causes. Recent data confirm that truly isolated cases represent the group with the best prognosis. This information was also fundamental to the counseling process in our study and influenced the clinical management process, particularly in cases with late diagnosis. [15-17].

Strengths and Limitations

The most significant strength of this study is that it comparatively evaluated pregnancies resulting in termination and those continuing to live birth, based on a large patient series including cases diagnosed with corpus callosum CC anomalies during the prenatal period. Conducting the study at a single tertiary referral center allowed for the evaluation of diagnosis, follow-up, and clinical decisions using a standardized approach. The evaluation of prenatal imaging findings in conjunction with fetal MRI, genetic testing, and perinatal outcomes contributes to the existing literature on clinical follow-up and family counseling.

A limitation of our study is that the length and width of the CC were not measured. The 2025 Delphi consensus states that these measurements are not necessary in routine screening and should only be considered if there is suspicion regarding the shape. Furthermore, there is no accepted reference curve or international consensus for length and thickness. Although the lack of standardized measurements makes interpretation somewhat difficult, the isolated nature of the cases and our focus on genetic analysis enabled us to obtain clinically meaningful results.

CONCLUSION

In this study, we describe key considerations in the evaluation of pregnancies diagnosed with CC anomalies. The distinction between isolated and non-isolated cases, together with MRI findings and genetic test results, played an important role in clinical decision-making in our cohort. Our experience is consistent with the literature, but each case should be evaluated individually.

Ethics Approval and Consent to Participate

This study was approved by the University of Health Sciences İstanbul Kanuni Sultan Süleyman Training and Research Hospital Scientific Research Ethics Committee (Decision No: KAEK/2025.07.177; date: 03.07.2025). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Informed consent was waived because of the retrospective nature of the study and the analysis used anonymous clinical data.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: TA; Study Design: TA, SS, DKA; Supervision: TA; TA, SS; Materials: N/A; Data Collection and/or Processing: TA, SS, DKA, AAA; Statistical Analysis and/or Data Interpretation: TA, SS; Literature Review: TA, SS; Manuscript Preparation: TA; and Critical Review: TA, SS, DKA, AAA.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Bartek V, Szabó I, Harmath Á, et al. Prenatal and Postnatal Diagnosis and Genetic Background of Corpus Callosum Malformations and Neonatal Follow-Up. Children (Basel). 2024;11(7):797. doi: 10.3390/children11070797.
2. Huang R, Chen J, Hou X, et al. Retrospective analysis of the prognostic factors of fetal corpus callosum dysplasia. BMC Pregnancy Childbirth. 2024;24(1):101. doi: 10.1186/s12884-024-06300-w.
3. Tsai P, Shinar S. Agenesis of the corpus callosum: What to tell expecting parents? Prenat Diagn. 2023;43(12):1527-1535. doi: 10.1002/pd.6447.
4. Siffredi V, Anderson V, Leventer RJ, Spencer-Smith MM. Neuropsychological profile of agenesis of the corpus callosum: a systematic review. Dev Neuropsychol. 2013;38(1):36-57. doi: 10.1080/87565641.2012.721421.
5. Sun H, Li K, Wang L, et al. Fetal agenesis of the corpus callosum: Clinical and genetic analysis in a series of 40 patients. Eur J Obstet Gynecol Reprod Biol. 2024;298:146-152. doi: 10.1016/j.ejogrb.2024.05.005.
6. Corroenne R, Paladini D, Papastefanou I, et al; Collaborators. Prenatal evaluation, diagnosis and management of fetal corpus callosal abnormalities: international Delphi consensus. Ultrasound Obstet Gynecol. 2025;66(5):582-588. doi: 10.1002/uog.70003.
7. Shakes P, Cashin A, Hurley J. Scoping Review of the Prenatal Diagnosis of Agenesis of the Corpus Callosum. J Obstet Gynecol Neonatal Nurs. 2020;49(5):423-436. doi: 10.1016/j.jogn.2020.06.003.
8. Rollins NK. Diffusion imaging of the congenitally thickened corpus callosum. AJNR Am J Neuroradiol. 2013;34(3):660-665. doi: 10.3174/ajnr.A3245.
9. Glenn OA, Barkovich AJ. Magnetic resonance imaging of the fetal brain and spine: An increasingly important tool in prenatal diagnosis. Part 2. AJNR Am J Neuroradiol. 2006;27(9):1807-1814.
10. Tang PH, Bartha AI, Norton ME, Barkovich AJ, Sherr EH, Glenn OA. Agenesis of the corpus callosum: an MR imaging analysis of associated abnormalities in the fetus. AJNR Am J Neuroradiol. 2009;30(2):257-263. doi: 10.3174/ajnr.A1331.
11. Sileo FG, Di Mascio D, Rizzo G, et al. Role of prenatal magnetic resonance imaging in fetuses with isolated agenesis of corpus callosum in the era of fetal neurosonography: A systematic review and meta-analysis. Acta Obstet Gynecol Scand. 2021;100(1):7-16. doi: 10.1111/aogs.13958.
12. Bedeschi MF, Bonaglia MC, Grasso R, et al. Agenesis of the corpus callosum: clinical and genetic study in 63 young patients. Pediatr Neurol. 2006;34(3):1861-93. doi: 10.1016/j.pediatrneurol.2005.08.008.
13. Zhou C, Li H, Han R, et al. Partial agenesis of the corpus callosum: Prenatal ultrasound characteristics, associations, and outcome. Acta Obstet Gynecol Scand. 2025;104(7):1304-1317. doi: 10.1111/aogs.15121.
14. Mustafa HJ, Barbera JP, Sambatur EV, et al. Diagnostic yield of exome sequencing in prenatal agenesis of corpus callosum: systematic review and meta-analysis. Ultrasound Obstet Gynecol. 2024;63(3):312-320. doi: 10.1002/uog.27440.
15. Glass HC, Shaw GM, Ma C, Sherr EH. Agenesis of the corpus callosum in California 1983-2003: a population-based study. Am J Med Genet A. 2008;146A(19):2495-500. doi: 10.1002/ajmg.a.32418.
16. de Wit MC, Boekhorst F, Mancini GM, et al. Advanced genomic testing may aid in counseling of isolated agenesis of the corpus callosum on prenatal ultrasound. Prenat Diagn. 2017;37(12):1191-1197. doi: 10.1002/pd.5158.
17. Arslanoğlu T, Alpay V, Özyılmaz İ. Prenatal diagnosis and postnatal outcomes of absent pulmonary valve syndrome: A case series with genetic and hemodynamic insights. Eur Res J. 2025;11(6):11081116. doi: 10.18621/eurj.1741967.

The European Research Journal 2026;12(3):297-304 Original Article

DOI: https://doi.org/10.18621/eurj.1829890 *Obstetrics and Gynaecology*

Impact of Ovarian Cyst Size on Pregnancy Course and Neonatal Outcomes: Does Size Really Matter?

Dilara Duygulu Bulan[1], Ruken Dayanan[1], Merve Ayas Ozkan[1] Rıfat Taner Aksoy[2], Ali Turhan Çağlar[1]

[1]Department of Obstetrics and Gynecology, Division of Perinatology, Ankara Etlik City Hospital, Ankara, Türkiye; [2]Department of Obstetrics and Gynecology, Ankara Bilkent City Hospital, Ankara, Türkiye

ABSTRACT

Objectives: This study aimed to evaluate the impact of benign ovarian cyst size on maternal and neonatal outcomes by comparing pregnancies complicated by cysts of different diameters.

Methods: In this retrospective cohort study, 266 pregnant women diagnosed with benign ovarian cysts were categorized into three groups according to cyst diameter: <5 cm, 5–10 cm, and >10 cm. Maternal characteristics, pregnancy complications, delivery outcomes, and neonatal parameters were analyzed. Composite adverse perinatal outcomes (CAPO) included prematurity, low Apgar scores, neonatal intensive care unit (NICU) admission, and neonatal morbidities.

Results: Most small cysts (<5 cm) regressed spontaneously, while larger cysts showed significantly higher rates of complications. Cysts >10 cm were associated with increased risks of preterm delivery (23.1%), composite adverse perinatal outcomes (26.9%), and higher incidences of torsion, rupture, and surgical intervention (P<0.05 for all). No significant differences were found among groups regarding Apgar scores, NICU admission, or neonatal morbidities.

Conclusions: While benign ovarian cysts generally do not impair pregnancy outcomes, cyst size is a key determinant of risk. Cysts >10 cm carry a markedly higher likelihood of maternal complications and adverse perinatal outcomes, supporting closer surveillance and individualized, size-based management strategies.

Keywords: Pregnancy, Ovarian Cyst, Adnexal Mass, Perinatal Outcomes, Cyst Size, Torsion, Preterm Birth

Ovarian cysts during pregnancy are lesions that are detected more frequently with the widespread use of routine obstetric ultrasonography and are mostly identified incidentally [1, 2]. Most of these cysts are benign, may resolve spontaneously during pregnancy, and generally do not cause any significant symptoms. However, in some cases, monitoring or surgical intervention may be necessary due to the size, structure, or risk of complications (such as torsion or rupture) of the cysts [3]. Ultrasonography is the first-line diagnostic method, while advanced imaging techniques may also be used in cases of suspected malignancy or complex cysts [4]. Management of ovarian cysts during pregnancy is individualized based on factors such as cyst size, presence of symptoms, and risk of malignancy; while most cases are monitored conservatively, surgical intervention is generally

Received: November 25, 2025. **Accepted:** January 4, 2026. **Available Online:** January 6, 2026 **Published:** March 1, 2026

How to cite this article: Duygulu Bulan D, Dayanan R, Ayas Ozkan M, Aksoy RT, Çağlar AT. Impact of Ovarian Cyst Size on Pregnancy Course and Neonatal Outcomes: Does Size Really Matter? Eur Res J. 2026;12(3):297-304. doi: 10.18621/eurj.1829890

Corresponding author: Dilara Duygulu Bulan, MD., Phone: +90 312 797 00 00, E-mail: duyguludilara4@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

preferred in the second trimester and in selected patients. Therefore, evaluating current approaches to the diagnosis and management of ovarian cysts detected during pregnancy is of great importance for both maternal and fetal health.

he impact of maternal ovarian cysts on pregnancy and fetal outcomes may vary depending on the cyst's characteristics and management strategy. In general, benign ovarian cysts identified during pregnancy are not expected to adversely influence obstetric or perinatal outcomes, and both maternal and fetal prognoses are usually favorable [5]. Nevertheless, when complications such as torsion, rupture, or rapid enlargement occur, early diagnosis and timely management are essential to prevent serious consequences. Although these complications are uncommon, they can result in adverse outcomes such as preterm delivery, fetal distress, or procedure-related risks associated with surgical intervention.

However, limited data are available regarding the effect of cyst size on maternal and neonatal outcomes in pregnancies complicated by benign ovarian cysts. This study aimed to evaluate the relationship between ovarian cyst size and pregnancy outcomes by comparing maternal and perinatal parameters across three cyst size groups.

METHODS

Study Design

This study was designed as a retrospective cohort analysis to evaluate the impact of ovarian cyst size on maternal and neonatal outcomes in pregnancies complicated by benign ovarian cysts. Data were collected from pregnant women who were evaluated and followed up at the Department of Perinatology between January 2023 and April 2025.

Selection of Case, and Selection of Control

This retrospective study included a total of 266 pregnant women diagnosed with benign ovarian cysts during routine obstetric ultrasonography. All patients were evaluated and followed up at the Department of Perinatology of a tertiary care center between January 2023 and April 2025. During the study period between January 2023 and April 2025, all pregnant women evaluated at our tertiary center for adnexal masses detected during routine obstetric ultrasonography were consecutively screened using hospital electronic medical records and ultrasound archives. Patients were included consecutively if they met the inclusion criteria and had complete antenatal follow-up and delivery records. After applying exclusion criteria, a total of 266 pregnancies with confirmed benign ovarian cysts were included in the final analysis. The number of patients was not predetermined but represents the complete cohort of eligible consecutive cases during the specified study period. Patients were divided into three subgroups according to the maximum diameter of the ovarian cysts measured by ultrasonography: Group 1 (<5 cm, n = 160), Group 2 (5–10 cm, n = 80), and Group 3 (>10 cm, n = 26). Only cases in which the cysts were confirmed to be benign—either through histopathology, spontaneous resolution, or stable benign morphology on serial imaging—were included in the study. Cysts diagnosed as malignant, borderline, or indeterminate, as well as paraovarian or tubal masses, were excluded. Patients with singleton pregnancies who had regular antenatal follow-up and complete delivery records were included. Exclusion criteria were: multiple pregnancies, pre-existing chronic systemic diseases (such as diabetes mellitus or hypertension), fetal structural or chromosomal abnormalities, and pregnancies with incomplete clinical data or follow-up. The diagnosis of ovarian cysts was made by transabdominal or transvaginal ultrasonography during the first or second trimester. Serial ultrasound examinations were performed at routine prenatal visits to monitor cyst characteristics and pregnancy course. Delivery information, including gestational age at birth, birth weight, Apgar scores, and perinatal outcomes, were obtained from hospital records. All patients were followed up and delivered at the same tertiary center, ensuring data consistency and uniformity in management.

Data Collection and Laboratory Procedures

Patient data were retrospectively obtained from hospital electronic medical records and the institutional obstetric database. Demographic, obstetric, and ultrasonographic variables were recorded for all participants. Maternal age, gravidity, parity, height, weight, and pre-pregnancy body mass

index (BMI, kg/m²) were collected. Additional clinical parameters included history of infertility, previous ovarian cysts, conception method (spontaneous or in vitro fertilization), and gestational age at diagnosis of the cyst (weeks). The ultrasonographic characteristics of the ovarian cysts, including maximum diameter (cm) and bilaterality, were noted. Patients were then categorized into three subgroups based on cyst size: <5 cm, 5–10 cm, and >10 cm. Pregnancy-related outcomes were evaluated, including preterm delivery (<37 weeks), premature rupture of membranes (PROM), preterm premature rupture of membranes (PPROM), malpresentation, fetal growth restriction (FGR), fetal distress, cyst rupture, torsion, regression during follow-up, and surgical intervention during pregnancy. Neonatal outcomes included gestational age at delivery (weeks), birth weight (grams), 1-minute and 5-minute Apgar scores, cesarean section rate, and admission to the neonatal intensive care unit (NICU). Adverse perinatal outcomes were defined as the presence of at least one of the following: prematurity (<37 weeks), low Apgar score (<7 at 5 minutes), NICU admission, respiratory distress syndrome (RDS), transient tachypnea of the newborn, neonatal sepsis, hypoglycemia, need for phototherapy, continuous positive airway pressure (CPAP) or mechanical ventilation, or perinatal mortality. A composite adverse perinatal outcome (CAPO) variable was created to encompass these parameters.All data were cross-checked by two independent researchers to ensure accuracy and consistency before statistical analysis.

Statistical Analysis

Statistical analysis was performed utilising IBM SPSS version 22.0 (IBM Corporation, Armonk, NY, USA). The Kolmogorov-Smirnov test was employed to assess adherence to normal distribution. Descriptive statistics for continuous variables were reported as "mean ± standard deviation" for normally distributed data and "median (interquartile range)" for non-normally distributed data. Categorical variables were analysed using the chi-square test or Fisher's exact test. Continuous variables were analysed using the independent sample t-test or the Mann-Whitney U test, contingent upon their normal distribution status.The statistical analysis performed in this study involved the utilization of One-Way ANOVA, followed by post hoc tests to assess intergroup comparisons in cases where significant differences were observed. Pairwise comparisons were performed using Bonferroni-adjusted independent samples t-tests following one-way ANOVA and Bonferroni-corrected Mann–Whitney U tests following Kruskal–Wallis tests, as appropriate. Statistical significance for all tests was defined as a P value less than 0.05.

The study was conducted in accordance with the principles stated in the Declaration of Helsinki and ethical approval was obtained from Ethics Committee (approval number: AESH-BADEK-2025-318). Considering the retrospective nature of the study, Ethics Committee granted an exemption for informed consent.

RESULTS

A total of 266 pregnant women with benign ovarian cysts were included in the analysis. According to cyst diameter, 160 (60.2%) patients were classified into the <5 cm group, 80 (30.1%) into the 5–10 cm group, and 26 (9.8%) into the >10 cm group. The mean maternal age, height, weight, body mass index (BMI), parity, and rates of nulliparity and in vitro fertilization did not differ significantly among the three groups. The mean gestational age at diagnosis was significantly earlier in women with smaller cysts (12.5±3.4 weeks) compared with those with moderate and large sized cysts (14.2±3.8 weeks and 16.1±4.2 weeks, respectively; P<0.001). No significant differences were observed in bilaterality or history of previous ovarian cysts across groups (Table 1).

The benign nature of the ovarian cysts was confirmed through multiple diagnostic pathways. Histopathological confirmation following surgery accounted for 12.0% of cases, whereas radiologic or clinical resolution was observed during pregnancy in 34.2% and within 12 months postpartum in 24.1% of patients. In addition, 29.7% of cysts maintained a stable benign morphology on serial ultrasound examinations, confirming the non-malignant character of all included cases (Table 2). Analysis of pregnancy outcomes revealed a progressive increase in adverse maternal events with increasing cyst size. The rates of preterm delivery rose from 8.8% in the <5 cm group to 12.5% in the 5–10 cm group and 23.1% in the >10

TABLE 1. Maternal Demographic and Ultrasonographic Characteristics According to Ovarian Cyst Size Groups

	<5 cm (60.2%, n=160)	5–10 cm (30.1%, n=80)	>10 cm (9.8%, n=26)	P-value
Maternal age (years)	29.4±5.2	30.1±5.4	30.5±5.6	0.312
Height (cm)	163.2±6.1	163.0±6.0	162.8±6.2	0.874
Weight (kg)	66.5±11.2	67.8±11.5	68.4±12.0	0.546
BMI (kg/m$^{2)}$	25.1±4.1	25.5±4.3	25.8±4.5	0.623
Gravida (n)	2 (1)	2 (1)	2 (1)	0.901
Parity (n)	1 (0)	1 (1)	1 (0)	0.842
Nulliparous	61 (38.1%)	29 (36.2%)	9 (34.6%)	0.918
In vitro fertilization	10 (6.2%)	6 (7.5%)	3 (11.5%)	0.511
History of previous ovarian cyst	19 (11.9%)	12 (15.0%)	5 (19.2%)	0.412
Gestational age at diagnosis (weeks)	12.5±3.4	14.2±3.8	16.1±4.2	**<0.001**[a,b,c]
Bilaterality	10 (6.2%)	6 (7.5%)	3 (11.5%)	0.588

Data are expressed as mean±standars deviation or median (interquartile range) or n (%) where appropriate. BMI, body mass index. [a]The difference between Group 1 and Group 2 is significant, [b]The difference between Group 2 and Group 3 is significant, [c]The difference between Group 1 and Group 3 is significant.

Statistically significant P-value is shown in bold.

cm group (P=0.047). Cyst regression during follow-up was significantly more frequent in smaller cysts (71.9%) compared with moderate (45.0%) and large cysts (15.4%; P<0.001). Conversely, complications such as rupture, torsion, and surgical intervention were significantly more common in larger cysts. The rates of PPROM, PROM, malpresentation, fetal growth restriction, and fetal distress did not differ significantly between groups (Table 3).

TABLE 2. Methods of Benign Confirmation of Ovarian Cysts

	n (%)
Histopathology (post-surgery)	32 (12.0%)
Radiologic/clinical resolution during pregnancy	91 (34.2%)
Radiologic/clinical resolution postpartum (≤ 12 months)	64 (24.1%)
Stable benign morphology on serial scans	79 (29.7%)
Total	266 (100%)

Mean gestational age at delivery and birth weight tended to decrease with increasing cyst size, though not reaching statistical significance (P=0.061 and P=0.082, respectively). The incidence of prematurity (<37 weeks) increased significantly with cyst size (8.8%, 12.5%, 23.1%; P=0.047). Similarly, the rate of composite adverse perinatal outcomes was higher in the >10 cm group (26.9%) compared with the smaller cyst groups (P=0.042). Cesarean section rates, Apgar scores, and rates of NICU admission, transient tachypnea, neonatal sepsis, respiratory distress syndrome, and hypoglycemia did not differ significantly among the groups (Table 4).

DISCUSSION

In the present study, we investigated the impact of benign ovarian cyst size on pregnancy outcomes and found that cyst size is a significant determinant of risk. Overall, our results indicate that most ovarian cysts detected during pregnancy are clinically benign and do not adversely affect the course of pregnancy. However, as cyst size increases, there is a clear rise in

TABLE 3. Pregnancy Course and Maternal Outcomes According to Ovarian Cyst Size Groups

	<5 cm (60.2%, n=160)	5–10 cm (30.1%, n=80)	>10 cm (9.8%, n=26)	P-value
Preterm delivery	14 (8.8%)	10 (12.5%)	6 (23.1%)	**0.047[c]**
PPROM	5 (3.1%)	3 (3.8%)	2 (7.7%)	0.462
PROM	16 (10.0%)	10 (12.5%)	4 (15.4%)	0.542
Malpresentation	6 (3.8%)	5 (6.2%)	3 (11.5%)	0.214
FGR	11 (6.9%)	7 (8.8%)	3 (11.5%)	0.614
Fetal distress	10 (6.2%)	6 (7.5%)	3 (11.5%)	0.542
Regression during follow-up	115 (71.9%)	36 (45.0%)	4 (15.4%)	**<0.001[a,b,c]**
Rupture	2 (1.2%)	2 (2.5%)	2 (7.7%)	**0.041[c]**
Torsion	3 (1.9%)	4 (5.0%)	3 (11.5%)	**0.018[b,c]**
Requirement for surgery	8 (5.0%)	12 (15.0%)	12 (46.2%)	**<0.001[b,c]**

Data are expressed as n (%). BMI, body mass index. FGR, Fetal Growth Restriction; PPROM, Preterm Premature Rupture of Membranes; PROM, Premature Rupture of Membranes.

[a]The difference between Group 1 and Group 2 is significant, [b]The difference between Group 2 and Group 3 is significant, [c]The difference between Group 1 and Group 3 is significant.

Statistically significant P-values are shown in bold.

TABLE 4. Birth Characteristics and Neonatal Outcomes According to Ovarian Cyst Size Groups

	<5 cm (60.2%, n=160)	5–10 cm (30.1%, n=80)	>10 cm (9.8%, n=26)	P-value
Gestational age at delivery (week)	38.3±1.6	38.0±1.8	37.2±2.1	0.061
Prematurity (<37 weeks)	14 (8.8%)	10 (12.5%)	6 (23.1%)	**0.047[c]**
Birth weight (gram)	3170±520	3120±540	2950±610	0.082
Cesarean section	51 (31.9)	29 (36.2%)	12 (46.2%)	0.231
Apgar score at 1st minute	8.3±0.7	8.2±0.8	7.9±0.9	0.115
Apgar score at 5th minute	9.6±0.5	9.5±0.6	9.3±0.7	0.173
CAPO	16 (10.0%)	11 (13.8%)	7 (26.9%)	**0.042[c]**
NICU admission	13 (8.1%)	8 (10.0%)	5 (19.2%)	0.091
Transient tachypnea of the newborn	6 (3.8%)	4 (5.0%)	2 (7.7%)	0.619
Neonatal sepsis	2 (1.2%)	2 (2.5%)	1 (3.8%)	0.551
Respiratory distress syndrome	2 (1.2%)	2 (2.5%)	1 (3.8%)	0.551
Continuous positive airway pressure	5 (3.1%)	3 (3.8%)	2 (7.7%)	0.512
Mechanical ventilation	2 (1.2%)	1 (1.2%)	1 (3.8%)	0.571
Phototherapy for neonates	16 (10.0%)	10 (12.5%)	5 (19.2%)	0.283
Neonatal hypoglycemia	3 (1.9%)	2 (2.5%)	2 (7.7%)	0.199

Data are expressed as mean±standars deviation or n (%) where appropriate. CAPO, composite adverse perinatal outcome; NICU, neonatal intensive care unit.

[a]The difference between Group 1 and Group 2 is significant, [b]The difference between Group 2 and Group 3 is significant, [c]The difference between Group 1 and Group 3 is significant.

Statistically significant P-values are shown in bold.

the likelihood of complications and interventions. In particular, pregnancies involving large ovarian cysts (>10 cm) showed higher rates of preterm delivery and a greater incidence of CAPO compared to those with smaller cysts. These large cysts were also far less likely to regress spontaneously and more often led to acute events such as ovarian torsion or cyst rupture, necessitating surgical management. This size-related risk gradient underscores the importance of stratifying antenatal management based on cyst dimensions.

Our findings align with the broader literature, noting that ovarian cysts in pregnancy are relatively uncommon and usually benign. The incidence of adnexal masses during pregnancy is estimated to range from about 0.2% to 2% of pregnancies, with the vast majority being benign, asymptomatic lesions [1, 6]. In a large prospective study by Gaughran *et al*. [7] İN 2024, adnexal masses were detected in 1% of more than 28,000 pregnancies, the majority of which were considered simple/functional cysts, and 74% of which disappeared spontaneously after birth. Similarly, in our series, a significant portion of the cysts regressed during the follow-up period. The literature reports that 85-96% of corpus luteum and follicle cysts, especially those detected in the first trimester, undergo spontaneous resolution by 16-20 weeks of gestation [8]. Parallel to this phenomenon, in our study, most of the small and simple cysts disappeared during the second trimester. Numerous studies demonstrate the safety of a conservative approach in the management of benign masses. In a retrospective study by Barcroft *et al*. [9], 92% of 267 adnexal masses detected during pregnancy were managed without the need for surgery; only 1.5% required emergency surgery during pregnancy. In the same study, it was stated that one-quarter of the cysts disappeared spontaneously during pregnancy, and most of the remaining were removed electively in the postpartum period. Similarly, in our series, the rate of emergency intervention was found to be low, and the majority of cases reached term pregnancy without any problems. While there are concerns about the possibility of poor obstetric outcomes in the presence of an adnexal mass, recent data suggest that these concerns may be unfounded. A 2021 review reported that planned elective surgery can be performed safely during pregnancy and does not increase the rate of miscarriage, premature rupture of membranes, or preterm birth [10]. In our study, no significant difference was found in the rates of preterm birth or composite adverse perinatal outcome due to the presence of cysts. This finding suggests that the presence of benign adnexal masses does not have a significant negative impact on fetal growth and timing of birth and is consistent with similar studies in the literature [11, 12].

There are also studies in the literature examining the relationship between cyst size and the risk of complications. It has been reported that the risk of torsion is significantly higher in medium-sized cysts (approximately 5-10 cm). In a classic study by Koo *et al*. [13], the risk of torsion was approximately 2.7 times higher in masses measuring 6-10 cm compared to cysts measuring <6 cm. However, it has been stated that the risk of torsion does not increase again in very large cysts (>10-15 cm), perhaps due to the restriction of mobility of the masses. In our study, complication rates - including torsion, rupture, and the need for surgical intervention - were highest in the >10 cm group, rather than in the medium-sized group, suggesting that cysts exceeding a critical size threshold may behave differently in pregnancy compared with the patterns reported in smaller cohorts. In light of these findings, cyst size emerges as a clinically meaningful determinant of risk. Larger cysts are more prone to complications due to increased mass effect, stretching of the utero-ovarian ligament, impaired venous and lymphatic drainage, and a higher likelihood of rupture or hemorrhage as intracystic pressure increases [14, 15]. Several contemporary studies have similarly emphasized that cysts exceeding 8–10 cm are associated with a significantly higher need for intervention, greater rates of acute abdominal pain, and a higher probability of requiring either emergency or elective surgery during pregnancy [16, 17]. Moreover, mass effect from large adnexal lesions may alter uterine irritability or exert mechanical pressure on surrounding structures, potentially contributing to preterm uterine activity.

Taken together, these observations highlight that cyst size remains one of the most clinically relevant parameters in predicting pregnancy-related risks in women with benign adnexal masses. An important strength of our study is the inclusion of a well-defined cohort in which all cysts were confirmed as benign through radiological, clinical, or histopathological criteria, allowing a more precise evaluation of size-

dependent behavior without the confounding effects of malignant or borderline lesions. Moreover, all patients were followed and delivered at the same tertiary referral center, ensuring uniform imaging protocols, consistent obstetric management, and reliable outcome documentation. The stratification of patients into three clearly defined size groups further enabled a more nuanced interpretation of how cyst diameter influences both maternal complications and perinatal outcomes.

Strengths and Limitations

Nevertheless, certain limitations should be acknowledged. First, the retrospective design inherently introduces the possibility of selection bias and limits the ability to establish causal relationships. Although the study population is relatively large, the number of women with cysts >10 cm remained small, reflecting the rarity of giant benign adnexal masses in pregnancy; therefore, risk estimates for this subgroup should be interpreted cautiously. Finally, the single-center design may restrict external generalizability, as surgical thresholds, ultrasound expertise, and follow-up practices can vary across institutions. Despite these limitations, our findings contribute meaningful evidence to the ongoing discussion regarding the management of benign adnexal masses in pregnancy. The results suggest that while most cysts identified during gestation follow a benign course and frequently regress, cysts exceeding 10 cm warrant closer surveillance due to their higher likelihood of torsion, rupture, surgical intervention, and preterm birth. In contrast, smaller cysts rarely require intervention and are associated with favorable maternal and neonatal outcomes. These observations support an individualized, size-based approach in which conservative management is appropriate for the majority of patients, while large or symptomatic cysts may benefit from proactive monitoring and timely intervention. Future prospective multicenter studies with larger samples of giant cysts are needed to refine risk stratification and further inform management algorithms for this unique patient population.

CONCLUSION

Benign ovarian cysts detected during pregnancy are generally considered to have a favorable clinical course and are commonly managed conservatively, with most lesions remaining asymptomatic or regressing spontaneously. However, evidence regarding the influence of cyst size on maternal and perinatal outcomes has been limited and inconsistent. The present study adds to the existing literature by demonstrating a clear size-dependent risk pattern, showing that cysts larger than 10 cm are associated with significantly higher rates of maternal complications, including torsion, rupture, and the need for surgical intervention, as well as increased risk of preterm birth and composite adverse perinatal outcomes. These findings highlight cyst size as a key determinant in risk stratification and support an individualized, size-based management approach to optimize both maternal and neonatal outcomes in pregnancies complicated by benign ovarian cysts.

Ethics Approval and Consent to Participate

This study was approved by the Ankara Etlik City Hospital Scientific Research Evaluation Ethics Committee No. 2 (Decision No: AEŞH-BADEK2-2025-318; date: 05.08.2025). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Informed consent was waived because of the retrospective nature of the study and the analysis used anonymous clinical data.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: DDB; Study Design: DDB, ATÇ; Supervision: DDB; Funding: N/A; Materials: N/A; Data Collection and/or Processing: MAO, RD, RTA; Statistical Analysis and/or Data Interpretation: RTA, DDB; Literature Review: DDB; Manuscript Preparation: DDB, MAO, RD, ATÇ; and Critical Review: DDB, ATÇ.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Shakya S, Shrestha A , Lama S. Pregnancy Complicated by Ovarian Cysts. Kathmandu Univ Med J (KUMJ). 2024;22(88):499-502.
2. Bhagat N, Gajjar K. Management of ovarian cysts during pregnancy. Obstet Gynaecol Reprod Med. 2022;32(9):205-210. doi: 10.1016/j.ogrm.2022.06.002.
3. El-Shawarby SA, Henderson AF, Mossa MA. Ovarian cysts during pregnancy: dilemmas in diagnosis and management. J Obstet Gynaecol. 2005;25(7):669-75. doi: 10.1080/01443610500291686.
4. Testa AC, Mascilini F, Quagliozzi L, et al. Management of ovarian masses in pregnancy: patient selection for interventional treatment. Int J Gynecol Cancer. 2021;31(6):899-906. doi: 10.1136/ijgc-2020-001996.
5. Katz L, Levy A, Wiznitzer A, Sheiner E. Pregnancy outcome of patients with dermoid and other benign ovarian cysts. Arch Gynecol Obstet. 2010;281(5):811-815. doi: 10.1007/s00404-009-1158-1.
6. Cathcart AM, Nezhat FR, Emerson J, Pejovic T, Nezhat CH, Nezhat CR. Adnexal masses during pregnancy: diagnosis, treatment, and prognosis. Am J Obstet Gynecol. 2023;228(6):601-612. doi: 10.1016/j.ajog.2022.11.1291.
7. Gaughran J, Magee C, Mitchell S, Knight CL, Sayasneh A. Adnexal Masses in Pregnancy: A Single-Centre Prospective Observational Cohort Study. Diagnostics (Basel). 2024;14(19):2182. doi: 10.3390/diagnostics14192182.
8. Kobenge FM, Elong F-A, Mpono EP, Egbe TO. Management of a Huge Ovarian Cyst in Pregnancy at the Douala General Hospital, Cameroon: A Case Report and Review of the Literature. Adv Reprod Sci. 2024;12(3):165–78. doi: 10.4236/arsci.2024.123014.
9. Barcroft J, Pandrich M, Del Forno S, et al. Evaluating use of two-step International Ovarian Tumor Analysis strategy to classify adnexal masses identified in pregnancy: pilot study. Ultrasound Obstet Gynecol. 2024;64(6):808-817. doi: 10.1002/uog.27707.
10. Martone S, Troìa L, Luisi S. Adnexal masses during pregnancy: management for a better approach. Gynecol Surg 2021;18:3. doi: 10.1186/s10397-021-01084-9.
11. Balinskaite V, Bottle A, Sodhi V, et al. The Risk of Adverse Pregnancy Outcomes Following Nonobstetric Surgery During Pregnancy: Estimates From a Retrospective Cohort Study of 6.5 Million Pregnancies. Ann Surg. 2017;266(2):260-266. doi: 10.1097/SLA.0000000000001976.
12. Djokovic D, Pinto P, Reis I. Maternal adnexal masses in pregnancy. Best Pract Res Clin Obstet Gynaecol. 2025;102:102643. doi: 10.1016/j.bpobgyn.2025.102643.
13. Koo YJ, Kim TJ, Lee JE, et al. Risk of torsion and malignancy by adnexal mass size in pregnant women. Acta Obstet Gynecol Scand. 2011;90(4):358-361. doi: 10.1111/j.1600-0412.2011.01070.x.
14. Sen S, Maurya R, Sen G, Srivastava H. Atypical Presentation of a Giant Hemorrhagic Ovarian Cyst. J Obstet Gynaecol India. 2022;72(Suppl 2):435-438. doi: 10.1007/s13224-022-01625-y.
15. Lim WH, Woods N, Lamaro VP. Trends and outcomes of ruptured ovarian cysts. Postgrad Med J. 2022;98(1161):e9. doi: 10.1136/postgradmedj-2020-138833.
16. Leiserowitz GS. Managing ovarian masses during pregnancy. Obstet Gynecol Surv. 2006;61(7):463-470. doi: 10.1097/01.ogx.0000224614.51356.b7.
17. Testa AC, Mascilini F, Quagliozzi L, et al. Management of ovarian masses in pregnancy: patient selection for interventional treatment. Int J Gynecol Cancer. 2021;31(6):899-906. doi: 10.1136/ijgc-2020-001996.

The European Research Journal 2026;12(3):305-314 Original Article

DOI: https://doi.org/10.18621/eurj.1844291 *Cardiology*

Association Between Kidney Function and Morphology-Voltage-P Wave Duration (MVP) Score: A Cross-Sectional Electrocardiographic Study

Ömer Doğan[1], Abdullah Ömer Ebeoğlu[2], Şevval İlke Ebeoğlu[1], Ümit Acar[1], Hasan Ali Barman[1]

[1]Department of Cardiology, İstanbul University-Cerrahpaşa, Institute of Cardiology, İstanbul, Türkiye; [2]Department of Cardiology, Bağcılar Training and Research Hospital, İstanbul, Türkiye

ABSTRACT

Objectives: Chronic kidney disease (CKD) is associated with a high burden of atrial arrhythmias, primarily atrial fibrillation (AF), due to structural and electrical remodelling of the atria. The Morphology–Voltage–P wave duration (MVP) score, a composite electrocardiographic parameter derived from surface electrocardiography (ECG), has been proposed as a simple, non-invasive marker of atrial electrical dysfunction. However, its relationship with renal function has not been previously explored.

Methods: This single-centre, retrospective, cross-sectional study included 90 adults with varying levels of estimated glomerular filtration rate (eGFR). MVP scores were calculated from standard 12-lead ECGs, and echocardiographic, demographic, and laboratory variables were recorded. Patients were grouped by eGFR (>60, 30–59, and <30 mL/min/1.73 m²). Correlation and multiple linear regression analyses were performed to identify independent predictors of MVP score.

Results: MVP score increased across worsening eGFR categories (1.33±0.61, 2.63±1.03, and 4.64±1.26, respectively; P<0.001). A strong inverse correlation was observed between eGFR and MVP score (r = –0.774, P<0.001). In multivariable analysis, eGFR (β = –0.519, P<0.001), left atrial diameter (β=0.396, P<0.001), and male sex (β=0.133, P=0.029) were independent determinants of higher MVP values, explaining 63% of MVP variance.

Conclusions: Reduced renal function is independently associated with higher MVP scores, reflecting atrial electrical abnormalities in patients with CKD. The MVP score may serve as a low-cost, easily applicable ECG marker for early detection of subclinical atrial dysfunction and arrhythmia risk in patients with renal impairment. Prospective, multicentre studies integrating MVP with advanced imaging and artificial intelligence–based ECG analysis are warranted to validate its clinical utility.

Keywords: Morphology-Voltage-P Wave Duration (MVP) Score, Chronic Kidney Disease, Atrial Electrical Remodelling

Received: December 18, 2025 **Accepted:** January 18, 2026 **Available Online:** January 26, 2026 **Published:** March 1, 2026

How to cite this article: Doğan Ö, Ebeoğlu AÖ, Ebeoğlu Şİ, Acar Ü, Barman HA. Association Between Kidney Function and Morphology–Voltage–P Wave Duration (MVP) Score: A Cross-Sectional Electrocardiographic Study. Eur Res J. 2026;12(3):305-314. doi: 10.18621/eurj.1844291

Corresponding author: Abdullah Ömer Ebeoğlu, MD., Phone: +90 212 440 40 00, E-mail: ebeoglu995@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

Chronic kidney disease (CKD) is one of the leading contributors to cardiovascular morbidity and mortality, and atrial arrhythmias, particularly atrial fibrillation (AF), are frequently observed in this population [1]. In CKD, volume overload, hypertension, left ventricular hypertrophy, and electrolyte disturbances accelerate atrial remodelling and fibrosis, thereby predisposing patients to the development of AF [2]. Accordingly, early evaluation of atrial electrical activity in patients with CKD is of substantial clinical relevance.

On surface electrocardiography (ECG), P-wave parameters - particularly morphology, duration, and voltage - serve as inexpensive and non-invasive markers that provide insight into atrial conduction abnormalities and interatrial block, and may predict the risk of AF [3]. The Morphology–Voltage–P wave duration (MVP) score, which integrates these parameters, was first described by Alexander *et al*. [4] and was shown to be a significant predictor of new-onset AF in patients with coronary artery disease. Subsequent studies have demonstrated that the MVP score can predict in-hospital and long-term AF in patients with acute ischaemic stroke, AF recurrence following pulmonary vein isolation, early left atrial (LA) dysfunction in hypertensive individuals, and long-term AF development in patients with systolic heart failure [5-8]. Moreover, recent population-based data suggest that the MVP score may also aid in stroke risk stratification [9].

The MVP score is a low-cost parameter that can be easily integrated into routine clinical practice, as it is derived solely from surface ECG, can be measured objectively, and does not require any additional equipment [4]. However, no study in the current literature has evaluated the relationship between renal function and the MVP score.

This study aims to investigate the association between renal function and the MVP P-wave score in individuals with different estimated glomerular filtration rate (eGFR) levels, thereby assessing the clinical potential of the MVP score as an easily obtainable electrocardiographic marker for the early detection of atrial electrical dysfunction.

METHODS

Study Population and Data Collection

This single-centre, retrospective, cross-sectional observational study was conducted to evaluate the relationship between the MVP score and renal function parameters in individuals with varying degrees of kidney function. Data were obtained from electronic medical records and archived digital ECG recordings. The study included a total of 90 adult individuals who presented to our outpatient clinic between January 2025 and July 2025. Participants were selected to represent a broad spectrum of renal function. Accordingly, patients were enrolled and stratified by renal function into three eGFR categories: >60 mL/min/1.73 m², 30–59 mL/min/1.73 m² and <30 mL/min/1.73 m². This approach allowed the assessment of the association between renal function and electrocardiographic parameters not only in advanced kidney disease, but across the full physiological range.

All participants were in sinus rhythm at the time of evaluation, and standard 12-lead ECG and transthoracic echocardiography (TTE) were performed for each individual. Conditions that could affect P-wave morphology or atrial conduction were excluded. Exclusion criteria comprised a history of atrial fibrillation or atrial flutter, the presence of a permanent pacemaker or implantable cardioverter-defibrillator (ICD), Wolff–Parkinson–White (WPW) syndrome, complete left or right bundle branch block, acute myocardial ischaemia, moderate-to-severe valvular heart disease, and systolic heart failure (left ventricular ejection fraction [LVEF] < 50%). In addition, patients with active infection or sepsis, known thyroid dysfunction, malignancy, or significant electrolyte imbalance (serum potassium > 5.5 or < 3.0 mmol/L) were excluded. ECG and echocardiographic recordings with inadequate image or signal quality were also excluded from the final analysis.

Clinical, demographic, laboratory, and echocardiographic data were obtained from the hospital's electronic medical record system. Demographic variables included age, sex, body mass index (BMI), and comorbidities such as hypertension, diabetes mellitus, and coronary artery disease. Laboratory parameters included serum creatinine, hemoglobin, electrolytes, and C-reactive protein (CRP) levels.

Electrocardiographic Assessment and Calculation of MVP Score

A standard 12-lead ECG was recorded in all participants at a paper speed of 50 mm/s and a voltage calibration of 10 mm/mV using a Philips PageWriter TC70 device (Philips Healthcare, Andover, MA, USA). To minimise autonomic influences, all recordings were obtained in the supine position after at least 10 minutes of rest. Digital ECG recordings were subsequently reviewed and analysed in electronic format.

P-wave analyses were performed offline on digital ECG recordings at ×200 magnification by two experienced cardiologists who were blinded to all clinical, laboratory, and echocardiographic data. For each lead, measurements were averaged over three consecutive sinus beats. The MVP score was calculated as previously described:

•P-wave duration: <120 ms = 0 points; 120–140 ms = 1 point; >140 ms = 2 points

•P-wave voltage in lead I: >200 µV = 0 points; 100–200 µV = 1 point; <100 µV = 2 points

•P-wave morphology in inferior leads: normal = 0 points; partial interatrial block = 1 point; advanced interatrial block = 2 points

The total MVP score ranged from 0 to 6, with higher values indicating more pronounced atrial conduction abnormalities and atrial electrical remodelling [4]. All ECG measurements were performed independently by two blinded cardiologists using a standardized protocol, and discrepancies were resolved by consensus.

Echocardiographic and Laboratory Assessment

TTE was performed in all participants using a Philips EPIQ 7C ultrasound system (Philips Healthcare, Andover, MA, USA) in accordance with the current recommendations of the American Society of Echocardiography (ASE) [10]. Standard echocardiographic measurements were obtained in all patients. All measurements were performed independently by two experienced cardiologists who were blinded to the MVP score results, and in cases of discrepancy, the final value was determined by consensus.

Venous blood samples were collected in the morning after an overnight fast. Serum creatinine, electrolytes (Na^+, K^+, Ca^{2+}, Mg^{2+}), hemoglobin, and CRP levels were measured using an Abbott Architect c8000 autoanalyser (Abbott Diagnostics, Illinois, USA). The eGFR was calculated using the CKD-EPI 2021 equation [11].

Statistical Analysis

All statistical tests were conducted using the Statistical Package for the Social Sciences 25.0 for Windows (SPSS Inc., Chicago, IL, USA). The Kolmogorov-Smirnov test was used to analyze the normality of the data. Continuous data are expressed as mean±SD, and categorical data are expressed as percentages. Chi-square test was used to assess differences in categorical variables between groups. Comparisons of independent numerical variables were performed using repeated-measures ANOVA and/or the Kruskal–Wallis test, depending on the distribution of the data. The relationships among parameters were assessed using Pearson's or Spearman's correlation analysis according to the normality of the data. Correlations between variables were evaluated by Pearson's rank correlation test. Multiple linear regression analyses using the stepwise method were performed to assess the independent variables affecting the dependent variable MVP score. All independent variables in the multiple linear regression were tested for multicollinearity. If the variance inflation factor exceeded 3.0, the variable was considered to be collinear. All reported confidence interval (CI) values are calculated at the 95% level. Significance was assumed at a 2-sided $P<0.05$.

RESULTS

The demographic, clinical, electrocardiographic, and echocardiographic characteristics of the patients included in the study are summarized in Table 1. A total of 90 participants were included in the study. The mean age of the population was comparable across the three eGFR categories ($P=0.182$). The distribution of sex, BMI, hypertension, diabetes mellitus, and coronary artery disease did not differ significantly between the groups (all $P>0.05$).

Electrocardiographic and echocardiographic findings demonstrated progressive alterations in atrial conduction and structure with decreasing renal function. P-wave duration increased significantly from

TABLE 1. Demographic and Clinical Characteristics of the Study Population.

Parameters	eGFR (>60) (n=30)	eGFR (30-59) (n=38)	eGFR (<30) (n=22)	P-value
Age (years)	58.3±9.9	56.4±8.6	56.4±11.2	0.677
BMI (kg/m²)	28.6±2.8	28.4±2.9	29.7±2.4	0.176
Male	20 (66.6%)	17 (44.7%)	12 (54.5%)	0.443
Smoking	16 (53.3%)	17 (44.7%)	14 (63.6%)	0.365
Hypertension	10 (33.3%)	17 (44.7%)	12 (54.5%)	0.304
Diabetes mellitus	11 (36.7%)	22 (57.9%)	10 (45.5%)	0.213
Coronary artery disease	9 (30.0%)	6 (15.8%)	4 (18.2%)	0.336
ECG findings				
P-wave duration (ms)	100.3±12.7	116.7±12.8	129.1±19.7	**<0.001**
P-wave voltage (μV)	116.3±23.6	116.2±47.2	87.3±10.5	**0.004**
P-wave morphology				**<0.001**
Normal	28 (93.3%)	15 (39.5%)	2 (9.1%)	
Partial block	2 (6.7%)	20 (52.6%)	6 (27.3%)	
Complete block	0 (0.0%)	3 (7.9%)	14 (63.6%)	
MVP score	1.33±0.61	2.63±1.03	4.64±1.26	**<0.001**
QRS (ms)	99.3±12.3	99.5±11.6	108.6±17.3	**0.023**
QTc (ms)	392.4±16.0	391.6±14.5	405.9±25.2	**0.009**
Echocardiography findings				
IVS (mm)	10.0±1.1	10.9±1.2	10.6±0.9	**0.004**
PW (mm)	9.6±0.7	9.7±0.9	9.3±1.0	0.219
LVd (mm)	48.2±2.4	50.1±3.1	50.6±3.2	**0.005**
LA (mm)	36.2±3.9	39.6±4.2	43.0±4.0	**<0.001**
PASP (mmHg)	26.6±4.9	31.7±7.6	28.1±6.8	**0.007**
Laboratory findings				
Creatinine (mg/dL)	1.00±0.18	1.65±0.38	4.05±1.56	**<0.001**
eGFR (mL/min/1.73 m²)	78.2±12.6	41.4±7.7	16.8±7.3	**<0.001**
Na (mmol/L)	139.3±2.7	137.1±3.3	140.5±2.6	**<0.001**
K (mmol/L)	4.06±0.44	4.07±0.34	4.85±0.45	**<0.001**
Hemoglobin (g/dL)	14.9±1.4	14.6±1.3	14.6±1.2	0.597
Total cholesterol (mg/dL)	182.0±36.8	181.6±29.9	185.5±24.3	0.891
HDL (mg/dL)	37.4±7.3	37.3±10.9	36.7±8.2	0.965
LDL (mg/dL)	131.0±33.6	126.0±31.5	138.5±24.1	0.323
Triglyceride (mg/dL)	185.4±67.1	170.7±88.6	147.5±59.9	0.221
Medical history				
Beta-blocker use	17 (56.7%)	21 (55.2%)	14 (63.6%)	0.341
Calcium channel blocker use	6 (20.0%)	8 (21.0%)	5 (22.7%)	0.643
Other antiarrhythmics	7 (23.3%)	7 (18.4%)	0 (0.0%)	0.059

Data are shown as mean±standard deviation or n (%) where appropriate. BMI, body mass index; eGFR, estimated glomerular filtration rate; HDL, high-density lipoprotein; IVS, interventricular septum diameter; K, potassium; LA, left atrium; LDL, low-density lipoprotein; LVd, left ventricular diameter; MVP, morphology-voltage-P wave duration score; Na, sodium; PASP, pulmonary artery systolic pressure; PW, posterior wall diameter; QTc, corrected QT interval.

Statistically significant P-values are shown in bold.

TABLE 2. Spearman Correlation and Multivariable Linear Regression Analyses of Determinants of the MVP Score.

Coefficients[a]				
Model	**Unstandardized coefficients**		**Standardized coefficients**	**P-value**
	B	**SE**	**Beta**	
(Constant)	-1.541	1.109		0.168
eGFR (mL/min/1.73m²)	-0.032	0.004	-0.519	**<0.001**
LA (mm)	0.131	0.024	0.396	**<0.001**
Gender (male)	0.417	0.188	0.133	0.029
Excluded Variables[a,b]				
Model	**B**	**Partial correlation**	**Collinearity statistics**	**P-value**
			Tolerance	
Age	-0.102	-0.166	0.804	0.124
BMI	-0.076	-0.120	0.744	0.269
HT	-0.059	-0.097	0.826	0.369
DM	0.027	0.049	0.971	0.450
IVS	-0.019	-0.031	0.819	0.776

BMI, body mass index; DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; HT, Hhypertension; IVS, interventricular septum diameter; LA, left atrium; MVP score, morphology-voltage-P-wave duration score; SE, standard error.

[a]Dependent variable: MVP score

[b]Correlates in the Model: (Constant), eGFR, LA, Gender

Statistically significant P-values are shown in bold.

100.3±12.7 ms in the eGFR > 60 mL/min/1.73 m² group to 129.1±19.7 ms in patients with eGFR <30 mL/min/1.73 m² (P<0.001). Conversely, P-wave voltage decreased as renal function declined (P=0.004). The prevalence of advanced interatrial block was markedly higher in patients with lower eGFR values (P<0.001).

The MVP score showed a stepwise increase across eGFR categories, with mean values of 1.33±0.61, 2.63±1.03, and 4.64±1.26, respectively (P<0.001). Moreover, QTc interval and QRS duration were both significantly prolonged in patients with eGFR <30 mL/min/1.73 m² (P=0.009 and P=0.023, respectively). Echocardiographic parameters revealed significant structural changes associated with reduced renal function. LA diameter and left ventricular end-diastolic diameter differed significantly across eGFR categories and were highest in the eGFR < 30 mL/min/1.73 m² group (P<0.001 and P=0.005, respectively). Similarly, interventricular septal diameter (IVS) increased with worsening renal function (P=0.004), while pulmonary artery systolic pressure (PASP) reached its highest value in the eGFR 30–59 mL/min/1.73 m² group (P=0.007).

Laboratory findings were consistent with the degree of renal dysfunction. Serum creatinine and potassium levels were significantly higher in the eGFR < 30 mL/min/1.73 m² group compared with the other eGFR categories (both P<0.001), whereas hemoglobin, total cholesterol, LDL, HDL, and triglycerides showed no significant differences among groups (P>0.05). Regarding medical therapy, beta-blocker and calcium channel blocker use did not differ significantly across eGFR groups (P>0.05).

In the Spearman correlation analysis (Table 2), a strong and inverse relationship was identified between the MVP score and eGFR (r = –0.774, P<0.001). In addition, the MVP score was positively correlated with LA diameter (r = 0.574, P<0.001). No significant correlations were observed between MVP score and age, BMI, hypertension, diabetes mellitus, or IVS (all P>0.05).

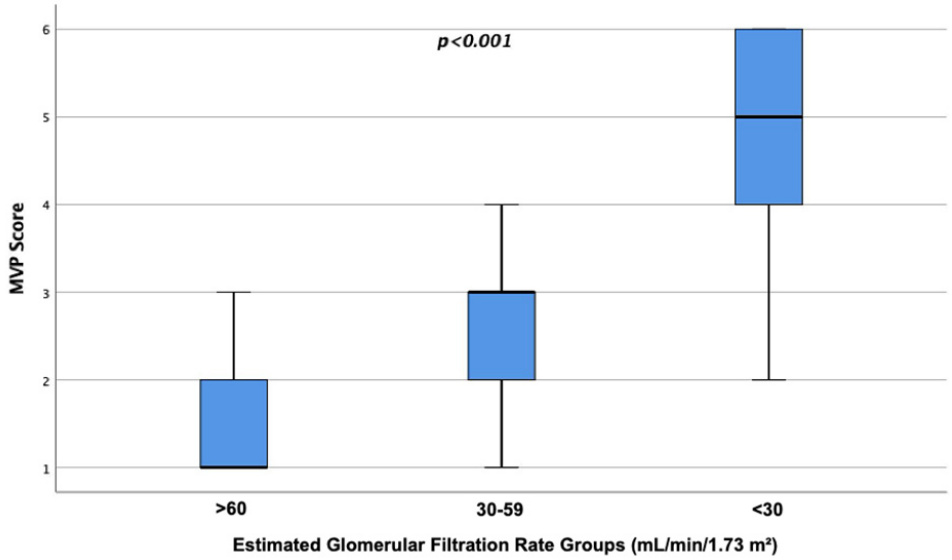

FIGURE 1. Comparison of Morphology-Voltage-P wave duration score (MVP) scores according to estimated glomerular filtration rate categories.

In the multivariable linear regression analysis (Table 2), eGFR (β = –0.519, P<0.001), LA diameter (β = 0.396, P<0.001), and male sex (β = 0.133, P=0.029) were identified as independent predictors of higher MVP scores. eGFR was the most powerful determinant of MVP score, indicating that lower renal function was significantly associated with higher MVP values. LA diameter also showed a positive association with MVP score, whereas male sex was found to be an independent but relatively weaker predictor. Together, these three variables explained approximately 63% of the variance in MVP score (adjusted $R^2 = 0.63$). Age, BMI, hypertension, diabetes mellitus, and IVS did not make a statistically significant contribution when added to the model (all P>0.05).

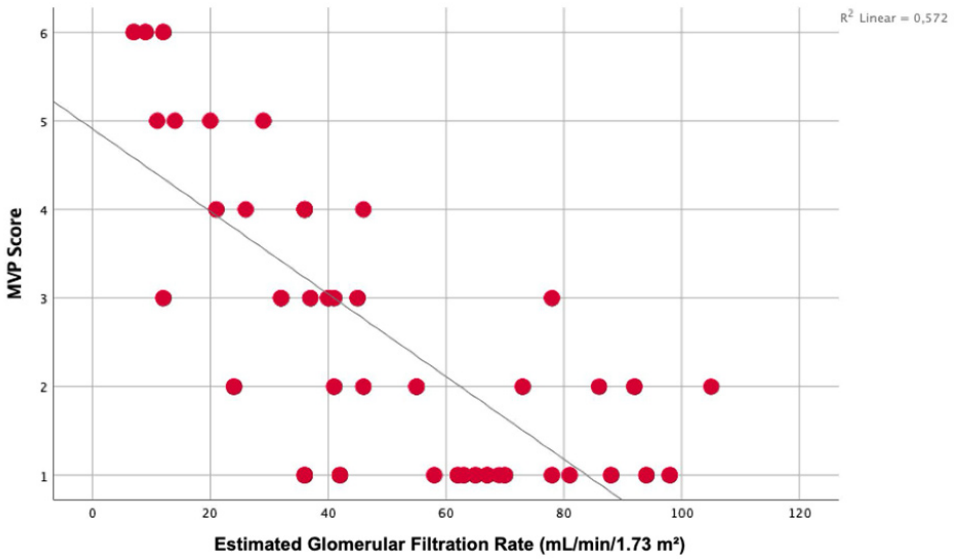

FIGURE 2. Negative correlation between Morphology-Voltage-P wave duration score (MVP) score and estimated glomerular filtration rate ($R^2 = 0.572$, P<0.001)

Figure 1 illustrates the distribution of MVP scores across eGFR categories, demonstrating a progressive and statistically significant increase in MVP score with worsening renal function (P<0.001). Figure 2 presents the scatter plot depicting the correlation between MVP score and eGFR, confirming a strong, inverse, and linear relationship between the two parameters (R^2 = 0.572, P<0.001). These graphical findings are consistent with the regression analyses, visually reinforcing the close association between reduced renal function and elevated MVP score.

DISCUSSION

In this study, we investigated the relationship between the MVP score and renal function parameters in individuals with varying degrees of kidney function. Our findings demonstrated a significant increase in MVP score with decreasing eGFR. In addition, LA diameter and male sex were independently associated with higher MVP scores, providing further evidence of an independent association between impaired renal function and atrial electrical remodelling. These results suggest a meaningful association between declining kidney function and disruption of atrial electrical and structural integrity.

CKD is a systemic condition characterised not only by loss of renal function, but also by chronic inflammation, oxidative stress, excessive activation of the renin–angiotensin–aldosterone system (RAAS), and endothelial dysfunction [12-13]. These mechanisms promote atrial fibrosis, conduction slowing, and increased electrical heterogeneity, thereby facilitating atrial remodelling [14-15]. Progressive deterioration in renal function was associated with electrocardiographic features consistent with atrial conduction abnormalities, as reflected by higher MVP scores [16-17]. The increase in MVP score observed in our cohort may therefore be interpreted as a non-invasive electrocardiographic marker of this underlying pathophysiological process. In patients with chronic kidney disease, even in the absence of overt electrolyte abnormalities, subclinical shifts in electrolyte balance and chronic volume overload are common and may influence atrial conduction properties. Such factors may contribute to prolongation of P-wave duration, attenuation of P-wave voltage, and alterations in P-wave morphology, thereby affecting MVP score components. These mechanisms should be regarded as potential contributors to atrial electrical vulnerability rather than direct causal pathways, and residual confounding related to these subclinical changes cannot be fully excluded [12-17].

The MVP score is an electrocardiographic index developed by Alexander *et al*. that combines P-wave duration, voltage, and morphology into a composite measure quantitatively reflecting atrial conduction abnormalities. Previous studies have reported that the MVP score predicts new-onset AF in patients with coronary artery disease, serves as a marker of LA dysfunction in hypertensive individuals, and is associated with both in-hospital and long-term AF risk in patients with acute ischaemic stroke. In addition, it has been shown to correlate with arrhythmic events in patients with systolic heart failure [4-8]. However, to date, no study has directly evaluated the relationship between renal function and the MVP score. In this context, our study provides the first evidence suggesting that the MVP score may be linked not only to cardiac pathology but also to systemic organ dysfunction, particularly impaired renal function.

Our findings demonstrate that the MVP score increases significantly with declining eGFR (β = –0.519, P<0.001), and this association remains independent in multivariable analysis. As renal function deteriorates, fibrotic changes within atrial conduction pathways and prolongation of depolarisation time are expected, which are reflected by prolonged P-wave duration, reduced P-wave voltage, and consequently higher MVP scores. The independent association of LA diameter with elevated MVP scores further indicates that this process encompasses not only electrical, but also mechanical atrial remodelling. In patients with CKD, volume overload, left ventricular hypertrophy, and diastolic dysfunction increase atrial wall stress and accelerate fibrosis [18]. As a result of this pathological cascade, P-wave duration is prolonged, voltage is attenuated, and the MVP score rises. Therefore, the MVP score may represent one of the earliest and most accessible surface ECG markers of subclinical atrial remodelling associated with renal dysfunction. Although a definitive clinically meaningful cut-off for the MVP score has not been universally established, prior

studies have consistently shown that higher MVP values - particularly scores ≥3–4 - are associated with an increased risk of new-onset or recurrent AF [14-18]. In this context, the progressive increase in MVP score observed across worsening eGFR categories in our study suggests a shift toward a higher atrial arrhythmic risk profile rather than an isolated electrocardiographic change. Therefore, the observed differences in MVP score may be considered clinically relevant in terms of atrial vulnerability.

The positive association between male sex and higher MVP scores is also noteworthy. Sex-related differences in hormonal milieu, atrial wall thickness, and fibrotic response may contribute to this finding. Consistent with previous reports indicating a higher prevalence of interatrial conduction disturbances and advanced interatrial block in men, our results suggest that sex should be taken into account when interpreting MVP scores in clinical practice [19-20].
In CKD, sodium–water retention, activation of the RAAS, and heightened sympathetic tone promote fibrosis, oxidative stress, and electrical delays at both atrial and ventricular levels. This pathological process is reflected on the surface ECG by prolonged P-wave duration, reduced P-wave voltage, and altered P-wave morphology [14-15]. Accordingly, an increased MVP score may be interpreted as a direct surface ECG manifestation of atrial fibrosis and conduction slowing. Our findings provide non-invasive electrophysiological evidence supporting the link between atrial remodelling and arrhythmogenesis in CKD.

From a clinical perspective, the MVP score is an easily obtainable, low-cost, and non-invasive parameter. It may therefore serve as a useful tool for the early detection of subclinical atrial dysfunction, prediction of arrhythmic risk, and individualisation of follow-up strategies in patients with CKD. Particularly in populations at high risk for AF or other atrial arrhythmias, incorporation of the MVP score into routine ECG assessment could facilitate early identification of vulnerable patients and guide the need for further diagnostic evaluation.

Strengths and Limitations

The present study is strengthened by its methodological rigor and clinical relevance. To our knowledge, it provides the first evidence linking renal function to the MVP ECG score in a cohort spanning a broad eGFR range. P-wave indices were derived from standardized ECG recordings (50 mm/s) and quantified via high-resolution offline measurements, with independent blinded assessment by two cardiologists and consensus adjudication to ensure reproducibility. Moreover, robust multivariable modelling was applied to delineate independent determinants of MVP score while controlling for key clinical and echocardiographic confounders.

Despite these strengths, the present study has certain limitations that warrant consideration. First, its cross-sectional design precludes any causal inference. Second, the sample size is relatively limited and prospective follow-up data are lacking. In addition, the single-centre design may limit the generalizability of the findings to broader populations. Although the MVP score was assessed manually, the high interobserver agreement supports the reliability of our measurements. In addition, we did not employ advanced imaging modalities - such as atrial strain analysis or cardiac magnetic resonance - to further characterise atrial fibrosis and electrical remodelling. Accordingly, LA diameter was used as the structural echocardiographic marker, which may have limited the detection of more subtle atrial structural and functional changes. Furthermore, patient recruitment from an outpatient setting may introduce selection bias, potentially under-representing patients with more advanced or unstable clinical conditions. The potential effects of antihypertensive therapies, electrolyte disturbances, and dynamic changes in volume status were also not systematically evaluated. However, these variables showed no meaningful variation or significant association with the MVP score and were therefore excluded from the multivariable model to avoid overfitting. Quantitative interobserver and intraobserver reproducibility metrics, such as intraclass correlation coefficients, were not formally calculated. Given the retrospective and cross-sectional design of the study, causal relationships or temporal directionality cannot be established, and the observed associations should be interpreted as hypothesis-generating. Prospective longitudinal studies are warranted to determine whether deterioration in renal function precedes or accelerates atrial electrical remodelling.

CONCLUSION

This study demonstrates a close association between the MVP score and renal function. The progressive increase in MVP score with declining kidney function supports its role as a non-invasive electrocardiographic marker of atrial electrical remodelling in patients with CKD. Integration of the MVP score into routine ECG assessment may facilitate earlier identification of patients with renal dysfunction who are at increased risk for atrial arrhythmias. Future prospective multicentre studies incorporating MVP scoring with atrial strain measurements, advanced imaging techniques (including MRI-based fibrosis assessment), and artificial intelligence–driven ECG analysis are warranted to validate these findings. Such an approach may enable the incorporation of the MVP score into AI-supported decision models for the early detection of subclinical atrial remodelling and prediction of atrial arrhythmogenesis.

Ethics Approval and Consent to Participate

This study was approved by the Istanbul University-Cerrahpasa Medical Research Ethics Committee (Decision No: E-24687260-604.01-1498608; date: 25.11.2025). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. The requirement for written informed consent was waived due to the retrospective design and the use of anonymized clinical data.

Data Availability

The datasets used and/or analysed during the current study are included in this published article. Additional data supporting the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contribution

Study Conception: HAB; Study Design: ÖD; Supervision: ÖD, HAB; Funding: N/A; Materials: ŞİE, ÜA; Data Collection and/or Processing: ŞİE, AÖE; Statistical Analysis and/or Data Interpretation: ÖD, AÖE; Literature Review: HAB, ÜA; Manuscript Preparation: ÖD, AÖE; and Critical Review: ŞİE, ÜA.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgment

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

During the preparation of this manuscript, the authors used an artificial intelligence–based language model (ChatGPT, OpenAI, San Francisco, CA, USA) solely to improve language and readability. The authors reviewed, edited, and take full responsibility for the content of this manuscript.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Zhang D, Feng Y, Leung FC, Wang L, Zhang Z. Does Chronic Kidney Disease Result in High Risk of Atrial Fibrillation? Front Cardiovasc Med. 2019;6:82. doi: 10.3389/fcvm.2019.00082.
2. Kim SM, Jeong Y, Kim YL, et al. Association of Chronic Kidney Disease With Atrial Fibrillation in the General Adult Population: A Nationwide Population-Based Study. J Am Heart Assoc. 2023;12(8):e028496. doi: 10.1161/JAHA.122.028496.
3. Power DA, Lampert J, Camaj A, et al. Cardiovascular Complications of Interatrial Conduction Block: JACC State-of-the-Art Review. J Am Coll Cardiol. 2022;79(12):1199-1211. doi: 10.1016/j.jacc.2022.01.030.
4. Alexander B, Milden J, Hazim B, et al. New electrocardiographic score for the prediction of atrial fibrillation: The MVP ECG risk score (morphology-voltage-P-wave duration). Ann Noninvasive Electrocardiol. 2019;24(6):e12669. doi: 10.1111/anec.12669.
5. Hayıroğlu Mİ, Çınar T, Selçuk M, et al. The significance of the morphology-voltage-P-wave duration (MVP) ECG score for prediction of in-hospital and long-term atrial fibrillation in ischemic stroke. J Electrocardiol. 2021;69:44-50. doi: 10.1016/j.jelectrocard.2021.09.006.

6. Yang N, Yan N, Cong G, Yang Z, Wang M, Jia S. Usefulness of Morphology-Voltage-P-wave duration (MVP) score as a predictor of atrial fibrillation recurrence after pulmonary vein isolation. Ann Noninvasive Electrocardiol. 2020;25(6):e12773. doi: 10.1111/anec.12773.
7. Kahyaoglu M, Gecmen C, Candan O, et al. The usefulness of morphology-voltage-P wave duration ECG score for predicting early left atrial dysfunction in hypertensive patients. Clin Exp Hypertens. 2021;43(6):572-578. doi: 10.1080/10641963.2021.1916945.
8. Pay L, Yumurtaş AÇ, Tezen O, et al. Efficiency of MVP ECG Risk Score for Prediction of Long-Term Atrial Fibrillation in Patients With ICD for Heart Failure With Reduced Ejection Fraction. Korean Circ J. 2023;53(9):621-631. doi: 10.4070/kcj.2022.0353.
9. Xing LY, Diederichsen SZ, Højberg S, et al. Electrocardiographic Morphology-Voltage-P-Wave-Duration (MVP) Score to Select Patients for Continuous Atrial Fibrillation Screening to Prevent Stroke. Am J Cardiol. 2023;205:457-464. doi: 10.1016/j.amjcard.2023.08.042.
10. Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr. 2015;28(1):1-39.e14. doi: 10.1016/j.echo.2014.10.003.
11. Inker LA, Eneanya ND, Coresh J, et al; Chronic Kidney Disease Epidemiology Collaboration. New Creatinine- and Cystatin C-Based Equations to Estimate GFR without Race. N Engl J Med. 2021;385(19):1737-1749. doi: 10.1056/NEJMoa2102953.
12. Rapa SF, Di Iorio BR, Campiglia P, Heidland A, Marzocco S. Inflammation and Oxidative Stress in Chronic Kidney Disease-Potential Therapeutic Role of Minerals, Vitamins and Plant-Derived Metabolites. Int J Mol Sci. 2019;21(1):263. doi: 10.3390/ijms21010263.
13. Dumler F, Kilates C. Body composition analysis by bioelectrical impedance in chronic maintenance dialysis patients: comparisons to the National Health and Nutrition Examination Survey III. J Ren Nutr. 2003;13(2):166-172. doi: 10.1053/jren.2003.50022.
14. Gadde S, Kalluru R, Cherukuri SP, Chikatimalla R, Dasaradhan T, Koneti J. Atrial Fibrillation in Chronic Kidney Disease: An Overview. Cureus. 2022;14(8):e27753. doi: 10.7759/cureus.27753.
15. Qiu H, Ji C, Liu W, et al. Chronic Kidney Disease Increases Atrial Fibrillation Inducibility: Involvement of Inflammation, Atrial Fibrosis, and Connexins. Front Physiol. 2018;9:1726. doi: 10.3389/fphys.2018.01726.
16. Huang JC, Wei SY, Chen SC, et al. P wave dispersion and maximum P wave duration are associated with renal outcomes in chronic kidney disease. PLoS One. 2014;9(7):e101962. doi: 10.1371/journal.pone.0101962.
17. Kumar Pandey U, Chaudhary N, Singh S. Utility of P Wave Indices on Electrocardiography in Predicting the Left Atrial Volume Index in Chronic Kidney Disease Patients. Cureus. 2025;17(1):e77809. doi: 10.7759/cureus.77809.
18. Law JP, Pickup L, Pavlovic D, Townend JN, Ferro CJ. Hypertension and cardiomyopathy associated with chronic kidney disease: epidemiology, pathogenesis and treatment considerations. J Hum Hypertens. 2023;37(1):1-19. doi: 10.1038/s41371-022-00751-4.
19. Istolahti T, Eranti A, Huhtala H, et al. The prevalence and prognostic significance of interatrial block in the general population. Ann Med. 2020;52(3-4):63-73. doi: 10.1080/07853890.2020.1731759.
20. Murariu E, Frigy A. Bayés' Syndrome-A Comprehensive Short Review. Medicina (Kaunas). 2020;56(8):410. doi: 10.3390/medicina56080410.

The European Research Journal 2026;12(3):315-322 Original Article

DOI: https://doi.org/10.18621/eurj.1828131 *Orthopedics and Traumatology*

Impact of Radial Bowing Alteration on Functional Outcomes Following Locked Intramedullary Nailing of Adult Both-Bone Forearm Fractures

Mehmed Nuri Tütüncü[1], Muhammed Muvahhid Sevgin[2], Salim Çağatay Akbulut[3], Fuat Akpınar[2]

[1]Department of Orthopedics and Traumatology, Göztepe Prof. Dr. Süleyman Yalçın City Hospital, İstanbul, Türkiye; [2]Department of Orthopedics and Traumatology, İstanbul Medeniyet University, Faculty of Medicine, İstanbul, Türkiye; [3]Department of Orthopedics and Traumatology, Kurtalan State Hospital, Siirt, Türkiye

ABSTRACT

Objectives: Plate osteosynthesis is the standard treatment for adult both-bone forearm fractures, but locked intramedullary nailing (IMN) offers a minimally invasive alternative. Its impact on restoring radial bow and functional outcomes is unclear. This study assessed changes in radial bow magnitude and location after IMN and their relationship with functional outcomes.

Methods: Twenty-six adult patients who underwent locked IMN for diaphyseal both-bone forearm fractures between 2015 and 2024 were retrospectively analyzed. Demographic data, radiographic measurements, union time, functional scores, and range of motion were recorded. Radial bow magnitude and bow apex location were measured on standardized radiographs and compared between the operated and contralateral forearms. Correlation analyses were performed to evaluate associations between bow changes and functional outcomes.

Results: All fractures achieved union, with no cases of nonunion or radioulnar synostosis. The mean time of union was 10.7±1.8 weeks. Extensor pollicis longus (EPL) rupture occurred in two (7.7%) patients. Postoperative radiographs showed reduced bow magnitude (Δ=0.61±1.13 mm; P=0.010) and a distal shift in bow location (Δ = –5.38±5.39%; P=0.001) on the operated side. A significant negative correlation was found between distal bow displacement and supination (r = –0.561; P=0.003). No correlation was observed with pronation, VAS, or QuickDASH scores.

Conclusions: Locked IMN ensures reliable union and excellent function. While IMN changes radial bowing, only excessive distal displacement impairs supination.

Keywords: Intramedullary Nail, Forearm Fractures, Radial Bow, Range of Motion

Forearm fractures in adults are common and usually treated surgically. Open reduction and internal fixation (ORIF) with plates remains the standard due to reliable reduction and fixation [1], but has drawbacks such as vascular disruption, delayed union, implant irritation, refracture after removal, and soft-tissue issues [2, 3]. Intramedullary nailing (IMN) is a minimally invasive alternative that

Received: November 24, 2025 **Accepted:** January 20, 2026 **Available Online:** January 22, 2026 **Published:** March 1, 2026

How to cite this article: Tütüncü MN, Sevgin MM, Akbulut SÇ, Akpınar F. Impact of Radial Bowing Alteration on Functional Outcomes Following Locked Intramedullary Nailing of Adult Both-Bone Forearm Fractures. Eur Res J. 2026;12(3):315-322. doi: 10.18621/eurj.1828131

Corresponding author: Mehmed Nuri Tütüncü, MD., Phone: +90 216 606 52 00, E-mail: mnuritutuncu@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

preserves periosteal integrity and vascularity, requires less dissection, and results in fewer refractures and soft-tissue problems [4]. First-generation nails, however, often failed to provide rotational stability and had high nonunion rates [5].

Diaphyseal fractures of the radius and ulna are considered intra-articular fractures because of the unique biomechanics of the forearm. Proper axial, angular, and rotational alignment following such injuries is crucial for forearm function, particularly pronation–supination movements [4]. Therefore, anatomical reduction, restoration of axial length, and accurate reconstruction of the radial bow are considered critical for achieving successful functional outcomes [6]. Recently developed locked intramedullary nails, with proximal and distal locking options, have improved fixation strength by enhancing both rotational and axial stability. Several studies have reported decreased complication rates, shorter union times, and improved functional outcomes with these implants [7-9].

Some studies suggest locked intramedullary nailing and plate fixation offer comparable functional outcomes, but the effect of changes in radial bowing on forearm function remains unclear [10, 11]. This study evaluates changes in radial bow after locked intramedullary nailing of both-bone forearm fractures and examines their relationship with functional results.

METHODS

After obtaining local ethics committee approval (Number: 2025/0127 Date: 28/08/2025), 32 patients who underwent surgery for both forearm fractures in a single center between 2015 and 2024 were retrospectively analyzed. The study included patients who were older than 16 years, had sustained both traumatic forearm fractures, grade 1-2 open fractures, and had a follow-up period of at least 12 months. Exclusion criteria were a floating elbow, Monteggia or Galeazzi fracture, Grade 3 open fractures, pathologic fractures, rheumatoid arthritis, and previous limb fractures or deformities. Twenty-six patients meeting the inclusion criteria were enrolled in the study. Written informed consent was obtained from each patient. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Demographic and clinical data, including age, sex, fracture side, trauma mechanism, time of surgery, and follow-up duration, were collected from medical records. Radiographic data included fracture type (Arbeitsgemeinschaft für Osteosynthesefragen/ Orthopaedic Trauma Association [AO/OTA] classification), time to union, and radial bowing measurement.

The forearms of all participants were assessed using standardized anteroposterior (AP) and lateral radiographs taken at a distance of one meter in the same position. The site of the radial bow was determined by calculating the ratio between the distance from the midpoint of the bicipital tuberosity to the point of maximal radial curvature and the total length of the radius from the midpoint of the bicipital tuberosity to the most ulnar aspect of the distal radius, as previously described in the literature [11]. Identical radiographic positioning was ensured for both forearms. Patients were followed up with serial radiographs until radiological union was achieved. Fractures were classified as nonunions if, after 6 months from the injury, radiographs showed no evidence of healing and patients experienced either persistent pain or abnormal mobility at the fracture site.

All operations were performed by a single surgeon (FA), with 3 decades of trauma surgery experience, using interlocking forearm nails (TST Rakor Tıbbi Aletler, Istanbul, Turkey). The surgical procedure has been previously detailed in the literature. While the radial nail followed a uniform design, two variations of the ulnar nail were available, distinguished by their distal interlocking mechanisms. The appropriate diameter and length of the implants were selected based on both preoperative planning and intraoperative evaluation. Following fixation, passive assessment of supination, pronation, and distal radioulnar joint alignment was performed to verify reduction and radial bowing restoration. Postoperatively, patients were advised to begin early mobilization. Elbow and wrist movements were initiated immediately, whereas pronation and supination exercises were introduced after three weeks.

All patients included in the study were followed up for a minimum of 12 months. At each follow-up visit, clinical and radiographic evaluations were performed to assess bone union and functional recovery. Union was defined as the presence of both

clinical and radiographic findings, including callus formation bridging the fracture at three cortices, continuity of the cortex, disappearance of the fracture line, and absence of tenderness or pain on examination (Figure 1). Forearm pronation–supination, wrist flexion–extension, and elbow flexion–extension were measured with a goniometer. Functional status was assessed at the final follow-up using the Visual Analog Scale, shortened version of the Disabilities of the Arm, Shoulder and Hand (QuickDASH) questionnaire, and the Grace–Eversmann scoring system. According to the Grace–Eversmann classification, outcomes were rated as excellent (union and ≥90% contralateral pronation–supination), good (union and ≥80%), acceptable (union and ≥60%), or poor (nonunion and <60%). All examinations were performed by an independent observer (M.M.S.). All radiological measurements were performed independently by two orthopedic surgeons blinded to group allocation (MMS and SÇA). The mean values of their measurements were used for analysis. Interobserver reliability was evaluated using the intraclass correlation coefficient (ICC). Agreement was considered excellent if the ICC was between 0.75 and 1.00; good between 0.60 and 0.74; fair between 0.40 and 0.59; and poor below 0.40. The interobserver ICC value was 0.84, indicating excellent reliability.

Statistical Analysis

Statistical analyses were performed using NCSS (Number Cruncher Statistical System) 2020 Statistical Software (NCSS LLC, Kaysville, Utah, USA). Continuous variables were expressed as mean ± standard deviation, median, and minimum–maximum values, while categorical variables were presented as frequencies and percentages. The normality of data distribution was assessed using the Shapiro–Wilk test, as well as skewness–kurtosis values and Box Plot visual inspection. For paired quantitative variables with a normal distribution, the paired-samples t-test was used. Relationships between variables were evaluated using Spearman's rank correlation analysis. All statistical analyses were conducted at a 95% confidence interval, and a P-value <0.05 was considered statistically significant.

RESULTS

A total of 26 patients (7 females and 19 males) with a mean age of 35.0 ± 16.8 years (range, 18–73 years) were included in the study. The left side was affected in 76.9% of cases. The mean follow-up duration was 55.5±34.1 months (range, 12–120 months). Regarding

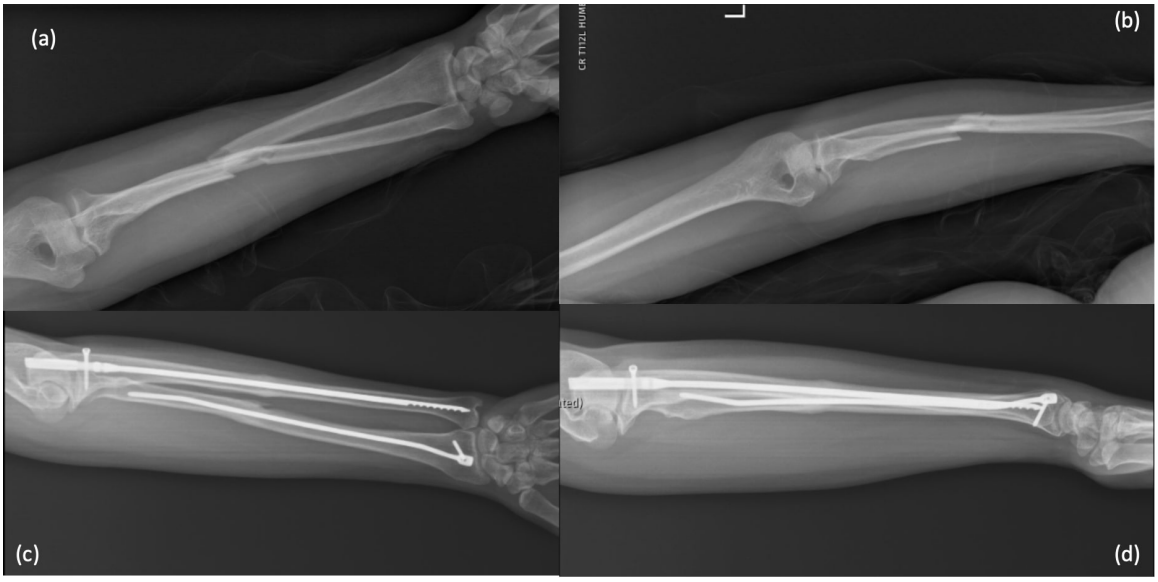

FIGURE 1. A 36-year-old female patient with a both-bone forearm fracture after a motor vehicle accident. (a) Preoperative anteroposterior radiograph. (b) Preoperative lateral radiograph. (c) Postoperative anteroposterior radiograph at final follow-up. (d) Postoperative lateral radiograph at final follow-up.

TABLE 1. Demographic and Clinical Characteristics of the Patients

		n (%) / Mean±SD	Median (Min–Max)
Sex	Female	7 (26.9%)	–
	Male	19 (73.1%)	–
Age (years)		35.04±16.78	29 (18–73)
Side	Right	6 (23.1%)	–
	Left	20 (76.9%)	–
Follow-up (months)		55.50±34.08	45 (12–120)
Time to union (weeks)		10.69±1.80	10 (8–13)
Complications	None	24 (92.3%)	–
	EPL tendon rupture	2 (7.7%)	–
VAS		0.85±0.34	0 (0–4)
QuickDASH score		4.54±6.66	2.26 (0–22.5)
Grace–Eversmann outcome	Excellent	24 (92.3%)	–
	Good	1 (3.8%)	–
	Acceptable	1 (3.8%)	–

EPL, extensor pollicis longus; VAS, Visual Analog Scale; QuickDASH, Shortened version of the Disabilities of the Arm, Shoulder, and Hand; SD, standard deviation; Min, minimum; Max, maximum.

the AO/OTA classification of the forearm fractures (n=26), the most common fracture patterns were 2R2A2-2U2A2 and 2R2A3-2U2A3, each accounting for 34.6% (n=9) of the cases. These were followed by the 2R2B3-2U2A2 pattern, which was observed in 11.5% (n=3) of the patients. The remaining fracture types (2R2A1-2U2A1, 2R2A2-2U2B3, 2R2B2-2U2B2, 2R2B3-2U2B3, and 2R2C2-2U2A2) each represented 3.8% (n=1) of the study population.

The average time to radiological union was 10.7±1.8 weeks. All fractures achieved union, and no cases of nonunion or radioulnar synostosis were observed. Two (7.7%) patients developed extensor pollicis longus (EPL) tendon rupture as a postoperative complication. One patient presented with acute EPL rupture in the early postoperative period and was successfully treated with tendon transfer. The other patient was diagnosed with a chronic EPL rupture at late follow-up but declined further surgical intervention (Table 1).

At final follow-up, 92.3% had excellent, 3.8% good, and 3.8% acceptable Grace–Eversmann outcomes. Mean VAS was 0.85±1.34 and QuickDASH 4.54±6.66, indicating minimal pain and disability (Table 2).

Radiographic evaluation showed that the mean

TABLE 2. Functional and Radiographic Measurements

	Mean±SD	Median (Min–Max)
Supination (°)	75.77±13.09	80 (15–80)
Pronation (°)	79.23±2.72	80 (70–80)
Radial bow magnitude–operated side (mm)	7.22±1.62	7.5 (4.0–10.1)
Radial bow magnitude–intact side (mm)	7.84±1.70	8.4 (5.1–10.5)
Bow location–operated side (% of radial length)	67.36±4.90	67 (58.6–76.4)
Bow location–intact side (% of radial length)	61.98±3.89	62.1 (54.4–70.2)

SD, standard deviation; Min, minimum; Max, maximum.

TABLE 3. Comparison of Bow Magnitude and Location Between Operated and Intact Sides

	Operated	Intact	Δ	P-value*
Magnitude (mm)	7.22±1.62	7.84±1.70	0.61±1.13	**0.010**
Location (% of length)	67.36±4.90	61.98±3.89	–5.38±5.39	**<0.001**

Data are shown as mean±standard deviation.
*Paired-samples t-test, Statistically significant P-values are shown in bold.

radial bow magnitude on the operated side was slightly lower than that of the contralateral forearm, while the bow apex was located more distally (Table 2). The differences were statistically significant for both bow magnitude (Δ=0.61±1.13 mm; P=0.010) and bow location (Δ= –5.38±5.39%; P=0.001) (Table 3). Correlation analysis demonstrated a significant negative relationship between the change in bow location (Δ) and the degree of supination (r = –0.561; P=0.003), suggesting that a more distally displaced radial bow was associated with decreased supination (Table 4). No statistically significant correlations were observed between bowing parameters and pronation, VAS, or QuickDASH scores (P>0.05) (Table 4).

DISCUSSION

The primary finding of this study is that IMN of adult both-bone forearm fractures achieved union in all cases, without any instances of nonunion or radioulnar synostosis, while providing excellent functional outcomes. However, a statistically significant negative correlation was detected between distal displacement of the radial bow and supination range, indicating that alterations in the anatomical curvature of the radius may impair rotational function.

Although plate osteosynthesis has long been considered the gold standard for treating adult forearm fractures, intramedullary nailing has gained increasing popularity and acceptance in recent years. Recent systematic reviews and meta-analyses have reinforced the clinical equivalence of IMN and ORIF in adult forearm fractures. Box *et al*. [12] and Zhao *et al*. [13] reported that IMN achieves union rates and functional outcomes comparable to those with plating, while offering shorter operative time and a lower implant removal rate in both forearm fractures. Lari *et al*. [14] concluded that interlocking nails provide sufficient radiographic restoration and patient-reported outcomes while causing less soft-tissue damage. This study supports these findings, as all fractures in this series achieved union, and functional recovery was excellent in over 90% of cases.

In the present study, the operated forearms demonstrated a reduction in radial bow magnitude and a distal shift in bow location compared with the contralateral side. However, these alterations did not appear to affect functional scores or fracture healing. The only clinically relevant finding was that a more distally displaced radial bow was associated with decreased forearm supination. Prior studies have reported varying interpretations of postoperative bow

Table 4. Correlation Between Changes in Radial Bow and Functional Outcomes

	Δ Magnitude (Operated–Intact)		Δ Location (Operated–Intact)	
	r	P-value	r	P-value
VAS	0.132	0.520	0.078	0.706
QuickDASH	0.023	0.911	0.324	0.107
Supination (°)	0.099	0.630	–0.561	**0.003**
Pronation (°)	–0.135	0.512	–0.212	0.299

VAS, Visual Analog Scale; QuickDASH, Shortened version of the Disabilities of the Arm, Shoulder, and Hand; r, Spearman's correlation. Statistically significant P-value is shown in bold.

alteration. Saka *et al*. [15] noted similar distal migration and reduction in magnitude after IMN but found no functional impact. Köse *et al*. [11] reported no significant differences in radial bow parameters between IMN and plating. Conversely, Yörükoğlu *et al*. [16] observed a significant association between bow alteration and nonunion, while Dave *et al*. [17] reported that reduced bow magnitude may compromise supination. Collectively, the literature suggests that IMN typically reduces bow magnitude and shifts its apex distally, but the functional consequences of these changes remain inconsistent.

In this study, no cases of nonunion, clinically significant malunion, or radioulnar synostosis were observed; however, two (7.7%) cases of extensor pollicis longus (EPL) tendon rupture were identified. In the patient with acute EPL rupture, early surgical intervention was initially recommended but postponed due to a concurrent pulmonary disease. Reconstruction was subsequently performed nine months after the index procedure using extensor indicis proprius tendon transfer. No entry point malposition or distal nail protrusion was identified in this case, suggesting inadvertent tendon injury during dorsal entry point exposure as the most likely mechanism. In contrast, the patient with delayed EPL rupture declined reoperation, and follow-up radiographs demonstrated distal nail prominence, indicating chronic mechanical irritation as a probable cause.

EPL rupture is a recognized complication of intramedullary nailing, attributed to the close anatomical relationship between the dorsal entry point and the EPL tendon, resulting either from inadvertent intraoperative injury or delayed attritional rupture due to distal nail prominence or bony overgrowth. Reported EPL injury rates vary widely in the literature, ranging from 4.8% to 23.6%, while a recent meta-analysis reported a pooled incidence of approximately 1% [14, 16, 18]. Management of EPL rupture is guided by patient-related factors and the timing of presentation and may include direct tendon repair, tendon transfer, or tendon grafting [19]. Taken together, these findings indicate that EPL rupture following intramedullary nailing may occur through timing-dependent mechanisms and underscore the importance of meticulous soft-tissue handling, accurate entry point selection, appropriate nail length determination, and adequate recessing of the nail tip to prevent tendon-related complications.

Strengths and Limitations

This study has several limitations that should be acknowledged. First, the relatively small sample size may limit the statistical power of correlation analyses and subgroup evaluations, increasing the risk of Type II error and restricting the generalizability of the findings. Second, the absence of a comparative control group treated with plate osteosynthesis precludes direct comparison between fixation methods; therefore, the observed anatomical and functional changes should be interpreted as findings following intramedullary nailing rather than effects specific or exclusive to this technique. Finally, the retrospective design of the study may be associated with inherent selection and reporting biases.

The strengths of this study include standardized radiographic measurements, excellent interobserver reliability, and a relatively long follow-up period.

CONCLUSION

Intramedullary nailing of both-bone forearm fractures was associated with reliable fracture union and favorable functional outcomes, with a low incidence of major complications in this series. The findings suggest that alterations in the physiological contour and location of the radial bow may occur following intramedullary nailing, and excessive distal displacement may be associated with reduced forearm supination. Although these observations do not imply technique-specific effects when compared with other fixation methods, careful attention to restoration of radial bow anatomy, appropriate nail length selection, and adequate distal nail management may help optimize functional outcomes and reduce the risk of extensor tendon-related complications.

Ethics Approval and Consent to Participate

The study was approved by the Göztepe Prof. Dr. Süleyman Yalçın City Hospital Non-Interventional Clinical Research Ethics Committee (Decision no.: 2025/0127 and date: 28.08.2025). All procedures performed during data collection, review of patient

records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Written informed consent was obtained from all individual participants included in the study.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: MNT, SÇA, MMS; Study Design: MMS, MNT, FA, SÇA; Supervision: FA, MNT; Funding: N/A; Materials: FA, MNT; Data Collection and/or Processing: MNT, MMS, SÇA; Statistical Analysis and/or Data Interpretation: MNT, FA; Literature Review: MNT, FA Manuscript Preparation: MNT, SÇA, MMS; and Critical Review: FA, MNT.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

We would like to express our sincere gratitude to Dr. Tolga Onay for his valuable advice, insightful comments, and meaningful contributions to this study.

Generative Artificial Intelligence Statement

The authors declare that artificial intelligence-based tools were used solely for language editing and grammar correction during the manuscript preparation process. No content generation, data analysis, or scientific interpretation was performed using artificial intelligence.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Jones DB Jr, Kakar S. Adult diaphyseal forearm fractures: intramedullary nail versus plate fixation. J Hand Surg Am. 2011;36(7):1216-1219. doi: 10.1016/j.jhsa.2011.03.020.
2. Azboy I, Demirtas A, Uçar BY, Bulut M, Alemdar C, Ozkul E. Effectiveness of locking versus dynamic compression plates for diaphyseal forearm fractures. Orthopedics. 2013;36(7):e917-922. doi: 10.3928/01477447-20130624-23.
3. Saikia K, Bhuyan S, Bhattacharya T, Borgohain M, Jitesh P, Ahmed F. Internal fixation of fractures of both bones forearm: Comparison of locked compression and limited contact dynamic compression plate. Indian J Orthop. 2011;45(5):417-421. doi: 10.4103/0019-5413.83762.
4. Uygur E, Özkut A, Akpınar F. Synostosis after fracture of both forearm bones treated by intramedullary nailing. Hand Surg Rehabil. 2021;40(1):25-31. doi: 10.1016/j.hansur.2020.07.002.
5. Sage FP. Medullary fixation of fractures of the forearm. A study of the medullary canal of the radius and a report of fifty fractures of the radius treated with a prebent triangular nail. J Bone Joint Surg Am. 1959;41-A:1489-1516.
6. Schemitsch EH, Richards RR. The effect of malunion on functional outcome after plate fixation of fractures of both bones of the forearm in adults. J Bone Joint Surg Am. 1992;74(7):1068-1078.
7. Bansal H. Intramedullary fixation of forearm fractures with new locked nail. Indian J Orthop. 2011 Sep;45(5):410-6. doi: 10.4103/0019-5413.83760.
8. Gao H, Luo CF, Zhang CQ, Shi HP, Fan CY, Zen BF. Internal fixation of diaphyseal fractures of the forearm by interlocking intramedullary nail: short-term results in eighteen patients. J Orthop Trauma. 2005;19(6):384-391. doi: 10.1097/01.bot.0000157911.76433.db.
9. Lee YH, Lee SK, Chung MS, Baek GH, Gong HS, Kim KH. Interlocking contoured intramedullary nail fixation for selected diaphyseal fractures of the forearm in adults. J Bone Joint Surg Am. 2008;90(9):1891-1898. doi: 10.2106/JBJS.G.01636.
10. Polat O, Toy S. Comparison of the clinical and radiographic outcomes of plate fixation versus new-generation locked intramedullary nail in the management of adult forearm diaphyseal fractures. Acta Orthop Traumatol Turc. 2022;56(5):321-326. doi: 10.5152/j.aott.2022.21190.
11. Köse A, Aydın A, Ezirmik N, Yıldırım ÖS. A comparison of the treatment results of dpen reduction internal fixation and intramedullary nailing in adult forearm diaphyseal fractures. Ulus Travma Acil Cerrahi Derg. 2017;23(3):235-244. doi: 10.5505/tjtes.2016.66267.
12. Box MW, Stegelmann SD, Domingue GA, et al. Intramedullary nail fixation versus open reduction and internal fixation for treatment of adult diaphyseal forearm fractures: a systematic review and meta-analysis. J Orthop Surg Res.

2024;19(1):719. doi: 10.1186/s13018-024-05158-0.
13. Zhao L, Wang B, Bai X, Liu Z, Gao H, Li Y. Plate Fixation Versus Intramedullary Nailing for Both-Bone Forearm Fractures: A Meta-analysis of Randomized Controlled Trials and Cohort Studies. World J Surg. 2017;41(3):722-733. doi: 10.1007/s00268-016-3753-1.
14. Lari A, Hassan Y, Altammar A, et al. Interlocking intramedullary nail for forearm diaphyseal fractures in adults-A systematic review and meta-analysis of outcomes and complications. J Orthop Traumatol. 2024;25(1):16. doi: 10.1186/s10195-024-00761-7.
15. Saka G, Saglam N, Kurtulmuş T, et al. New interlocking intramedullary radius and ulna nails for treating forearm diaphyseal fractures in adults: a retrospective study. Injury. 2014;45 Suppl 1:S16-23. doi: 10.1016/j.injury.2013.10.040.
16. Yörükoğlu AÇ, Demirkan AF, Akman A, Kitiş A, Usta H. The effects of radial bowing and complications in intramedullary nail fixation of adult forearm fractures. Eklem Hastalik Cerrahisi. 2017;28(1):30-34. doi: 10.5606/ehc.2017.05.
17. Dave MB, Parmar KD, Sachde BA. The Radial Bow following Square Nailing in Radius and Ulna Shaft Fractures in Adults and its Relation to Disability and Function. Malays Orthop J. 2016;10(2):11-15. doi: 10.5704/MOJ.1607.003.
18. Blažević D, Benčić I, Ćuti T, et al. Intramedullary nailing of adult forearm fractures: Results and complications. Injury. 2021;52 Suppl 5:S44-S48. doi: 10.1016/j.injury.2020.11.012.
19. Ganon S, Bellity J, Zbili D, Boccara D. Reconstruction strategies after rupture of the extensor pollicis longus tendon: A systematic review. Hand Surg Rehabil. 2020;39(6):502-507. doi: 10.1016/j.hansur.2020.07.005.

The European Research Journal 2026;12(3):323-328 Original Article

DOI: https://doi.org/10.18621/eurj.1829668 *Dermatology*

Distribution of Blood Groups in Patients with Atopic Dermatitis: A Retrospective Case-Control Study

Candan Çelik[1], Mehmet Semih Çelik[2]

[1]Department of Dermatology, Malatya Training and Research Hospital, Malatya, Türkiye; [2]Department of Dermatology, University of Health Sciences, Diyarbakır Gazi Yaşargil Training and Research Hospital, Diyarbakır, Türkiye

ABSTRACT

Objectives: Atopic dermatitis is a chronic inflammatory skin condition with a multifactorial etiology involving genetic, environmental, and immunologic influences. Although several mechanisms have been proposed, the potential association between ABO/Rh blood groups and the risk of developing atopic dermatitis is still uncertain. The present study sought to investigate the distribution of ABO and Rh blood groups in individuals with atopic dermatitis compared with healthy, age- and sex-matched controls.

Methods: A retrospective analysis was performed covering a 10-year interval (2014–2024). The study included 3,470 patients diagnosed with atopic dermatitis and 7,449 matched controls, yielding a total sample of 10,919 individuals. Demographic data and blood group characteristics (ABO and Rh typing) were collected and evaluated.

Results: The O Rh– phenotype was observed significantly more often in patients with atopic dermatitis than in controls (5.2% vs. 3.5%; odds ratio [OR]=1.52, 95% confidence interval [CI]: 1.25–1.85; P<0.001). Conversely, the A Rh– blood group was notably less frequent in atopic dermatitis patients compared with the control group (3.7% vs. 4.7%; OR=0.79, 95% CI: 0.64–0.96; P=0.021).

Conclusions: The findings indicate that the general distribution of ABO and Rh blood groups among patients with atopic dermatitis is comparable to that of healthy individuals. However, the increased prevalence of the O Rh– type and the lower frequency of A Rh– among atopic dermatitis patients suggest that specific blood group antigens may influence susceptibility to atopic dermatitis. Future large-scale, multicenter, and prospective studies are required to further elucidate this potential relationship.

Keywords: Atopic Dermatitis, ABO Blood Group, Rh Factor, Retrospective Study

Atopic dermatitis (AD) is a chronic, recurrent, and remitting inflammatory skin disease that affects one in ten people during their lifetime. Atopic dermatitis results from a complex interaction of immune dysregulation, epidermal gene mutations, and environmental factors that disrupt the epidermis, leading to intensely itchy skin lesions. Repeated scratching triggers a self-sustaining itch–scratch cycle, which can have a significant impact on the patient's quality of life [1]. The etiology is complex and not fully understood. Immune dysregulation, including abnormalities in the skin barrier and excessive activity of T helper 2 (TH2 and TH22) cells, is thought to contribute, along with genetic and environmental factors [2].

Received: November 24, 2025 **Accepted:** January 4, 2026 **Available Online:** January 6, 2026 **Published:** March 1, 2026

How to cite this article: Çelik C, Çelik MS. Distribution of Blood Groups in Patients with Atopic Dermatitis: A Retrospective Case-Control Study. Eur Res J. 2026;12(3):323-328. doi: 10.18621/eurj.1829668

Corresponding author: Mehmet Semih Çelik, MD., Phone: +90 412 258 00 60, E-mail: drmsemihcelik@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

The diagnosis of atopic dermatitis is based on clinical presentation and medical history, after excluding multiple erythematous and eczematous diseases. Although diagnosing AD is generally easier in infants and young children, it can become difficult in severe cases and in adults. Adult-onset AD tends to present with more diverse clinical features. The morphology and distribution of lesions are more variable, with a greater likelihood of involvement of the head, neck, hands, and feet. Over time, multiple diagnostic criteria have been proposed to support the identification of AD. The Hanifin–Rajka (HR) criteria are extensive and are widely regarded as the "gold standard" for diagnosing AD [3, 4]. Blood groups have been suggested to be associated with AD and other allergic conditions [5].

In this study, we aimed to explore this potential association by evaluating the distribution of blood groups among patients with AD.

METHODS

This retrospective study was approved by the Clinical Research Ethics Committee of the University of Health Sciences, Gazi Yaşargil Training and Research Hospital (Approval No: 440/2025, Date: 25 April 2025). The study was conducted in accordance with the principles of the Declaration of Helsinki. Due to the retrospective design of the study, the requirement for informed consent was waived by the ethics committee.

The study population was retrospectively identified from patients who attended the dermatology outpatient clinic during the study period. Patients diagnosed with atopic dermatitis were included in the case group, while the control group consisted of individuals who attended the same clinic during the same period and had no diagnosis of atopic dermatitis or other chronic inflammatory skin diseases. Control subjects were selected from patients with non-inflammatory dermatological conditions. Age and sex distributions were comparable between the atopic dermatitis and control groups. A sum of 3,470 patients diagnosed with AD according to the Hanifin and Rajka diagnostic criteria were included in the study. Patients who did not meet the diagnostic criteria or had missing demographic or laboratory data were excluded.

Demographic data (age, sex) and clinical data (blood group distribution: ABO and Rh factor) were obtained from hospital records. The diagnosis of AD was confirmed based on the fulfillment of at least three major and three minor criteria as proposed by Hanifin and Rajka [3] (Table 1).

TABLE 1. Hanifin and Rajka Diagnostic Criteria for Atopic Dermatitis

Major Criteria (at least 3 required)	Minor Criteria (at least 3 required)
Pruritus	Early age of onset
Typical morphology and distribution	Xerosis
Chronic or chronically relapsing course	Ichthyosis
Personal or family history of atopy	Immediate skin test reactivity
	Elevated serum IgE
	Tendency toward cutaneous infections
	Nonspecific hand/foot dermatitis
	Nipple eczema
	Cheilitis
	Recurrent conjunctivitis
	Dennie-Morgan infraorbital folds
	Keratoconus
	Anterior subcapsular cataracts
	Orbital darkening ('allergic shiners')

Statistical Analysis

Statistical analyses were conducted using the JAMOVI software (version 2.3). Categorical variables were summarized with frequency and percentage values. Differences between patients with AD and the control group were assessed using the chi-square (χ^2) test. To evaluate the association between blood groups and AD, odds ratios (OR) and relative risks (RR) along with their 95% confidence intervals (CIs) were calculated. A P-value below 0.05 was considered indicative of statistical significance. The obtained P-values were organized and interpreted in a table format.

RESULTS

Overall, the study included 10,919 participants, consisting of 3,470 patients with atopic dermatitis and 7,449 controls. The mean age was 38.2±15.2 years in the patient group and 41.8±15.7 years in the control group, with a range of 18 to 75 years in both groups. The median ages were 35.0 and 39.0 years, respectively. Regarding sex distribution, females were more frequent in both groups. In the atopic dermatitis group, there were 1,156 (33.3%) males and 2,314 (66.7%) females. In the control group, there were 2,793 (37.5%) males and 4,656 (62.5%) females.

The distribution of blood groups among patients and controls is also presented, with O Rh+ and A Rh+ being the most common blood groups, and AB Rh– the least frequent in both groups (Table 2). The frequency of A Rh– was lower in patients with atopic dermatitis (3.7%) compared to controls (4.7%). Case-control analysis indicated that A Rh– was significantly underrepresented among patients (odds ratio [OR]=0.79, 95% confidence interval [CI]: 0.64–0.96, P=0.021). The frequency of A Rh+ did not differ significantly between patients with atopic dermatitis (33.0%) and controls (33.8%) (OR=1.04, 95% CI: 0.95–1.13, P=0.398).

The frequency of B Rh– did not differ significantly between patients with atopic dermatitis (1.7%) and controls (1.6%) (OR=0.93, 95% CI: 0.68–1.27, P=0.652). The frequency of B Rh+ was similar between patients with atopic dermatitis (12.4%) and controls (12.2%) (OR=0.98, 95% CI: 0.86–1.10, P=0.685).

The frequency of O Rh– was higher in patients with atopic dermatitis (5.2%) compared to controls

TABLE 2. Demographic and Clinical Characteristics of the Participants

Variable	Atopic Dermatitis (n=3470)	Controls (n=7449)	Total (n=10919)
Age (years)	38.2±15.2 35 (18–75)	41.8±15.7 39 (18–75)	-
Gender, n (%)			
Male	1156 (33.3%)	2793 (37.5%)	3949 (36.2%)
Female	2314 (66.7%)	4656 (62.5%)	6970 (63.8%)
Blood groups, n (%)			
A Rh–	129 (3.7%)	349 (4.7%)	478 (4.4%)
A Rh+	1145 (33.0%)	2519 (33.8%)	3664 (33.6%)
B Rh–	60 (1.7%)	120 (1.6%)	180 (1.6%)
B Rh+	432 (12.4%)	907 (12.2%)	1339 (12.3%)
O Rh–	179 (5.2%)	257 (3.5%)	436 (4.0%)
O Rh+	1301 (37.5%)	2822 (37.9%)	4123 (37.8%)
AB Rh–	25 (0.7%)	43 (0.6%)	68 (0.6%)
AB Rh+	199 (5.7%)	432 (5.8%)	631 (5.8%)

Data are shown mean±standard deviation or median (minimum-maximum) or n (%) where appropriate.

TABLE 3. Demographic and Clinical Characteristics of the Study Population

Blood groups	Atopic Dermatitis n (%)	Controls n (%)	Odds ratio	95% CI	P-value
A Rh–	129 (3.7%)	349 (4.7%)	0.79	0.64–0.96	**0.021**
A Rh+	1145 (33.0%)	2519 (33.8%)	1.04	0.95–1.13	0.398
B Rh–	60 (1.7%)	120 (1.6%)	0.93	0.68–1.27	0.652
B Rh+	432 (12.4%)	907 (12.2%)	0.98	0.86–1.10	0.685
O Rh–	179 (5.2%)	257 (3.5%)	1.52	1.25–1.85	**<0.001**
O Rh+	1301 (37.5%)	2822 (37.9%)	1.02	0.94–1.10	0.694
AB Rh–	25 (0.7%)	43 (0.6%)	0.80	0.49–1.31	0.376
AB Rh+	199 (5.7%)	432 (5.8%)	1.01	0.85–1.20	0.893

CI, confidence interval. Statistically significant P-values are shown in bold.

(3.5%). Case-control analysis showed that O Rh– was significantly associated with atopic dermatitis (OR=1.52, 95% CI: 1.25–1.85, P<0.001). The frequency of O Rh+ was similar in patients with atopic dermatitis (37.5%) and controls (37.9%) (OR=1.02, 95% CI: 0.94–1.10, P=0.694).

The frequency of AB Rh– did not differ significantly between patients with atopic dermatitis (0.7%) and controls (0.6%) (OR=0.80, 95% CI: 0.49–1.31, P=0.376). The frequency of AB Rh+ did not differ significantly between patients with atopic dermatitis (5.7%) and controls (5.8%) (OR=1.01, 95% CI: 0.85–1.20, P=0.893) (Table 3).

DISCUSSION

In this study, the frequency of O Rh– blood group was found to be significantly higher among individuals with AD compared with the control group, whereas the A Rh– blood group was significantly less common. These results raise additional questions regarding the potential influence of blood group types on allergic disorders, including AD.

AD is a chronic, inflammatory, and non-contagious skin disease, characterized primarily by persistent pruritus. Its pathophysiology is multifaceted and involves numerous interacting mechanisms, making it difficult to fully elucidate. Key contributing factors include genetic predisposition, impairment of the epidermal barrier, dysregulated immune pathways, and disturbances in the skin microbiome [6]. IgE is also believed to play a role in the disease process. Additionally, accumulating evidence indicates a possible involvement of TLR2 in the development of AD [7, 8]. Associations have further been reported between TLR2 expression and levels of the high-affinity IgE receptor (FcεRI) [9].

A key feature of AD is the tendency of $CD4^+$ lymphocytes to differentiate predominantly into the Th2 subtype. An overactive Th2 response results in elevated levels of the cytokines (interleukin [IL]-4, IL-5, and IL-13). These cytokines promote the production of immunoglobulin E (IgE) antibodies and activate eosinophils in both the bloodstream and tissues. The resulting inflammatory process further impairs the epidermal barrier, compounding existing primary barrier defects, and this interplay is thought to contribute to the pathogenesis of AD [10, 11].

Blood groups have also been studied in certain allergic diseases associated with elevated IgE. One such disease is allergic rhinitis, where research has identified a meaningful association between blood group O of the ABO system and allergic rhinitis [12, 13]. Similarly, in another IgE-related disease, asthma, it has been noted that individuals with blood group O show increased susceptibility to asthma [14–16]. In a systematic review, the aim was to examine the links between ABO blood groups and allergic diseases to investigate the root causes of immunological mechanisms. It was shown that blood group antigens significantly influence immune responses through interactions with IgE and differences in cytokine profiles [17]. Blood groups are also thought to alter

the immune response by modifying the gut microbiota, thereby affecting allergy risk [18]. Elevated IgE levels have been observed to be involved in the development of allergic diseases, including AD [19].

Studies have also reported an increased prevalence of allergic diseases in individuals with blood group O. This is thought to be related to alterations in immune recognition mechanisms and IgE binding due to alterations in the glycosylation of ABO antigens. Furthermore, in another study, Rh-negative individuals were observed to have significantly higher rates of allergic diseases, similar to the findings in our study [12, 13, 20]. It is also thought that individuals with blood group O might play a role in the onset of allergic diseases through the Th2 cytokines, IL-4 and IL-13, which function as mediators responsible for IgE release, thereby increasing IgE levels [21, 22].

Strengths and Limitations

The main strengths of this study include its large sample size, covering a 10-year period, and the inclusion of a substantial, age- and sex-matched control group from the same clinical setting. The evaluation of both ABO and Rh blood group systems and the use of well-established diagnostic criteria for atopic dermatitis further strengthen the reliability of the findings. In addition, the availability of detailed blood group data allowed for a robust statistical comparison, enhancing the validity of the observed associations.

This study has several limitations. First, its retrospective design precludes the establishment of a causal relationship. Additionally, the diagnosis of atopic dermatitis was based on established clinical diagnostic criteria rather than supportive laboratory biomarkers. Genetic and environmental factors, as well as comorbid conditions, were not comprehensively evaluated. Moreover, only the ABO and Rh blood group systems were assessed, while other potentially relevant blood group systems were not examined. The study period also included the COVID-19 pandemic, which may have influenced outpatient clinic attendance patterns; however, this factor is unlikely to have directly affected blood group distribution. Finally, as this was a single-center study, the generalizability of the findings may be limited. Therefore, the results should be interpreted with caution and validated through future multicenter prospective studies with larger and more comprehensive datasets.

CONCLUSION

In the present study, patients with atopic dermatitis exhibited a notably higher prevalence of the O Rh (–) blood group and a lower prevalence of the A Rh (–) blood group compared to controls. These observations imply that ABO and Rh blood groups could be involved in the development of atopic dermatitis. The data particularly highlight the potential susceptibility of individuals with the O Rh (–) blood group to this condition. However, additional prospective studies with larger populations are necessary to better understand the association between blood groups and allergic diseases.

Ethics Approval and Consent to Participate

This study was approved by the University of Health Sciences, Gazi Yaşargil Training and Research Hospital Clinical Research Ethics Committee (Decision No: 2025/440; date: 25.04.2025. All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Because the study was retrospective and no additional intervention was performed on the participants, the informed consent form was waived.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: MSÇ; Study Design: MSÇ; Supervision: MSÇ; Funding: CÇ; Materials: CÇ; Data Collection and/or Processing: CÇ; Statistical Analysis and/or Data Interpretation: MSÇ, CÇ; Literature Review: MSÇ, CÇ; Manuscript Preparation: MSÇ, CÇ; and Critical Review: MSÇ, CÇ.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Frazier W, Bhardwaj N. Atopic Dermatitis: Diagnosis and Treatment. Am Fam Physician. 2020;101(10):590-598.
2. Kim J, Kim BE, Leung DYM. Pathophysiology of atopic dermatitis: Clinical implications. Allergy Asthma Proc. 2019 Mar 1;40(2):84-92. doi: 10.2500/aap.2019.40.4202.
3. Fishbein AB, Silverberg JI, Wilson EJ, Ong PY. Update on Atopic Dermatitis: Diagnosis, Severity Assessment, and Treatment Selection. J Allergy Clin Immunol Pract. 2020;8(1):91-101. doi: 10.1016/j.jaip.2019.06.044.
4. Eichenfield LF, Tom WL, Berger TG, et al. Guidelines of care for the management of atopic dermatitis: section 2. Management and treatment of atopic dermatitis with topical therapies. J Am Acad Dermatol. 2014;71(1):116-132. doi: 10.1016/j.jaad.2014.03.023.
5. Dahalan NH, Tuan Din SA, Mohamad SMB. Association of ABO blood groups with allergic diseases: a scoping review. BMJ Open. 2020;10(2):e029559. doi: 10.1136/bmjopen-2019-029559.
6. Sroka-Tomaszewska J, Trzeciak M. Molecular mechanisms of atopic dermatitis pathogenesis. Int J Mol Sci. 2021;22(8):4130. doi: 10.3390/ijms22084130.
7. Kuo IH, Carpenter-Mendini A, Yoshida T, et al. Activation of epidermal toll-like receptor 2 enhances tight junction function: implications for atopic dermatitis and skin barrier repair. J Invest Dermatol. 2013;133(4):988-998. doi: 10.1038/jid.2012.437.
8. Gimalova GF, Karunas AS, Fedorova YY, et al. [Association of polymorphisms in toll-like receptor genes with atopic dermatitis in the Republic of Bashkortostan]. Mol Biol (Mosk). 2014;48(2):265-276. doi: 10.1134/S002689331402006X. [Article in Russian]
9. Song Z, Deng X, Chen W, et al. Toll-like receptor 2 agonist Pam3CSK4 up-regulates FcεRI receptor expression on monocytes from patients with severe extrinsic atopic dermatitis. J Eur Acad Dermatol Venereol. 2015;29(11):2169-2176. doi: 10.1111/jdv.13172.
10. Matsunaga MC, Yamauchi PS. IL-4 and IL-13 inhibition in atopic dermatitis. J Drugs Dermatol. 2016;15(8):925-929.
11. Kim KH. Overview of atopic dermatitis. Asia Pac Allergy. 2013;3(2):79-87. doi: 10.5415/apallergy.2013.3.2.79.
12. Topno N, Narvey VP, Jain AK. The correlation of allergic rhinitis with ABO phenotype. Indian J Otolaryngol Head Neck Surg. 2019;71(Suppl 3):1827-1831. doi: 10.1007/s12070-017-1215-1.
13. Falsarella N, Ferreira AI, Nakashima F, de Mattos CC, de Mattos LC. Evidence of an association between the O blood group and allergic rhinitis. Rev Bras Hematol Hemoter. 2011;33(6):444-448. doi: 10.5581/1516-8484.20110120.
14. Kauffmann F, Frette C, Pham QT, Nafissi S, Bertrand JP, Oriol R. Associations of blood group-related antigens to FEV1, wheezing, and asthma. Am J Respir Crit Care Med. 1996;153(1):76-82. doi: 10.1164/ajrccm.153.1.8542166.
15. Chen YL, Chen JC, Lin TM, et al. ABO/secretor genetic complex is associated with the susceptibility of childhood asthma in Taiwan. Clin Exp Allergy. 2005;35(7):926-932. doi: 10.1111/j.1365-2222.2005.02278.x.
16. Ronchetti F, Villa MP, Ronchetti R, et al. ABO/secretor genetic complex and susceptibility to asthma in childhood. Eur Respir J. 2001;17(6):1236-1238. doi: 10.1183/09031936.01.99109101.
17. Bahashwan AS. The mechanism underlying the association between ABO blood groups and allergic diseases: an evidence-based systematic review. Allergol Immunopathol (Madr). 2025;53(4):78-84. doi: 10.15586/aei.v53i4.1287.
18. Ewald DR, Sumner SCJ. Human microbiota, blood group antigens, and disease. Wiley Interdiscip Rev Syst Biol Med. 2018;10(3):e1413. doi: 10.1002/wsbm.1413.
19. Altaii HA, Al-Tae FMD. Investigation of serum total IgE and eosinophil levels in different allergic diseases together with the study of their correlations with various possible allergens in Mosul City. Immunol Invest. 2022;51(3):567-587. doi: 10.1080/08820139.2020.1853154.
20. Lal T, Sadhasivam M, K SPA, J AAS, R P, Khaleeluddin KB. An observational study to analyse the association of the ABO and Rh blood group systems with bronchial asthma. Cureus. 2023;15(4):e37675. doi: 10.7759/cureus.37675.
21. Tamayo-Velasco Á, Peñarrubia Ponce MJ, Álvarez FJ, et al. Can the cytokine profile according to ABO blood groups be related to worse outcome in COVID-19 patients? Yes, they can. Front Immunol. 2021;12:726283. doi: 10.3389/fimmu.2021.726283.
22. Abo-Shanab AM, Elnady H, Helwa I, et al. Role of Th2-type cytokines and IgE in asthmatic children. Biomed Pharmacol J. 2020;13(4):1765-1772. doi: 10.13005/bpj/2051.

The European Research Journal 2026;12(3):329-338 Original Article

DOI: https://doi.org/10.18621/eurj.1854589 *Perinatology*

Association of Maternal Creatinine-to-Body Weight and Urea-to-Creatinine Ratios with Gestational Diabetes Mellitus

Merve Ayas Özkan[1], Ruken Dayanan[1], Dilara Duygulu Bulan[1], Halis Doğukan Özkan[2], Gizem Aktemur[1], Gülşan Karabay[1], Betül Tokgöz Çakır[1], Zeynep Şeyhanlı[1], Ayşegül Atılgan Yıldırım[3], Furkan Akın[3], Zehra Vural Yılmaz[1]

[1]Department of Perinatology, Ankara Etlik City Hospital, Ankara, Türkiye; [2]Department of Obstetrics and Gynecology, Ankara Lösante Children and Adult Hospital, Ankara, Türkiye; [3]Department of Obstetrics and Gynecology, Ankara Etlik City Hospital, Ankara, Türkiye

ABSTRACT

Objectives: Gestational diabetes mellitus (GDM) is a multifactorial metabolic disorder associated with altered glucose metabolism and systemic physiological adaptations during pregnancy. Subtle changes in renal-related biochemical parameters may reflect underlying metabolic alterations accompanying GDM. This study aimed to evaluate the association between maternal creatinine-to-body weight (Cre/BW) and urea-to-creatinine (Urea/Cre) ratios and GDM, with a particular focus on their potential role as renal–metabolic indicators during pregnancy.

Methods: This retrospective case–control study included 1,064 pregnant women (532 with GDM and 532 matched controls) who underwent a 75 g oral glucose tolerance test between 24 and 28 weeks of gestation. Propensity score matching was applied to reduce baseline differences between groups. Maternal serum creatinine, urea, and body weight were recorded in the first and second trimesters, and Cre/BW and Urea/Cre ratios were calculated. Group comparisons, receiver operating characteristic (ROC) analyses, and multivariate logistic regression were performed.

Results: In the second trimester, the Cre/BW ratio was significantly lower, and the Urea/Cre ratio was significantly higher in women with GDM compared with controls (both P<0.001). The Cre/BW ratio demonstrated moderate discriminatory ability for distinguishing GDM status (Area under the curve [(AUC)]=0.774), whereas the Ure/Cre ratio showed modest performance (AUC=0.614). In multivariate analysis, second-trimester Cre/BW and Urea/Cre ratios remained independently associated with GDM after adjustment for clinically relevant covariates. These results suggest a potential association; however, the discriminatory performance of the Cre/BW and Urea/Cre ratios should be regarded as exploratory rather than diagnostically definitive.

Conclusions: Lower Cre/BW and higher Urea/Cre ratios were associated with GDM in this retrospective cohort. These findings suggest that simple renal-related biochemical ratios may reflect metabolic–renal adaptations in pregnancies complicated by GDM. However, given the modest diagnostic performance and retrospective design, these parameters should be considered exploratory indicators rather than standalone clinical tools. Further prospective studies are required to clarify their clinical relevance.

Keywords: Diabetes Mellitus, Pregnancy, Creatinine, Urea, Body Weight, Biomarkers

Received: January 5, 2026 **Accepted:** February 7, 2026 **Available Online:** February 11, 2026 **Published:** March 1, 2026

How to cite this article: Ayas Ozkan M, Dayanan R, Duygulu Bulan D, et al. Association of Maternal Creatinine-to-Body Weight and Urea-to-Creatinine Ratios with Gestational Diabetes Mellitus. Eur Res J. 2026;12(3):329-338. doi: 10.18621/eurj.1854589

Corresponding author: Merve Ayas Özkan, MD., Phone: +90 312 797 00 00, E-mail: merveayasozkan@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

Gestational diabetes mellitus (GDM) is defined as glucose intolerance first diagnosed during pregnancy [1, 2]. Its prevalence varies by region worldwide, averaging 10–18% [3, 4]. It is one of the most common complications of pregnancy, leading to maternal and fetal morbidity [3–5]. Insulin resistance and pancreatic beta cell dysfunction are the primary mechanisms in the pathophysiology of GDM. Hormonal changes, inflammation, genetic factors, oxidative stress, and alterations in lipid metabolism have been shown to contribute to this condition [6–10].

GDM can cause both short- and long-term changes in maternal kidney function. In normal pregnancy, renal plasma flow and glomerular filtration rate increase by approximately 40–50%. As a result, a physiological decrease in creatinine levels is observed during pregnancy [11, 12]. In pregnant women diagnosed with GDM, microvascular function, glomerular hyperfiltration, and renal hemodynamic adaptations may not occur fully. Even if creatinine levels are within the reference range, the expected decrease seen in normal pregnancy may not be evident in pregnancies complicated by GDM [13–15].

The creatinine-to-body weight ratio (Cre/BW) is used as a practical indicator of muscle mass. Muscle mass plays an important role in glucose metabolism, and low muscle mass has been associated with an increased risk of diabetes [16, 17]. Recent studies have shown that the Cre/BW ratio can predict metabolic diseases such as insulin resistance, diabetes mellitus, and liver diseases [16–18]. Similarly, increased urea levels and the urea/creatinine ratio (Urea/Cre) have also been reported to be associated with metabolic dysfunction and GDM [19].

The primary clinical goal in GDM management is to prevent maternal and neonatal complications such as preterm birth, fetal distress, and cesarean delivery. Therefore, identifying biochemical markers that not only predict GDM but may also be associated with adverse perinatal outcomes could be of clinical importance. Because the Cre/BW and Urea/Cre ratios reflect renal and metabolic function, they may also provide insight into maternal adaptation mechanisms linked to pregnancy complications.

Although physiological changes in renal function indicators during pregnancy are well defined, very few studies have directly evaluated the Cre/BW and Urea/Cre ratios in GDM [19, 20]. The oral glucose tolerance test (OGTT) remains the gold standard for diagnosing GDM, but its acceptability and tolerance among pregnant women have not been fully established. Therefore, there is a need for simple, inexpensive, and non-invasive biochemical markers that may provide supportive information in the evaluation of GDM, without replacing the OGTT. Considering the expected physiological decrease in urea and creatinine levels due to increased glomerular filtration rate (GFR) and hemodilution during pregnancy, relative increases in these parameters may be clinically significant [11].

In addition to identifying GDM itself, assessing biochemical markers that may also be linked to composite adverse perinatal outcomes (CAPO) could provide a more comprehensive understanding of the disease's clinical impact. Investigating whether renal function–related indices such as the Cre/BW and Urea/Cre ratios are associated with these complications may therefore have important implications for both diagnosis and prognosis. Accordingly, this study aims to examine trimester-specific changes in Cre/BW and Urea/Cre ratios in pregnant women and to investigate whether deviations in these ratios are associated with the development of GDM.

METHODS

This retrospective case–control study initially included 1,101 pregnant women who presented to Etlik City Hospital between September 2022 and September 2024. Propensity score matching was subsequently performed to create comparable groups, resulting in a final analytical sample of 1,064 pregnancies (532 women with GDM and 532 controls). Approval was obtained from the local ethics committee (Ethics Committee Decision No: AEŞH-BADEK-2024-848 date: 02/10/2024), and the research was conducted in accordance with the principles of the Helsinki Declaration. The medical records and laboratory data of all cases were reviewed retrospectively.

Pregnant women aged 18–45 years, with first trimester blood samples available at our hospital, and who underwent a 75 g OGTT between the 24th and 28th weeks of pregnancy were included in the study. Patients with additional systemic diseases (such as kidney disease, cardiac disease, or hypertension),

those previously diagnosed with type 1 or type 2 diabetes mellitus, those who developed any disease affecting kidney function during pregnancy, or those with multiple pregnancies were excluded. Among the cases meeting the inclusion and exclusion criteria, those diagnosed with GDM were classified as the "GDM group," while those with normal OGTT results were classified as the "control group."

GDM diagnosis was based on the criteria established by the International Association of Diabetes and Pregnancy Study Groups (IADPSG) [1]. Exceeding any of the following values in the 75 g OGTT was considered sufficient for a GDM diagnosis: fasting plasma glucose ≥92 mg/dL, 1-hour plasma glucose ≥180 mg/dL, or 2-hour plasma glucose ≥153 mg/dL.

First trimester blood samples consisted of serum values obtained from routine laboratory tests at the time of initial pregnancy diagnosis and body weight recorded on the same day. Second trimester blood samples consisted of blood taken at the time of the OGTT visit between 24 and 28 weeks of gestation and body weight measurements taken on that day. All blood samples in both trimesters were collected in the morning after an overnight fast. Cre/BW and Urea/Cre ratios were calculated using serum urea and creatinine levels for both periods.

Cases and controls were selected consecutively from the same hospital database during the study period (September 2022–September 2024). Each woman diagnosed with GDM according to the IADPSG criteria was paired with the next eligible pregnant woman who had a normal 75 g OGTT result according to IADPSG cut-offs and met the inclusion criteria.

CAPO was defined as the presence of at least one of the following: preterm birth (<37 weeks of gestation), fetal distress at birth, requirement for admission to the neonatal intensive care unit (NICU), or intrauterine death.

Age, gravida, parity, body mass index (BMI), gestational age, laboratory parameters (urea, creatinine, hemoglobin A1c [HbA1c]), and birth outcomes of the cases included in the study were recorded through the hospital information system.

Propensity Score Matching

To minimize baseline confounding and improve comparability between groups, propensity score matching was performed using a 1:1 nearest-neighbor algorithm without replacement and with a caliper width of 0.20 of the standard deviation of the logit of the propensity score. The propensity score was estimated through a logistic regression model including clinically relevant baseline covariates: maternal age, gravida, parity, pre-pregnancy BMI, first-trimester body weight, and family history of diabetes mellitus. After matching, 532 pregnancies with GDM and 532 matched controls constituted the final analytic cohort. All subsequent analyses - including univariate comparisons, ROC analyses, and multivariable logistic regression - were conducted exclusively within this matched sample.

Statistical Analysis

Statistical analyses were performed using SPSS Statistics for Windows, Version 27.0 (IBM Corp., Armonk, NY, USA). The distribution of continuous variables was assessed using the Shapiro–Wilk test. Variables with a normal distribution were presented as mean ± standard deviation, while those without a normal distribution were presented as median (interquartile range, IQR). Categorical variables were expressed as number (n) and percentage (%). For group comparisons, the independent samples t-test was used for normally distributed variables, and the Mann–Whitney U test was used for non-normally distributed variables. For categorical variables, the Pearson Chi-square test or, when appropriate, Fisher's exact test was applied. Receiver operating characteristic (ROC) curve analysis was performed to evaluate the diagnostic performance of Cre/BW and Urea/Cre ratios in the second trimester for predicting GDM diagnosis. Area under the curve (AUC) values were calculated from the ROC curve, and optimal cut-off values, sensitivity, and specificity were determined for statistically significant variables. In addition to univariate comparisons, a multivariate binary logistic regression analysis was performed to identify independent predictors of GDM. Variables that were considered clinically relevant (maternal age, BMI, fasting plasma glucose, Cre/BW ratio, Urea/Cre ratio, and family history of diabetes mellitus) were entered into the model using the enter method. Adjusted odds ratios (aOR) with 95% confidence intervals (CI) were calculated. Statistical significance was set at $P<0.05$.

TABLE 1. Comparison of Demographic and Clinical Characteristics of the Case and Control Groups.

	GDM group n=532	Control group n=532	P-value
Maternal age (year)	30.8±6.2	30.3±6.6	0.452[a]
Gravidty	2 (2)	2 (2)	0.911[b]
Parity	1 (2)	1 (2)	0.439[b]
BMI, prepregnancy (kg/m^2)	28.06±5.32	27.82±5.10	0.703[b]
First trimester weight (kg)	68.8±13.8	67.4±13.2	0.526[a]
Second trimester weight (kg)	77.3±14.0	72.3±13.9	**<0.001**[a]
Family history of DM	6 (1.12 %)	5 (0.93 %)	0.760[b]

Data are shown as mean ± standard deviation, median (interquartile range), or n (%) where appropriate. BMI, body mass index; DM, diabetes mellitus; GDM, gestational diabetes mellitus.

[a]Independent samples *t*-test, [b]Mann–Whitney *U* test or Chi-square test, as appropriate.

Statistically significant P-value is shown in bold.

RESULTS

The demographic and clinical characteristics of the study population are shown in Table 1. Maternal age, gravida, parity, pre-pregnancy BMI, and first-trimester body weight were similar between the two groups (all P>0.05). Body weight in the second trimester was significantly higher in women with GDM compared with controls (77.3±14.0 kg vs. 72.3±13.9 kg, P<0.001). Although the frequency of a family history of diabetes mellitus did not differ significantly between groups in univariate analysis, it was further evaluated in multivariate analysis.

Biochemical parameters measured in the first and

TABLE 2. Comparison of Biochemical Parameters Between the GDM and Control Groups

	GDM group n=532	Control group n=532	P-value
First trimester urea (mg/dL)	16.50 (6.92)	13.60 (3.60)	**<0.001**
Second trimester urea (mg/dL)	13.95 (5.42)	11.30 (5.50)	**<0.001**
First trimester Cre (mg/dL)	0.54 (0.12)	0.50 (0.14)	**<0.001**
Second trimester Cre (mg/dL)	0.47 (0.12)	0.46 (0.13)	**0.028**
HbA1C	5.1 (0.6)	-	-
Fasting blood glucose (mg/dL) (first trimester)	88 (58)	88 (17)	**<0.001**
75 g OGTT fasting plasma glucose	88.5 (22)	77 (11)	**<0.001**
75 g OGTT 1 hour plasma glucose	194.5 (27.3)	137.0 (25.0)	**<0.001**
75 g OGTT 2 hour plasma glucose	158 (52)	101 (11)	**<0.001**
First trimester Cre/BW	0.00743 (0.00239)	0.00769 (0.00295)	0.214
Firs trimester Urea/Cre	31.08 (10.81)	29.79 (12.26)	0.467
Second trimester Cre/BW	0.00600 (0.00158)	0.00758 (0.00251)	**<0.001**
Second trimester Urea/Cre	28.75 (11.71)	25.0 (8.78)	**<0.001**

Data are shown as median (interquartile range) unless otherwise indicated. GDM, gestational diabetes mellitus; OGTT, oral glucose tolerance test; HbA1c, hemoglobin A1c; BMI, body mass index; Cre, creatinine; BW, body weight; Urea/Cre, urea-to-creatinine ratio.

All comparisons were performed using the Mann–Whitney U test. Statistically significant P-values are shown in bold.

TABLE 3. Comparison of Obstetric and Neonatal Outcomes Between the GDM and Control Groups

	GDM group n=532	Control Group n=532	P-value
Gestational week at birth	38.0 (2.0)	39.0 (2.0)	**<0.001**[a]
Birth weight, grams	3240.0 (540.0)	3220 (494.0)	0.428[a]
Mode of delivery, n (%)			**<0.001**[b]
Vaginal delivery	208 (38.0%)	269 (48.4%)	
Primary cesarean section	141 (25.8%)	124 (22.3%)	
Repeat cesarean section	197 (36.0%)	162 (29.1%)	
1-min APGAR score	9 (1)	9 (1)	0.320[a]
5-min APGAR score	10 (1)	10 (1)	0.326[a]
Fetal compromise, n (%)	39 (7.1%)	37 (6.7%)	0.812[b]
Preterm birth, n (%)	87 (15.9%)	49 (8.8%)	**<0.001**[b]
NICU admission, n (%)	68 (12.5%)	49 (8.8%)	0.063[b]
Composite adverse outcomes, n (%)	133 (24.4%)	68 (12.3%)	**<0.001**[b]

Data are shown as median (interquartile range) or n (%) where appropriate. GDM, gestational diabetes mellitus, NICU, neonatal intensive care unit.

[a]Mann–Whitney *U* test, [b]Pearson chi-square test. Statistically significant P-values are shown in bold.

second trimesters are presented in Table 2. Urea levels were higher in women with GDM in both trimesters (P<0.001 for both). First-trimester median serum creatinine was also higher in the GDM group (0.54 [IQR: 0.12] vs. 0.50 [0.14] mg/dL, P<0.001), and a smaller but statistically significant difference persisted in the second trimester (P=0.028). Cre/BW was comparable between the groups in early pregnancy (P=0.214) but was significantly lower in the GDM group during the second trimester (0.00600 [IQR: 0.00158] vs. 0.00758 [0.00251], P<0.001). In contrast, Urea/Cre was similar in the first trimester (P=0.467) but significantly higher among women with GDM in the second trimester (28.75 [IQR: 11.71] vs. 25.0 [8.78], P<0.001). All glucose values obtained during the 75-g OGTT (fasting, 1-hour, and 2-hour) were significantly higher in the GDM group (all P<0.001).

Obstetric and neonatal outcomes are summarized in Table 3. Gestational age at birth was lower in the GDM group (38.0 [IQR: 2.0] vs. 39.0 [2.0] weeks, P<0.001), while birth weight was similar between groups (P=0.428). Cesarean delivery was more frequent among women with GDM (P<0.001). Preterm birth occurred more often in the GDM group (15.9% vs. 8.8%, P<0.001). Rates of fetal compromise and NICU admission were similar between the groups (P=0.063). The rate of composite adverse perinatal outcomes was significantly higher in women with GDM (24.6% vs. 12.2%, P<0.001).

The diagnostic performance of the Cre/BW and Urea/Cre ratios in the second trimester is summarized in Table 4. The Cre/BW ratio demonstrated moderate

TABLE 4. ROC Analysis for the Determination of Optimal Cutoff Values for Second Trimester Cre/BW and Urea/Cre in the Prediction of GDM

	Cut-off*	Sensitivity	Specificity	AUC	%95 CI	P-value
Second trimester Cre/BW	<0.00645	75.5%	62.6%	0.774	0.743-0.801	**<0.001**
Second trimester Ure/Cre	>27.03	60.1%	61.6%	0.614	0.578-0.650	**<0.001**

Cut-off values were determined using receiver operating characteristic (ROC) curve analysis. GDM, gestational diabetes mellitus; Cre, creatinine, BW, body weight, Urea/Cre, urea-to-creatinine ratio, AUC, Area under the curve; CI, Confidence interval. Statistically significant values are shown in bold.

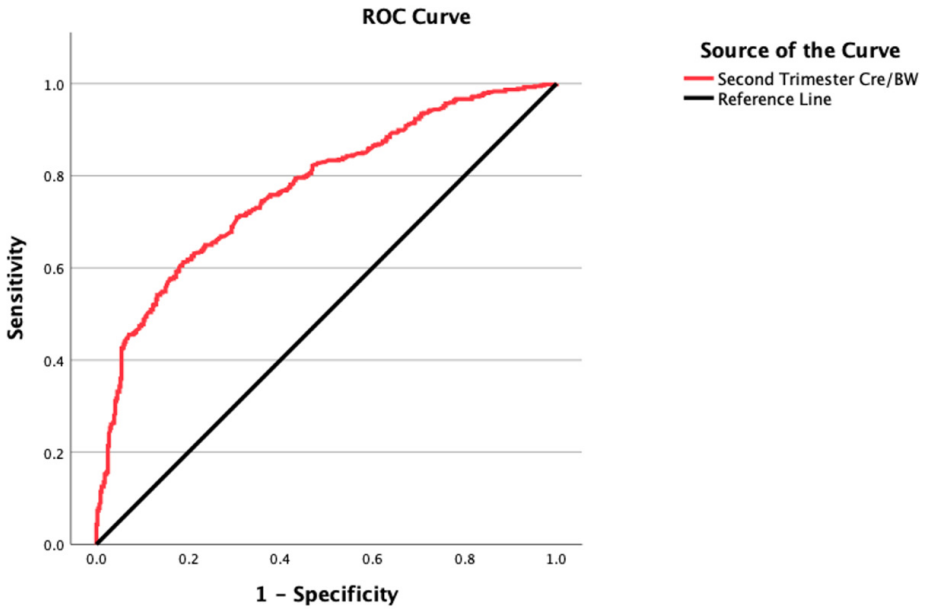

FIGURE 1. ROC curves of second trimester Cre/BW in predicting GDM. GDM, gestational diabetes mellitus; Cre/BW, creatinine-to-body weight; ROC, receiver operating characteristic.

discriminatory ability for identifying GDM, with an AUC of 0.774 (95% CI: 0.743–0.801) (Figure 1). The optimal cutoff value (<0.00645) yielded a sensitivity of 75.5% and a specificity of 62.6%. The Urea/Cre ratio showed more modest predictive value, with an AUC of 0.614 (95% CI: 0.578–0.650) (Figure 2). The optimal cutoff (>27.03) provided a sensitivity of 60.1% and a specificity of 61.6%.

Multivariate logistic regression analysis identified fasting plasma glucose, second-trimester Cre/BW, and second-trimester Urea/Cre as independent predictors of GDM (Table 5). Lower Cre/BW was strongly associated with increased odds of GDM (adjusted OR 0.68, 95% CI 0.55–0.84, P<0.001). Higher Urea/Cre ratio showed a modest but statistically significant association (adjusted OR 1.04, 95% CI 1.01–1.07,

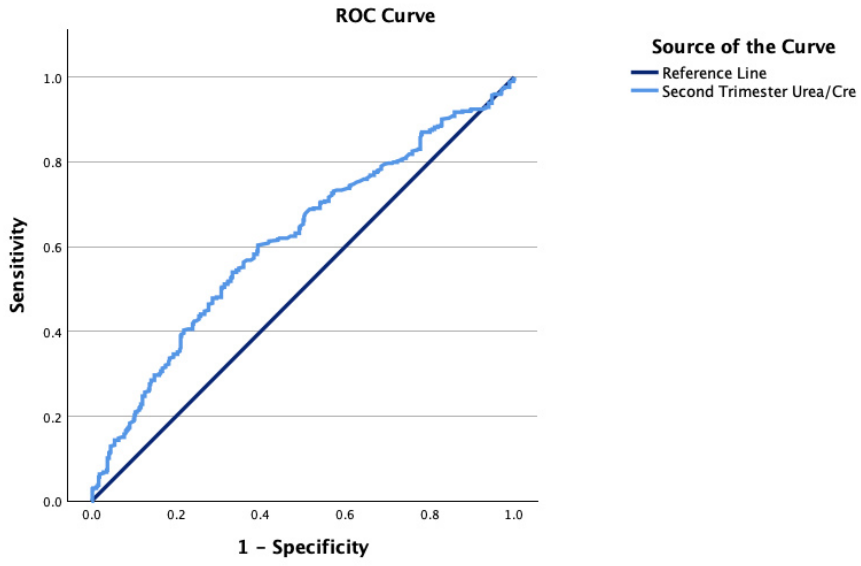

FIGURE 2. ROC curve of second trimester Ure/Cre in predicting GDM. GDM, gestational diabetes mellitus; Urea/Cre, urea-to-creatinine; ROC, receiver operating characteristic.

TABLE 5. Multivariate Logistic Regression Analysis for Independent Predictors of Gestational Diabetes Mellitus

Variable	aOR (95% CI)	P-value
Age (years)	1.01 (0.99–1.04)	0.121
BMI (kg/m²)	1.02 (1.00–1.05)	0.058
Fasting glucose (mg/dL)	1.05 (1.03–1.09)	**<0.001**
Cre/BW	0.64 (0.52–0.78)	**<0.001**
Urea/Cre	1.04 (1.00–1.08)	**0.041**
Family history of DM (yes/no)	4.35 (1.62–11.65)	**0.003**

Multivariate binary logistic regression model including maternal age, body mass index (BMI), fasting plasma glucose, creatinine-to-body weight (Cre/BW) ratio, urea-to-creatinine (Urea/Cre) ratio, and family history of diabetes mellitus (yes/no) for the prediction of gestational diabetes mellitus (GDM). Adjusted odds ratios (aOR) with 95% confidence intervals (CI) are presented.

Statistically significant P-values are shown in bold.

P=0.045). Fasting glucose remained the strongest predictor (adjusted OR 1.05, 95% CI 1.03–1.08, P<0.001). Maternal age and BMI were not independently associated with GDM (P>0.05 for both), whereas a family history of diabetes mellitus emerged as an independent predictor in the multivariate model (adjusted OR 4.35, 95% CI 1.62–11.65, P=0.003).

DISCUSSION

The observed alterations in renal-related biochemical ratios in pregnancies complicated by GDM may reflect underlying metabolic–renal adaptations rather than isolated laboratory changes. The Cre/BW ratio, which has been proposed as a surrogate marker of muscle mass and metabolic health, demonstrated a stronger association with GDM than the Urea/Cre ratio. This observation is consistent with previous studies reporting inverse relationships between creatinine-based indices and metabolic disorders, including diabetes mellitus [16–18]. Because body weight in the second trimester was higher among women with GDM, gestational weight gain may have partially influenced the Cre/BW ratio. Although this ratio accounts for body weight, dynamic changes in maternal weight during pregnancy may affect its interpretation. Therefore, the Cre/BW ratio should be considered a composite indicator influenced by both metabolic status and gestational factors.

Renal hemodynamic changes and metabolic stress may both contribute to the observed alterations in these ratios. In diabetes mellitus, hyperglycemia can induce glomerular hyperfiltration during early disease stages [21, 22]. However, GDM is also associated with endothelial dysfunction and microvascular alterations that may interfere with normal renal adaptation during pregnancy [23]. Previous studies have reported slightly higher serum creatinine levels in pregnant women with GDM compared with healthy controls [24] supporting the concept that renal physiological responses may differ in metabolically complicated pregnancies. In addition, a large cohort study demonstrated that higher creatinine and Cre/BW ratios in early pregnancy were associated with a lower risk of GDM, while altered renal-related indices were linked to increased metabolic risk [20].

In multivariate analysis, second-trimester Cre/BW and Ure/Cre ratios remained independently associated with GDM after adjustment for clinically relevant covariates. The inverse association between the Cre/BW ratio and GDM supports the hypothesis that reduced muscle mass or altered body composition may contribute to impaired glucose utilization and increased insulin resistance during pregnancy. Skeletal muscle represents the primary site of insulin-mediated glucose uptake, and lower muscle mass has been consistently associated with metabolic dysfunction in both pregnant and non-pregnant populations [16, 25, 26]. The modest but statistically significant association observed for the Urea/Cre ratio may reflect subtle alterations in renal handling of nitrogenous waste products in the setting of metabolic stress, as previously suggested by Chong *et al*. [19].

Despite statistically significant group differences, the diagnostic performance of these ratios requires cautious interpretation. Although the discriminatory performance was moderate, the association between the Cre/BW ratio and GDM supports the hypothesis that this index may capture subtle metabolic alterations during pregnancy. Importantly, these biochemical indices alone are unlikely to provide sufficient accuracy for clinical screening or diagnostic

decision-making and should not be considered replacements for established screening strategies or existing clinical and biochemical predictors of GDM.

The associations observed between these renal-related ratios and adverse perinatal outcomes should also be interpreted with caution. Although higher rates of composite adverse perinatal outcomes were noted among women with GDM, these findings were descriptive in nature, and causal inferences cannot be drawn due to the retrospective study design. Potential confounding factors such as hydration status, gestational weight gain patterns, renal hemodynamics, and precise assessment of muscle mass could not be fully accounted for and may have influenced the observed biochemical variations [27].

Strengths and Limitations

Several limitations should be acknowledged. The retrospective case–control design limits causal interpretation and complete control of confounding variables, despite the use of propensity score matching and multivariable adjustment. Serum creatinine was used as an indirect marker of muscle mass, which may not accurately reflect dynamic changes in body composition during pregnancy. Additionally, estimated glomerular filtration rate was not calculated, which could have provided further insight into renal functional adaptations. The single-center nature of the study may also limit generalizability. Nevertheless, the relatively large, well-matched cohort and standardized timing of biochemical measurements across trimesters strengthen the internal consistency of the observed associations.

CONCLUSION

In summary, lower creatinine-to-body weight and higher urea-to-creatinine ratios were associated with gestational diabetes mellitus in this retrospective cohort. These renal-related biochemical indices may reflect underlying metabolic–renal adaptations in pregnancies complicated by GDM. However, given their modest diagnostic performance and the limitations inherent to the study design, these ratios should be regarded as exploratory indicators rather than clinically actionable tools. Prospective, multicenter studies incorporating comprehensive assessment of renal function and body composition are needed to further elucidate their potential role in the context of GDM.

Ethics Approval and Consent to Participate

This study was approved by the Ankara Etlik City Hospital Scientific Research Evaluation and Ethics Committee (Decision No: AEŞH-BADEK-2024-848; date: 02.10.2024. All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Informed consent was not required in this study because this is a retrospective study.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: MAÖ, RD, DDB, HDÖ, GA, GK, BTÇ, ZŞ, FA, ZVY; Study Design: MAÖ, RD, DDB, HDÖ, GA, GK, BTÇ, ZŞ, FA, ZVY; Supervision: ZVY; Funding: N/A; Materials: MAÖ, RD, DDB, HDÖ, AAY; Data Collection and/or Processing: MAÖ, RD, DDB, HDÖ, AAY; Statistical Analysis and/or Data Interpretation: MAÖ, RD, DDB, HDÖ, BTÇ, ZŞ; Literature Review: MAÖ, HDÖ, BTÇ, ZŞ; Manuscript Preparation: MAÖ; and Critical Review: MAÖ.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial

intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. International Association of Diabetes and Pregnancy Study Groups Consensus Panel; Metzger BE, Gabbe SG, Persson B, et al. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33(3):676-682. doi: 10.2337/dc09-1848.
2. Hod M, Kapur A, Sacks DA, et al. The International Federation of Gynecology and Obstetrics (FIGO) Initiative on gestational diabetes mellitus: A pragmatic guide for diagnosis, management, and care. Int J Gynaecol Obstet. 2015;131 Suppl 3:S173-211. doi: 10.1016/S0020-7292(15)30033-3.
3. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. Nat Rev Dis Primers. 2019;5(1):47. doi: 10.1038/s41572-019-0098-8.
4. Sweeting A, Wong J, Murphy HR, Ross GP. A Clinical Update on Gestational Diabetes Mellitus. Endocr Rev. 2022;43(5):763-793. doi: 10.1210/endrev/bnac003.
5. Moon JH, Jang HC. Gestational Diabetes Mellitus: Diagnostic Approaches and Maternal-Offspring Complications. Diabetes Metab J. 2022;46(1):3-14. doi: 10.4093/dmj.2021.0335.
6. Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The Pathophysiology of Gestational Diabetes Mellitus. Int J Mol Sci. 2018;19(11):3342. doi: 10.3390/ijms19113342.
7. Modzelewski R, Stefanowicz-Rutkowska MM, Matuszewski W, Bandurska-Stankiewicz EM. Gestational Diabetes Mellitus-Recent Literature Review. J Clin Med. 2022;11(19):5736. doi: 10.3390/jcm11195736.
8. Filardi T, Catanzaro G, Mardente S, et al. Non-Coding RNA: Role in Gestational Diabetes Pathophysiology and Complications. Int J Mol Sci. 2020;21(11):4020. doi: 10.3390/ijms21114020.
9. Kopylov AT, Papysheva O, Gribova I, et al. Molecular pathophysiology of diabetes mellitus during pregnancy with antenatal complications. Sci Rep. 2020;10(1):19641. doi: 10.1038/s41598-020-76689-9.
10. Hadi SH, Hassan A-SU, Abdalredha RD. Interleukin 6 as a key inflammatory predictor of gestational diabetes – clinical and biochemical evidence. Eur J Clin Exp Med. 2025;23(3):633-640 2025. doi: 10.15584/ejcem.2025.3.17.
11. Cheung KL, Lafayette RA. Renal physiology of pregnancy. Adv Chronic Kidney Dis. 2013;20(3):209-214. doi: 10.1053/j.ackd.2013.01.012.
12. Wiles K, Bramham K, Seed PT, Nelson-Piercy C, Lightstone L, Chappell LC. Serum Creatinine in Pregnancy: A Systematic Review. Kidney Int Rep. 2018;4(3):408-419. doi: 10.1016/j.ekir.2018.10.015.
13. Duo Y, Song S, Zhang Y, et al. Relationship between serum uric acid in early pregnancy and gestational diabetes mellitus: a prospective cohort study. Endocrine. 2024;83(3):636-647. doi: 10.1007/s12020-023-03544-y.
14. Song S, Duo Y, Zhang Y, et al. Serum Uric Acid-to-Serum Creatinine Ratio, Thyroid Function, and Gestational Diabetes Mellitus Risk during Early Pregnancy among Chinese Women. Diabetes Metab Syndr Obes. 2025;18:2235-2246. doi: 10.2147/DMSO.S486695.
15. Güngör ES, Danişman N, Mollamahmutoğlu L. Relationship between serum uric acid, creatinine, albumin and gestational diabetes mellitus. Clin Chem Lab Med. 2006;44(8):974-977. doi: 10.1515/CCLM.2006.173.
16. He J. Creatinine-to-body weight ratio is a predictor of incident diabetes: a population-based retrospective cohort study. Diabetol Metab Syndr. 2022;14(1):7. doi: 10.1186/s13098-021-00776-8.
17. Abdullah N, Goh YX, Khuzaimi AN, et al. Linking Creatinine-to-Body Weight Ratio With Diabetes Incidence: A Multiethnic Malaysian Cohort Study. J Diabetes. 2025;17(1):e70039. doi: 10.1111/1753-0407.70039.
18. Lin J, Zheng J, Lin X, Chen Y, Li Z. A Low Creatinine to Body Weight Ratio Predicts the Incident Nonalcoholic Fatty Liver Disease in Nonelderly Chinese without Obesity and Dyslipidemia: A Retrospective Study. Gastroenterol Res Pract. 2020;2020:4043871. doi: 10.1155/2020/4043871.
19. Chong H, Li J, Chen C, Wang W, Liao D, Zhang K. The diagnostic model for early detection of gestational diabetes mellitus and gestational diabetic nephropathy. J Clin Lab Anal. 2022;36(9):e24627. doi: 10.1002/jcla.24627.
20. Yang X, Peng J, Sun F, et al. Renal Function and Risk of Gestational Diabetes Mellitus: The Potential Mediating Role of Carnitine Metabolites. J Clin Endocrinol Metab. 2025;110(9):e3052-e3062. doi: 10.1210/clinem/dgae881.
21. Zhang J, Wei J, Jiang S, et al. Macula Densa SGLT1-NOS1-Tubuloglomerular Feedback Pathway, a New Mechanism for Glomerular Hyperfiltration during Hyperglycemia. J Am Soc Nephrol. 2019;30(4):578-593. doi: 10.1681/ASN.2018080844.
22. Vallon V, Thomson SC. The tubular hypothesis of nephron filtration and diabetic kidney disease. Nat Rev Nephrol. 2020;16(6):317-336. doi: 10.1038/s41581-020-0256-y.
23. Fato BR, Beard S, Binder NK, et al. The Regulation of Endothelin-1 in Pregnancies Complicated by Gestational Diabetes: Uncovering the Vascular Effects of Insulin. Biomedicines. 2023;11(10):2660. doi: 10.3390/biomedicines11102660.
24. Khan R, Khan Z, Javed K, Ali K. Effect of gestational diabetes on blood sugar, liver and renal function tests. J Ayub Med Coll Abbottabad. 2012;24(2):95-98.
25. Harita N, Hayashi T, Sato KK, et al. Lower serum creatinine

is a new risk factor of type 2 diabetes: the Kansai healthcare study. Diabetes Care. 2009;32(3):424-426. doi: 10.2337/dc08-1265.
26. Hashimoto Y, Okamura T, Hamaguchi M, Obora A, Kojima T, Fukui M. Creatinine to Body Weight Ratio Is Associated with Incident Diabetes: Population-Based Cohort Study. J Clin Med. 2020;9(1):227. doi: 10.3390/jcm9010227.
27. Zhang B, Zhan Z, Xi S, Zhang Y, Yuan X. Association Between Maternal Creatinine to Body Weight Ratio and Small/Large for Gestational Age Newborns Among 11,734 Chinese Women. J Epidemiol Glob Health. 2025;15(1):70. doi: 10.1007/s44197-025-00414-w.

The European Research Journal 2026;12(3):339-346 Original Article

DOI: https://doi.org/10.18621/eurj.1861383 *Cardiovascular Surgery*

Association Between Cardiopulmonary Bypass-Induced Sirtuin-1 Suppression and Apoptosis

Bişar Amaç[1], Ömer Göç[1], Mesut Engin[2], Şenol Yavuz[2]

[1]Department of Perfusion, Faculty of Health Sciences, Harran University, Şanlıurfa, Türkiye; [2]Department of Cardiovascular Surgery, Bursa Yuksek Ihtisas Training and Research Hospital, University of Health Sciences, Bursa, Türkiye

ABSTRACT

Objectives: The aim of this study is to investigate, at the molecular level, the effect of cardiopulmonary bypass (CPB) on Sirtuin-1 levels, an epigenetic-metabolic regulatory protein, and the relationship between this change and Caspase-3, one of the key executor molecules of cellular apoptosis.

Methods: Patients undergoing open heart surgery under CPB were included in the study. Sirtuin-1 and Caspase-3 levels were measured in venous blood samples taken from patients during the preoperative, intraoperative, and postoperative periods. The Wilcoxon Signed Ranks test was used to analyze differences between time points.

Results: Sirtuin-1 levels decreased significantly during the intraoperative period compared to preoperative values ($Z= -6.212$, $P<0.001$). Similarly, postoperative Sirtuin-1 levels were significantly lower compared to both the preoperative ($Z= -6.229$, $P<0.001$) and intraoperative periods ($Z= -6.186$, $P<0.001$). Caspase-3 levels showed a significant increase in the intraoperative period compared to the preoperative period ($Z= −6.262$, $P<0.001$). Postoperative Caspase-3 levels were significantly higher than both the preoperative ($Z= -6.196$, $P<0.001$) and intraoperative periods ($Z= -6.203$, $p < 0.001$).

Conclusions: These findings indicate that CPB suppresses circulating Sirtuin-1 levels and that this suppression is associated with increased apoptosis. CPB-induced Sirtuin-1 inhibition may be one of the molecular mechanisms of surgical-related cellular damage. Sirtuin-1 may be considered a potential molecular target for preventing CPB-related myocardial and systemic damage.

Keywords: Cardiopulmonary Bypass, Sirtuin-1, Caspase-3, Apoptosis, Molecular Perfusion

Cardiopulmonary bypass (CPB) is a fundamental life-supporting technology in open-heart surgery; however, oxidative stress, inflammation, and energy imbalance that develop during this process cause damage at the cellular level [1]. Silent Information Regulator 2 Homolog 1 (Sirtuin-1 / SIRT1), an important regulator of intracellular energy metabolism, is notable for its functions in reducing oxidative stress, protecting mitochondrial function, and slowing cellular aging [2]. Changes in Sirtuin-1 levels during CPB are thought to be decisive for postoperative organ function and the recovery process. Therefore, examining the biochemical changes exhibited by Sirtuin-1 in response to CPB will contribute to redefining perfusion science not only in terms of systemic circulation but also in terms of maintaining cellular homeostasis at the molecular level [3].

Received: January 11, 2026 **Accepted:** February 11, 2026 **Available Online:** February 11, 2026 **Published:** March 1, 2026

How to cite this article: Amaç B, Göç Ö, Engin M, Yavuz S. Association Between Cardiopulmonary Bypass–Induced Sirtuin-1 Suppression and Apoptosis. Eur Res J. 2026;12(3):339-346. doi: 10.18621/eurj.1861383

Corresponding author: Bişar Amaç, PhD., Assist. Prof., Phone: +90 414 318 30 00, E-mail: amacbisar@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

CPR, while an indispensable component of open-heart surgery, triggers complex biological processes associated with systemic inflammatory response, oxidative stress, and cellular damage. During CPB, factors such as contact of blood with non-physiological surfaces, surgical trauma, hypothermia, and ischemia-reperfusion injury trigger complement activation and pro-inflammatory cytokine release, thereby increasing the systemic inflammatory response [4, 5]. The same conditions contribute to an increase in reactive oxygen species production, leading to insufficient endogenous antioxidant capacity and elevated oxidative stress; this condition is associated with postoperative organ dysfunction. [6]. This pro-inflammatory and oxidative environment activates cellular stress responses, paving the way for the triggering of apoptotic pathways, particularly in myocardial and other organ tissues [7]. These mechanistic findings demonstrate that CPB not only achieves hemodynamic goals but also affects the balance between life and death at the molecular level.

Apoptosis is a programmed cell death mechanism that plays a fundamental role in maintaining cellular homeostasis and is closely associated with myocardial dysfunction, acute organ damage, and postoperative complications following cardiac surgery. Caspase-3, which plays a central role in the execution of the apoptotic process, is the common final effector of both intrinsic and extrinsic apoptotic pathways, and its increased circulating levels are considered a biochemical marker of systemic cellular damage. Inflammation and oxidative stress associated with CPB have been reported to trigger apoptosis by increasing caspase-3 activation; however, the higher-level regulatory mechanisms of these processes have not yet been fully elucidated [7-9].

In recent years, Sirtuin-1 has attracted attention as a critical molecular sensor in the regulation of the cellular stress response. Sirtuin-1, a nicotinamide adenine dinucleotide (oxidized form) (NAD^+)-dependent deacetylase, has regulatory effects on mitochondrial function, inflammation, oxidative stress response, and cellular life–death balance. Experimental and clinical studies have shown that Sirtuin-1 exhibits anti-inflammatory and anti-apoptotic properties, particularly supporting cell survival under oxidative stress conditions. Decreased Sirtuin-1 activity has been associated with increased inflammation and apoptosis [10-13].

It has been suggested that systemic inflammation and redox imbalance developing during CPB may have an inhibitory effect on Sirtuin-1 expression and activity. However, in humans, changes in circulating Sirtuin-1 levels, particularly during the perioperative period, and their relationship with apoptotic markers have been evaluated in a limited number of studies. In this context, elucidating the effect of CPB on Sirtuin-1 and its possible relationship with the apoptotic response may contribute to a better understanding of the molecular mechanisms of CPB-related cellular damage [14, 15].

Caspase-3, which plays a central role in the execution phase of apoptosis, is one of the primary effector caspases activated under cellular stress conditions. Increased pro-apoptotic signals resulting from stimuli such as inflammation, oxidative stress, and ischemia-reperfusion injury lead to Caspase-3 activation, causing the cellular death process to enter its irreversible phase. It has been reported that the systemic inflammatory response and oxidative stress environment that arise during CPB trigger Caspase-3-mediated apoptotic pathways and that this contributes to myocardial and extracardiac cellular damage. In this context, Caspase-3 is considered an important apoptotic marker in the assessment of cellular damage associated with CPB [7, 9, 16, 17].

The aim of this study is to evaluate changes in circulating Sirtuin-1 and caspase-3 levels in patients undergoing CPB during the preoperative, intraoperative, and postoperative periods using the ELISA method, and to investigate the relationship between CPB-induced Sirtuin-1 suppression and increased apoptosis. This study aims to shed light on the molecular basis of cellular stress and apoptotic response associated with CPB, thereby contributing to the development of targeted protective strategies in the future.

METHODS

In this prospective clinical and translational molecular medicine study, changes in circulating Sirtuin-1 levels during the perioperative period in patients undergoing

CPB and the relationship between these changes and the caspase-3-mediated apoptotic response were evaluated using the ELISA method.

Ethical Dimension of the Research

Approval for this prospective study was obtained from the institutional administration and the Harran University Clinical Research Ethics Committee (Date: 17.11.2025; Approval No: HRÜ/25.18.33). The study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Prior to enrollment, written informed consent was obtained from all participants after providing detailed information regarding the purpose, procedures, and potential risks of the study. Blood samples were collected using existing vascular access without any additional invasive intervention, and patient confidentiality was strictly maintained throughout the study. All data were anonymized prior to analysis.

Data Collection Method

A total of 50 consecutive patients undergoing cardiac surgery with CPB and without planned additional surgical procedures were prospectively included in the study. The demographic and clinical data of the patients included in the study were systematically recorded [age, gender, height, weight, body surface area (BSA), flow, aortic cross-clamp time, total perfusion time, and type of surgery performed (number of coronary artery bypass grafts)]. Venous blood samples were collected from patients at three different time points: preoperatively before anesthesia induction, intraoperatively following aortic cross-clamping, and during the weaning phase from CPB, each sample being 10 mL. Blood samples were obtained using existing central venous catheters and the heart-lung machine manifold; no additional invasive procedures were performed on the patients. Circulating Sirtuin-1 and caspase-3 protein levels were analyzed from the collected blood samples using the ELISA method, and the obtained data were evaluated using appropriate statistical methods.

ELISA Analysis

Venous blood samples were centrifuged at 3000 rpm for 10 minutes at 4 °C after clotting to separate the serum. The serum samples obtained were stored at –80 °C until analysis. Samples were thawed only once prior to analysis and were not refrozen. Circulating Sirtuin-1 and caspase-3 protein levels were measured using a human serum-specific ELISA method according to the manufacturer's instructions. The minimum detectable dose of Sirtuin-1 in human serum was 0.28 ng/mL, with a typical range of 0.78–50 ng/mL.

Inclusion and Exclusion Criteria

Inclusion Criteria: Patients aged 20–85 years who were scheduled to undergo elective cardiac surgery with CPB and who provided informed consent following ethical committee approval were included in this study. All patients underwent surgery using the standard CPB protocol without additional surgical procedures.

Exclusion Criteria: Patients who received amiodarone therapy in the preoperative period, those with a history of atrial fibrillation, patients undergoing emergency cardiac surgery, and patients scheduled for additional major cardiac surgery such as aortic aneurysm or aortic dissection were excluded from the study. Additionally, patients with active infection, chronic or active respiratory disease such as pneumonia or chronic obstructive pulmonary disease (COPD), known systemic inflammatory or autoimmune disease, active malignancy, chronic corticosteroid or immunosuppressive therapy, previous cardiac surgery (redo) patients, individuals undergoing chronic hemodialysis treatment, those with hematological diseases, and patients with severe liver dysfunction were excluded from the study.

Statistical Analysis

Patient data collected within the scope of the study were analyzed using the IBM Statistical Package for the Social Sciences 25 (IBM SPSS Statistics 25®) software package (IBM Corporation, Armonk, NY, USA). The sample size was determined based on previous similar studies and power analysis. A minimum power of 80% and a significance level of 0.05 were considered sufficient to detect statistically meaningful differences between groups. Means and standard deviations were calculated for continuous and ordinal data. The Kolmogorov Smirnov test and Shapiro-Wilk test were used to assess normality of distribution (the Kolmogorov–Smirnov test was used

because the number of patients was over 30). Comparemeans (Paired-sample T Test) and Nonparametric test (2Related Samples>Willcoxon) tests were used to evaluate normal and non-normally distributed data, respectively, for the comparison of the same parameters before, during, and after surgery. Frequency and percentage analyses were performed for nominal data. A P value of less than 0.05 was considered statistically significant.

RESULTS

A total of 50 patients were included in the study. Forty-four percent (n=22) of patients were female, and 56% (n=28) were male. When surgical procedures were evaluated, 10% (n=5) of patients underwent coronary artery bypass grafting (CABG)×1, 16% (n=8)underwent CABG×2, 40% (n=20)underwent CABG×3, 24% (n=12) underwent CABG×4, and 10% (n=5) underwent CABG×5 surgery. The number of cases valid for all variables was determined as 50 (Table 1).

The ages of the patients included in the study ranged from 42 to 76 years, with an average age of 61.18±7.23 years. The patients' height ranged from 1.50 to 1.85 m, with an average of 1.66±0.08 m. Body weight ranged from 55 to 103 kg, and the average value was calculated as 78.62±13.39 kg. Body surface area (BSA) ranged from 1.50 to 2.20 m², with an average of 1.86±0.18 m². The average flow rate used during CPB was determined to be 4476.2±420.9 mL/min. The aortic cross-clamp time ranged from 16 to 100 minutes, with an average of 52.84±17.22 minutes. The total perfusion time ranged from 53 to 167 minutes, with an average of 86.08±21.96 minutes (Table 1).

In the 50 patients included in the study, Sirtuin-1 levels ranged from 3.80 to 3.90 ng/mL in the preoperative period, with a mean of 3.85±0.02 ng/mL. Intraoperative Sirtuin-1 levels ranged from 2.88 to 3.00 ng/mL, with a mean of 2.94±0.03 ng/mL. Postoperative Sirtuin-1 levels ranged from 2.40 to 2.56 ng/mL, with a mean of 2.48±0.03 ng/mL. Caspase-3 levels ranged from 145.40 to 145.80 pg/mL in the preoperative period, with a mean of 145.60±0.04 pg/mL. During the intraoperative period, Caspase-3 levels ranged from 198.20 to 198.40 pg/mL, with an average of 198.30±0.03 pg/mL. In the postoperative period, Caspase-3 levels ranged from 236.80 to 237.00 pg/mL,

TABLE 1. Demographic and Perioperative Characteristics of the Patients (n=50)

Categorical Variables		Frequency	Percent	Valid Percent	Cumulative Percent
Gender	Female	22	44.0	44.0	44.0
	Male	28	56.0	56.0	100.0
Operation	CABG×1	5	10.0	10.0	10.0
	CABG×2	8	16.0	16.0	26.0
	CABG×3	20	40.0	40.0	66.0
	CABG×4	12	24.0	24.0	90.0
	CABG×5	5	10.0	10.0	100.0
Continuous Variables		**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Age (year)		42.00	76.00	61.180	7.232
Height (m)		1.50	1.85	1.659	0.081
Weight (kg)		55.00	103.00	78.620	13.390
BSA (m^2)		1.50	2.20	1.861	0.180
Flow (L/min)		3670.00	5290.00	4476.200	420.911
Cross clamp time (min)		16.00	100.00	52.840	17.220
Total perfusıon time (min)		53.00	167.00	86.080	21.959

BSA, body surface area; CABG, coronary artery bypass grafting.

TABLE 2. Patients' Sirtuin-1 and Caspase 3 Levels Over Time

	Descriptive Statistics (n=50)			
	Minimum	Maximum	Mean	Std. Deviation
Preoperative Sirtuin-1 (ng/mL)	3.80	3.90	3.850	0.015
Intraoperative Sirtuin-1 (ng/mL)	2.88	3.00	2.940	0.025
Postoperative Sirtuin-1 (ng/mL)	2.40	2.56	2.480	0.026
Preoperative Caspase 3 (pg/mL)	145.40	145.80	145.600	0.043
Intraoperative Caspase 3 (pg/mL)	198.20	198.40	198.302	0.027
Postoperative Caspase 3 (pg/mL)	236.80	237.00	236.900	0.033

with an average of 236.90±0.03 pg/mL (Table 2).

Since the data were found not to follow a normal distribution in the normality analysis, the temporal change in Sirtuin-1 levels was evaluated using the Wilcoxon Signed Ranks test. Intraoperative Sirtuin-1 levels showed a significant difference compared to the preoperative period ($Z= -6.212$, $P<0.001$). Postoperative Sirtuin-1 levels showed a significant difference compared to the preoperative period ($Z= -6.229$, $P<0.001$). In addition, a statistically significant difference was found between postoperative and intraoperative Sirtuin-1 levels ($Z= -6.186$, $P<0.001$) (Table 3).

Since the data were found not to follow a normal distribution in the normality analysis, the temporal change in Caspase-3 levels was evaluated using the Wilcoxon Signed Ranks test. Intraoperative Caspase-3 levels showed a significant difference compared to the preoperative period ($Z= -6.262$, $P<0.001$). Postoperative Caspase-3 levels showed a significant difference compared to the preoperative period ($Z= -6.196$, $P<0.001$). Furthermore, a statistically significant difference was found between postoperative and intraoperative Caspase-3 levels ($Z= -6.203$, $P<0.001$) (Table 4).

DISCUSSION

This study demonstrated that CPB significantly suppressed circulating Sirtuin-1 levels and that this suppression was associated with increased apoptosis. The significant decrease in Sirtuin-1 levels detected during the intraoperative and postoperative periods compared to the preoperative period suggests that the oxidative stress, inflammatory response, and metabolic imbalances caused by CPB exert an inhibitory effect on epigenetic-metabolic regulatory mechanisms. Considering the anti-apoptotic and cytoprotective effects of Sirtuin-1, this suppression may pave the way for the activation of apoptotic pathways under cellular stress conditions. Indeed, the significant increase observed in Caspase-3 levels during the same time period indicates the activation of executive apoptosis pathways and supports a possible causal relationship between Sirtuin-1 suppression and apoptosis. It is known that Sirtuin-1 supports cell survival by deacetylating transcription factors such as Tumor Protein p53 (p53), Nuclear Factor Kappa-light-chain-enhancer of Activated B Cells (NF-κB), and Forkhead Box O; therefore, decreased Sirtuin-1 activity during CPB may contribute to the

TABLE 3. Comparison of Sirtuin-1 Levels Across Preoperative, Intraoperative, and Postoperative Periods

	Test Statistics[a]		
	Intraoperative Sirtuin-1 - Preoperative Sirtuin-1	Postoperative Sirtuin-1 - Preoperative Sirtuin-1	Postoperative Sirtuin-1 - Intraoperative Sirtuin-1
Z	-6.212[b]	-6.229[b]	-6.186[b]
Asymp. Sig. (2-tailed)	**<0.001**	**<0.001**	**<0.001**

[a]Wilcoxon Signed Ranks Test, [b]Based on positive ranks.
Statistically significant P-values are shown in bold.

TABLE 4. Comparison of Caspase-3 Levels Across Preoperative, Intraoperative, and Postoperative Periods

Test Statistics[a]			
	Intraoperative Caspase 3 - Preoperative Caspase 3	Postoperative Caspase 3 - Preoperative Caspase 3	Postoperative Caspase 3 - Intraoperative Caspase 3
Z	-6.262[b]	-6.196[b]	-6.203[b]
Asymp. Sig. (2-tailed)	**<0.001**	**<0.001**	**<0.001**

[a]Wilcoxon Signed Ranks Test, [b]Based on negative ranks.
Statistically significant P-values are shown in bold.

uncontrolled progression of the Caspase-3-mediated apoptotic process. These findings reveal that CPB causes cellular damage not only at the hemodynamic and inflammatory levels, but also at the epigenetic and molecular levels, and suggest that the Sirtuin-1 – Caspase-3 axis may play a critical role in understanding the molecular mechanisms of CPB-related organ damage.

During CPB, factors such as contact of blood with non-physiological surfaces, ischemia–reperfusion injury, hemodilution, and oxidative stress trigger a systemic inflammatory response. These processes result in mitochondrial dysfunction and redox imbalance at the cellular level, paving the way for the activation of apoptotic pathways [18]. In our study, the increase observed in caspase-3 levels after CPB can be considered a biochemical reflection of systemic apoptotic activation associated with CPB, consistent with previous studies.

Sirtuin-1 plays a central role in regulating the cellular stress response as an NAD^+-dependent deacetylase; it exhibits protective effects such as suppression of inflammation, reduction of oxidative stress, and inhibition of apoptosis. Through its NAD^+-dependent deacetylase activity, Sirtuin-1 directly deacetylates pro-inflammatory and pro-apoptotic transcription factors such as NF-κB p65 (RelA) and p53, thereby suppressing their activity and limiting the inflammatory response and cellular death pathways. It has been shown that a decrease in Sirtuin-1 activity under stress conditions is associated with the activation of p53 and NF-κB-mediated proapoptotic and proinflammatory pathways, leading to impaired This study demonstrated that Sirtuin-1 levels decreased significantly during the intraoperative and postoperative periods compared to the preoperative period in patients undergoing cardiac surgery with CPB, and that this suppression occurred simultaneously with an increase in Caspase-3 levels, an executor apoptosis marker. These time-dependent molecular changes reveal that CPB exerts an inhibitory effect on epigenetic-metabolic regulatory mechanisms and is associated with an increased apoptotic response. The findings suggest that the Sirtuin-1–Caspase-3 axis may play an important role in understanding the molecular basis of cellular damage associated with CPB and indicate that Sirtuin-1 could be a potential biomarker and therapeutic target for future protective or modulatory strategies.

Ethics Approval and Consent to Participate

This study was approved by the Harran University Clinical Research Ethics Committee (Decision No: HRÜ/25.18.33; date: 17.11.2025. All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Furthermore, informed consent was obtained from all participants involved in the study.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: BA; Study Design: BA, ÖG, ME, ŞY; Supervision: BA, ÖG, ME, ŞY; Funding: BA, ÖG; Materials: BA, ÖG; Data Collection and/or Processing: BA, ÖG, ME, ŞY; Statistical Analysis

and/or Data Interpretation: BA, ÖG, ME; Literature Review: BA, ÖG; Manuscript Preparation: BA, ÖG, ME; and Critical Review: BA, ÖG, ME, ŞY.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript. Two of the authors of this article (ME, ŞY) are the member of the Editorial Board of this journal. They were completely blinded to the peer review process of the article.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare..

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Zheng S, Yang L, Dai Q, Li X, Masuoka T, Lv J. Role of sirtuin 1 in depression induced coronary heart disease: Molecular pathways and therapeutic potential (Review). Biomed Rep. 2025;22(3):46. doi: 10.3892/br.2025.1924.
2. Heidari L, Ghaderian SMH, Bastami M, et al. Reverse expression pattern of sirtuin-1 and histone deacetylase-9 in coronary artery disease. Arch Physiol Biochem. 2023;129(1):46-53. doi: 10.1080/13813455.2020.1797100.
3. Doulamis IP, Tzani AI, Konstantopoulos PS, et al. A sirtuin 1/MMP2 prognostic index for myocardial infarction in patients with advanced coronary artery disease. Int J Cardiol. 2017;230:447-453. doi: 10.1016/j.ijcard.2016.12.086.
4. Raja SG, Berg GA. Impact of off-pump coronary artery bypass surgery on systemic inflammation: current best available evidence. J Card Surg. 2007;22(5):445-455. doi: 10.1111/j.1540-8191.2007.00447.x.
5. Naase H, Harling L, Kidher E, et al. Toll-like receptor 9 and the inflammatory response to surgical trauma and cardiopulmonary bypass. J Cardiothorac Surg. 2020;15(1):137. doi: 10.1186/s13019-020-01179-y.
6. Ferreira LO, Vasconcelos VW, Lima JS, et al. Biochemical Changes in Cardiopulmonary Bypass in Cardiac Surgery: New Insights. J Pers Med. 2023;13(10):1506. doi: 10.3390/jpm13101506.
7. Zakkar M, Guida G, Suleiman MS, Angelini GD. Cardiopulmonary bypass and oxidative stress. Oxid Med Cell Longev. 2015;2015:189863. doi: 10.1155/2015/189863.
8. Anselmi A, Abbate A, Girola F, et al. Myocardial ischemia, stunning, inflammation, and apoptosis during cardiac surgery: a review of evidence. Eur J Cardiothorac Surg. 2004;25(3):304-311. doi: 10.1016/j.ejcts.2003.12.003.
9. Fischer UM, Tossios P, Huebner A, Geissler HJ, Bloch W, Mehlhorn U. Myocardial apoptosis prevention by radical scavenging in patients undergoing cardiac surgery. J Thorac Cardiovasc Surg. 2004;128(1):103-108. doi: 10.1016/j.jtcvs.2003.11.034.
10. Vachharajani VT, Liu T, Wang X, Hoth JJ, Yoza BK, McCall CE. Sirtuins Link Inflammation and Metabolism. J Immunol Res. 2016;2016:8167273. doi: 10.1155/2016/8167273.
11. Yang Y, Liu Y, Wang Y, et al. Regulation of SIRT1 and Its Roles in Inflammation. Front Immunol. 2022;13:831168. doi: 10.3389/fimmu.2022.831168.
12. Santos L, Benitez-Rosendo A, Bresque M, Camacho-Pereira J, Calliari A, Escande C. Sirtuins: The NAD+-Dependent Multifaceted Modulators of Inflammation. Antioxid Redox Signal. 2023;39(16-18):1185-1208. doi: 10.1089/ars.2023.0295.
13. Autiero I, Costantini S, Colonna G. Human sirt-1: molecular modeling and structure-function relationships of an unordered protein. PLoS One. 2008;4(10):e7350. doi: 10.1371/journal.pone.0007350.
14. Chan SH, Hung CH, Shih JY, et al. SIRT1 inhibition causes oxidative stress and inflammation in patients with coronary artery disease. Redox Biol. 2017;13:301-309. doi: 10.1016/j.redox.2017.05.027.
15. Luo G, Jian Z, Zhu Y, et al. Sirt1 promotes autophagy and inhibits apoptosis to protect cardiomyocytes from hypoxic stress. Int J Mol Med. 2019;43(5):2033-2043. doi: 10.3892/ijmm.2019.4125.
16. Eefting F, Rensing B, Wigman J, et al. Role of apoptosis in reperfusion injury. Cardiovasc Res. 2004;61(3):414-426. doi: 10.1016/j.cardiores.2003.12.023.
17. Sulpizi M, Rothlisberger U, Carloni P. Molecular dynamics studies of caspase-3. Biophys J. 2003;84(4):2207-2215. doi: 10.1016/S0006-3495(03)75026-7.
18. Kovacević M, Simić O, Jonjić N, Stifter S. Apoptosis and cardiopulmonary bypass. J Card Surg. 2007;22(2):129-134. doi: 10.1111/j.1540-8191.2006.00355.x.
19. Liu TF, McCall CE. Deacetylation by SIRT1 Reprograms Inflammation and Cancer. Genes Cancer. 2013;4(3-4):135-147. doi: 10.1177/1947601913476948.
20. Yeung F, Hoberg JE, Ramsey CS, et al. Modulation of NF-kappaB-dependent transcription and cell survival by the SIRT1 deacetylase. EMBO J. 2004;23(12):2369-2380. doi:

10.1038/sj.emboj.7600244.
21. Cattelan A, Ceolotto G, Bova S, et al. NAD(+)-dependent SIRT1 deactivation has a key role on ischemia-reperfusion-induced apoptosis. Vascul Pharmacol. 2015;70:35-44. doi: 10.1016/j.vph.2015.02.004.
22. Hwang JW, Yao H, Caito S, Sundar IK, Rahman I. Redox regulation of SIRT1 in inflammation and cellular senescence. Free Radic Biol Med. 2013;61:95-110. doi: 10.1016/j.freeradbiomed.2013.03.015.

The European Research Journal 2026;12(3):347-357 Original Article

DOI: https://doi.org/10.18621/eurj.1849232 *Surgical Diseases Nursing*

Determination of Factors Affecting Oxygen Saturation During Transfer From the Operating Room to the Post-Anesthesia Care Unit

Emine Arıcı Parlak[1], Neslihan Ilkaz[1], Hatice Ayhan[1], Emine İyigün[1]

[1]Department of Surgical Diseases Nursing, Gülhane Faculty of Nursing, University of Health Sciences Türkiye, Ankara, Türkiye

ABSTRACT

Objectives: This study aimed to determine the incidence of desaturation (SpO_2 ≤94%) during the transfer of postoperative patients from the operating room to the post-anesthesia care unit (PACU) without oxygen support and to identify factors associated with desaturation.

Methods: This prospective cross-sectional study was conducted between July 2023 and July 2024 in the operating room and PACU of a training and research hospital. The sample consisted of 164 patients. Data were collected using the Oxygen Saturation Assessment Form. Post-extubation oxygen saturation levels were continuously monitored with a pulse oximeter during transfer to the PACU and until discharge.

Results: During transfer from the operating room to the PACU, 43.3% of patients had SpO_2 ≤94%, and 8.4% of these patients had SpO_2 <90%. Patients who developed desaturation had significantly higher mean age and body mass index (BMI) and lower preoperative SpO_2 values (P<0.001). Desaturation was significantly associated with female sex and the presence of chronic disease (P=0.032), whereas no significant association was found with smoking history or ASA score (P>0.05). In addition, desaturation was associated with general surgical procedures, the supine position, upper abdominal and thyroid surgeries, and the use of neuromuscular blocking agents (P<0.05).

Conclusions: Approximately half of postoperative patients experienced SpO_2 ≤ 94% in the early postoperative period. Advanced age, higher BMI, lower preoperative SpO_2 values, comorbidities, upper abdominal and thyroid surgical incisions, and the use of respiratory-depressant neuromuscular agents were identified as key factors negatively affecting postoperative oxygen saturation.

Keywords: Oxygen Saturation, Patient Transfer, Perioperative Nursing, Post-Anesthesia Care Unit, Postoperative Desaturation, Risk Factors

The early postoperative period is characterized by the persistence of physiological effects of anesthesia, one of which is hypoxemia [1, 2]. A decrease in oxygen saturation measured by pulse oximetry is the earliest and primary indicator of hypoxemia. Hypoxemia is a critical oxygen deficiency that may develop depending on the type of surgery or anesthesia and typically becomes evident within the first 20 minutes after extubation [3, 4]. It has been reported that a substantial proportion of desaturated

Received: December 25, 2025 **Accepted:** January 26, 2026 **Available Online:** January 28, 2026 **Published:** March 1, 2026

How to cite this article: Arıcı Parlak E, Ilkaz N, Ayhan H, İyigün E. Determination of Factors Affecting Oxygen Saturation During Transfer From the Operating Room to the Post-Anesthesia Care Unit. Eur Res J. 2026;12(3):347-357. doi: 10.18621/eurj.1849232

Corresponding author: Emine Arıcı Parlak, RN, PhD., Assist. Prof., E-mail: emine.ariciparlak@sbu.edu.tr

Available Online at https://dergipark.org.tr/en/pub/eurj

patients experience hypoxemia particularly during the initial minutes of transfer [4-6]. Therefore, administering supplemental oxygen via an oxygen mask or nasal cannula to all patients during transfer from the operating room to the post-anesthesia care unit (PACU) and in the early post-anesthesia period may reduce the development or severity of hypoxemia [4, 7, 8]. In this regard, the transfer to the PACU represents a critical period for nurses and anesthesiologists in terms of mitigating hypoxemia risk [2].

Postoperative hypoxemia may lead to cardiac and neurological complications, prolong hospital stay, and result in additional costs [5]. Normal postoperative oxygen saturation during anesthesia and emergence from anesthesia is expected to range between 95% and 100%; values of 94% and below are considered desaturation, while levels below 90% are regarded as a clinical emergency [9, 10-12]. Furthermore, guidelines for the prevention of surgical site infections recommend the administration of a high fraction of inspired oxygen in the postoperative period and maintaining oxygen saturation above 95% [13, 14]. In the literature, reported rates of desaturation during transfer from the operating room to the PACU or in the early postoperative period range from 13% to 35% [2, 4, 6], while hypoxemia rates range from 14% to 45% [9, 15, 16]. It is observed that the factors affecting oxygen saturation during transfer from the operating room to the PACU have not been clearly established and that desaturation incidence varies widely. One factor whose impact on oxygen saturation at initial PACU admission remains controversial is the provision of supplemental oxygen during transfer. The literature indicates that while some anesthesiologists transfer patients from the operating room to the PACU with supplemental oxygen [17, 18], others do not consider it necessary [19, 20]. This suggests that oxygen administration practices during transfer vary across institutions. Such variability may stem from the absence of clear, high-level, and binding recommendations regarding the transfer process in existing guidelines, as well as from inconsistencies in the available evidence. Within this context, the aim of the present study is to determine the incidence of desaturation (SpO_2 ≤94%) during transfer of postoperative patients from the operating room to the PACU without supplemental oxygen and to examine the factors affecting desaturation. It is anticipated that the findings of this study will enhance the quality of patient care and make a significant contribution to clinical practice.

METHODS

The study was designed as a prospective cross-sectional investigation and was conducted between July 2023 and July 2024 in the operating room and PACU affiliated with the Department of Anesthesiology and Reanimation of a training and research hospital in Ankara. The operating suite consists of a total of 14 operating rooms. Following surgery, patients are transferred to the PACU, which is located approximately 150 meters from the operating rooms, on a stretcher and under the supervision of an anesthesiologist. This transfer process takes an average of approximately 100 seconds, and due to the short distance, supplemental oxygen is not administered during transfer. The PACU is equipped to provide early postoperative monitoring and care for up to eight patients simultaneously. After completion of postoperative monitoring, patients who meet the PACU discharge criteria are transferred to their respective inpatient wards.

Study Population and Sample Size

The study population consisted of all elective surgical cases scheduled between July 2023 and July 2024. The total number of surgeries performed during the previous year was 22,459. The study sample comprised 164 cases, as determined using the Raosoft Sample Size Calculator, with a confidence level of 80% and a margin of error of 5% [21].

The inclusion criteria were as follows: willingness to participate in the study, age 18 years or older, and an American Society of Anesthesiologists (ASA) physical status classification of I, II, or III. The exclusion criteria included a preoperative oxygen saturation of ≤94% and the presence of anemia, defined as a hemoglobin level of <12 g/dL in women and <13 g/dL in men. Patients classified as ASA IV were excluded due to their higher baseline risk of hypoxemia and increased perioperative physiological instability, which could

have affected postoperative oxygen saturation measurements. Additionally, patients who developed intraoperative hypothermia (body temperature < 36.0°C) or who could not be extubated postoperatively were excluded from the study.

Oxygen Saturation Assessment Form

In this study, data were collected using a structured form developed by the researchers based on the relevant literature [4–6]. The form includes variables related to patients' sociodemographic and clinical characteristics, such as age, sex, height, weight, presence of chronic disease, history of smoking, and ASA score. In addition, it contains data describing surgical and anesthetic characteristics, including the type of surgical clinic, surgical position, incision site, type of anesthesia, use of neuromuscular blocking agents, sedation status, duration of anesthesia, and transfer time.

Operational Definitions

Early postoperative desaturation was defined as an peripheral oxygen saturation (SpO_2) value below 94%, measured by pulse oximetry, persisting for at least 30 seconds during the transfer from the operating room to the PACU and within the first 20 minutes of the postoperative period in the recovery unit [4, 9, 10].

Sedation score was defined as follows: 0 = awake; 1 = drowsy, arousable by verbal stimulation; 2 = drowsy, arousable by tactile stimulation; 3 = unarousable [4].

Data Collection Procedure

Patients who met the inclusion criteria were informed about the study by the researchers, and written and verbal informed consent was obtained. Study data were collected using a structured data collection form through face-to-face interviews conducted in the operating room prior to surgery, as well as from patient medical records, and were subsequently recorded. Patients' oxygen saturation levels were measured using a pulse oximeter. The study was conducted in two phases.

In the first phase, the first ten items of the researcher-developed data collection form were used to assess patients' demographic and baseline clinical characteristics, including age, sex, height, weight, presence of chronic disease, ASA score, history of smoking, and oxygen saturation. These data were collected and recorded through face-to-face interviews. Following surgery, the same data collection form was used to systematically record information related to the surgical procedure and anesthesia administered to the patients.

The second phase of the study involved monitoring and evaluating the oxygen saturation levels of postoperative extubated patients from the time of extubation through transfer to the PACU and until discharge from the PACU, using a pulse oximeter (Finger Clip Pulse Oximeter, 2019). Immediately after extubation, the patients' initial oxygen saturation values were measured and recorded by the researchers. Oxygen saturation monitoring continued during the transfer to the PACU, with patients being transferred with the pulse oximeter attached to their finger. The lowest oxygen saturation value was recorded every 30 seconds. Pulse oximeter devices typically provide averaged signal readings every 6–8 seconds, and threshold alarms are generally set at an average of 90%, depending on the unit or institutional protocols [15]. No patients received supplemental oxygen during the transfer process. Finally, patients' initial oxygen saturation values upon arrival at the PACU and their oxygen saturation values at the time of PACU discharge were recorded. Patients were discharged from the PACU when an Aldrete score of ≥9 was achieved.

Statistical Analysis

Study data were analyzed using SPSS version 22.0 (SPSS Inc., Chicago, IL, USA). Descriptive statistics, including frequencies, percentages, means, and standard deviations, were used to summarize the data. Normality of distribution was assessed using the Kolmogorov–Smirnov test. For variables that showed a normal distribution, Chi-square tests and Student's t-tests were applied to compare selected demographic, surgical, and anesthetic characteristics between patients with and without desaturation. The relationships between selected demographic and perioperative variables and minimum and mean SpO_2 values were examined using Pearson correlation analysis. In all statistical analyses, a P-value <0.05 was considered statistically significant.

RESULTS

Patients (n=32) who developed intraoperative hypothermia (body temperature <36°C) or those (n=13) who could not be extubated postoperatively were excluded from the analysis. According to the study findings, 43.3% (n=71) of patients experienced a decrease in oxygen saturation to ≤94% during the transfer to the PACU. Additionally, 8.4% (n=6) of desaturated patients were found to have oxygen saturation levels below 90%.

The findings related to the comparison of patients' demographic and clinical characteristics according to desaturation status are presented in Table 1. The mean age of patients who developed desaturation was 51.95±13.4 years, whereas it was 41.4±16.6 years in those without desaturation, and the difference between the two groups was statistically significant (P<0.001). The mean body mass index (BMI) of desaturated patients was 28.3±4.7, compared with 25.2±5.3 in non-desaturated patients, and this difference was also statistically significant (P<0.001). It was determined that 73.2% (n=52) of desaturated patients were female and 59.2% (n=42) had at least one chronic disease. Statistically significant differences were observed between the groups with respect to sex and the presence of chronic disease (P=0.032 and P<0.001, respectively). In contrast, no statistically significant differences were found between desaturated and non-desaturated patients in terms of smoking history and ASA score (P>0.05). Additionally, preoperative oxygen saturation levels were found to be significantly lower in desaturated patients compared with those without desaturation (P<0.001).

The findings regarding the comparison of patients' surgical and anesthetic characteristics according to

TABLE 1. Comparison of Patients' Sociodemographic and Clinical Characteristics According to Postoperative Desaturation Status (n=164)

Variables	Postoperative SpO_2		Test statistic	P-value
	Nondesaturation SpO_2 >94 (n=93)	Desaturation SpO_2 ≤94 (n=71)		
Age (year)	41.4±16.6	51.95±13.4	-4.367	**<0.001[b]**
BMI (kg/m²)	25.2±5.3	28.3±4.7	-3.829	**<0.001[b]**
Gender				
Female	53 (57.0%)	52 (73.2%)	4.616	**0.032[a]**
Male	40 (43.0%)	19 (26.8%)		
Chronic diseases				
Yes	19 (20.4%)	42 (59.2%)	25.846	**<0.001[a]**
No	74 (79.6%)	29 (40.8%)		
ASA score				
I	27 (29.0%)	12 (16.9%)	3.271	0.195[a]
II	58 (62.4%)	52 (73.2%)		
III	8 (8.6%)	7 (9.9%)		
History of smoking				
Yes	32 (34.4%)	25 (35.2%)	0.011	0.915[a]
No	61 (65.6%)	46 (64.8%)		
Preoperative SpO_2	97.7±1.36	96.6±1.5	4.941	**<0.001[b]**

Data are shown as Mean±standard deviation or n (%) where appropriate. BMI, body mass index; ASA, American Society of Anesthesiologists; SpO_2, peripheral oxygen saturation.

[a]Chi-square test. [b] Independent samples *t*-test. Statistically significan P-values are shown in bold.

TABLE 2. Comparison of Surgical and anesthetic Characteristics According to Postoperative Desaturation Status (n=164)

Variables	Postoperative SpO_2		Test Statistic	P-value
	Nondesaturation SpO_2 >94 (n=93)	Desaturation SpO_2 ≤94 (n=71)		
Type of surgical clinic				
General surgery	51 (54.8%)	52 (73.2%)		
Thoracic surgery	18 (19.4%)	8 (11.3%)		
Neurosurgery	17 (18.3%)	5 (7.0%)	7.665	**0.053**[a]
Obstetrics and gynecology	7 (7.5%)	6 (8.5%)		
Surgical Position				
Supine	59 (63.4%)	37 (52.1%)	10.683	**0.030**[a]
Reverse Trendelenburg	22 (23.7%)	22 (31.0%)		
Prone	6 (6.5%)	4 (5.6%)		
Lithotomy	6 (6.5%)	2 (2.8%)		
Lateral	0 (0.0%)	6 (8.5%)		
Incision site				
Upper abdominal region	25 (26.9%)	22 (31.0%)	22.741	**<0.001**[a]
Lower abdominal region	15 (16.1%)	9 (12.7%)		
Sternotomy incision	18 (19.4%)	8 (11.3%)		
Breast surgery	18 (19.4%)	8 (11.3%)		
Thyroid surgery	3 (3.2%)	19 (26.8%)		
Lumbar region / lower back	14 (15.1%)	5 (7.0%)		
Anesthesia type				
General anesthesia	76 (81.7%)	64 (90.1%)	2.287	0.319[a]
Laryngeal mask airway	12 (12.9%)	5 (7.0%)		
Spinal anesthesia	5 (5.4%)	2 (2.8%)		
Neuromuscular blocking agent				
Yes	52 (55.9%)	51 (71.8%)	4.367	**0.037**[a]
No	41 (44.1%)	20 (28.2%)		
Sedation status, n (%)				
Awake	51 (54.8%)	37 (52.1%)	0.120	0.729[a]
Drowsy, arousable to verbal stimulation	42 (45.2%)	34 (47.9%)		
Duration of anesthesia (min)	126.1±59.73	133.8±62.2	-0.803	0.423[b]
Transfer time (seconds)	96.3±27.1	99.0±31.4	-0.610	0.543[b]

Data are shown as Mean±standard deviation or n (%) where appropriate. BMI, body mass index; SpO_2, peripheral oxygen saturation. [a]Chi-square test. [b]Independent samples *t*-test. Statistically significan P-values are shown in bold.

desaturation status are presented in Table 2. As shown in Table 2, 73.2% (n=52) of desaturated patients underwent general surgical procedures, 52.1% (n=37) were operated on in the supine position, 31.0% (n=22) underwent upper abdominal surgery, 26.8% (n=19) underwent thyroid surgery, and 71.8% (n=51) received neuromuscular blocking agents. Statistically significant differences were observed between desaturated and non-desaturated patients with respect to these variables (P<0.05).

It was also observed that 90.1% (n=64) of desaturated patients received general anesthesia and 52.1% (n=37) were awake during the transfer process. However, no statistically significant differences were found between the groups in terms of type of anesthesia and sedation status (P>0.05). Similarly, when patients were compared according to duration of anesthesia and transfer time, no statistically significant differences were identified (P=0.423 and P=0.543, respectively).

The relationships between minimum and mean SpO_2 values and selected demographic and perioperative variables are presented in Table 3. According to the findings, age and mean BMI were moderately, negatively, and statistically significantly correlated with both minimum and mean SpO_2 values (P<0.05). In contrast, preoperative SpO_2 levels and the first oxygen saturation values measured after extubation showed a strong positive and statistically significant correlation with both minimum and mean SpO_2 values (P<0.001). However, no statistically significant correlations were identified between duration of anesthesia or transfer time and minimum or mean SpO_2 values (P>0.05).

Changes in patients' mean SpO_2 levels monitored from the preoperative period until discharge from the PACU are presented in Figure 1. Accordingly, the mean SpO_2, which was 97.3% in the preoperative period, gradually decreased during the first three measurements following extubation, reaching approximately 94.0%. The lowest mean SpO_2 value was observed at initial admission to the PACU (93.5%). Subsequently, patients' mean SpO_2 levels in the PACU increased to approximately 96.5%.

DISCUSSION

In this study, the incidence of desaturation (SpO_2 ≤94%) observed during the transfer of patients from the operating room to the PACU, as well as the factors affecting oxygen saturation, were comprehensively examined. It is recommended that patients' SpO_2 levels be maintained above 95% in the PACU [15]. In addition, oxygen saturation is expected to remain between 95% and 100% during anesthesia and the emergence period, with SpO_2 values of ≤94% considered desaturation and values below 90% regarded as a clinical emergency [10, 11, 22]. In the our study, 43.3% of patients experienced desaturation during transfer from the operating room to the PACU, and 8.4% of the desaturated patients had oxygen

TABLE 3. Relationship Between Patient and Procedure-Related Factors and Minimum/Mean Peripheral Oxygen Saturation*

Variables	Postoperative SpO_2			
	Minimum SpO_2		Mean SpO_2	
	r	P-value	r	P-value
Age (year)	-0.344	**<0.001**	-0.310	**<0.001**
BMI (kg/m²)	-0.287	**0.001**	-0.292	**<0.001**
Duration of anesthesia (min)	-0.043	0.581	-0.019	0.805
Transfer time (seconds)	-0.110	0.162	-0.128	0.104
Preoperative SpO_2	0.436	**<0.001**	0.398	**<0.001**
First SpO_2 after extubation	0.779	**<0.001**	0.908	**<0.001**

SpO_2, peripheral oxygen saturation; BMI, body mass index.

*Pearson correlation. Statistically significan P-values are shown in bold.

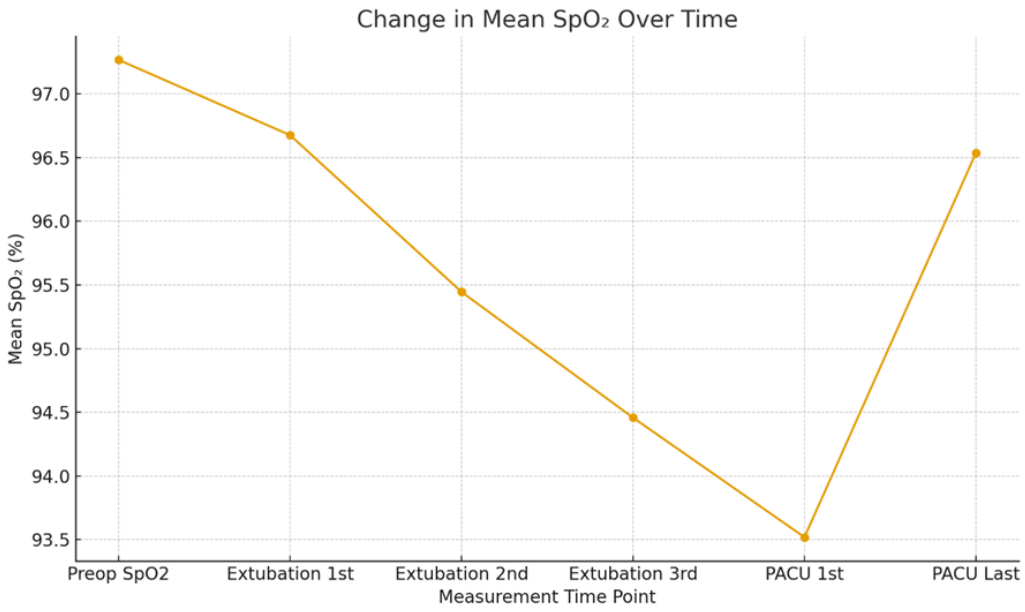

FIGURE 1. Mean oxygen saturation changes across six time points from the preoperative period to PACU transfer and discharge. SpO_2, peripheral oxygen saturation; PACU, post-anesthesia care unit.

saturation levels below 90%. In the literature, postoperative hypoxemia is most commonly defined using an SpO_2 threshold of <90%, with the reported incidence ranging between 13% and 35% [2, 4, 6,23]. Studies conducted in Ethiopia have reported hypoxemia rates ranging from 22.7% [24] to 26.7% [4], while earlier studies have indicated rates as high as 61.4% [25]. In the study, the relatively lower incidence of hypoxemia may be attributable to the short transfer duration. However, some studies define desaturation as SpO_2 ≤94% and report higher rates accordingly. Taye *et al*. [9] reported that 45.8% of patients experienced SpO_2 levels of ≤94% in the early postoperative period, while this rate was 14.5% in the study by Ramachandran *et al*. [15] and 13.5% in the study by Kaushal *et al*. [16]. In the study conducted by Taye *et al*. [9], hypoxemia was observed in patients who did not receive supplemental oxygen during transfer, whereas in the study by Kaushal *et al*. [16], oxygen saturation values were monitored while patients were breathing room air. In contrast, the study by Ramachandran *et al*. [15] did not provide information regarding supplemental oxygen use during transfer. Similar to the study by Taye *et al*. [9], our study found that 43.5% of patients had oxygen saturation levels of SpO_2 ≤94%. In addition, none of the aforementioned studies reported the distance between the operating room and the PACU. In our study, however, the PACU is located close to the operating rooms; therefore, supplemental oxygen was not administered to patients during transfer. It is thought that these differences in the incidence of hypoxemia may be attributable to variations in practices during the transfer process from the operating room to the PACU after extubation, as well as institutional differences in hospital physical layouts. The literature identifies various factors influencing the incidence of hypoxemia during transfer from the operating room to the PACU or in the early postoperative period, with advanced age consistently reported as one of the most prominent risk factors. Numerous studies have demonstrated a significant increase in hypoxemia incidence among individuals aged 55 years and older [4, 16, 19, 26, 27]. Maity *et al*. [23] similarly reported that the rate of hypoxemia in elderly patients was nearly twice as high as that observed in younger patients (43.47% vs. 17.24%) [26]. Consistent with these findings, our study revealed that patients who developed desaturation had a higher mean age, and a strong, negative, and statistically significant correlation was identified between age and oxygen saturation levels. These results suggest that age-related physiological changes in the lungs may adversely affect postoperative

oxygenation.

The literature indicates that BMI is an important determinant of postoperative oxygenation. Kaushal *et al*. [16] demonstrated that the risk of hypoxemia increases in patients with a BMI greater than 25, with this risk becoming particularly pronounced in individuals with a BMI ≥ 30 kg/m². Similarly, Taye *et al*. [9] reported a higher frequency of hypoxemia in patients with BMI >25 kg/m², while Labeste *et al*. [6] found increased hypoxemia rates among patients with BMI >30 kg/m². Walker *et al*. [27] also reported a similar BMI threshold value of approximately 29.5. Consistent with these findings, our study identified a significant negative relationship between BMI and postoperative oxygen saturation, and the mean BMI of desaturated patients was above 25. This finding suggests that pathophysiological mechanisms observed in obese patients, such as reduced chest wall compliance, upper airway obstruction, and central apnea, may contribute to an increased risk of postoperative hypoxemia.

The literature indicates that patients' preoperative hemodynamic status, particularly preoperative SpO_2 levels, is closely associated with postoperative oxygenation. Andualem and Yesuf [26] reported that the risk of postoperative hypoxemia was four times higher in patients with preoperative SpO_2 <95%. Similarly, numerous studies have demonstrated that preoperative SpO_2 levels below 95% significantly increase the risk of hypoxemia [4, 6, 9]. In contrast, some studies have reported that preoperative oxygen saturation does not significantly affect the incidence of hypoxemia in the early postoperative period [4, 25]. In our study, however, consistent with the majority of the literature, preoperative SpO_2 levels were significantly lower in patients who developed desaturation. This finding highlights that identifying low oxygen saturation during the preoperative period and implementing appropriate preventive measures constitute a critical area of assessment for nurses, both in reducing the risk of hypoxemia and in planning safe postoperative care.

The literature has demonstrated through studies that smoking is an important risk factor for postoperative hypoxemia [9, 16, 28]. However, in our study, no statistically significant difference was found between desaturated and non-desaturated patients with respect to smoking status. This finding may be related to the fact that desaturation observed in this study was predominantly mild in severity. In contrast, when comorbidity status was evaluated, the frequency of comorbid conditions was found to be higher in the desaturated group. Consistent with this result, the literature indicates that the risk of hypoxemia is significantly increased in individuals with chronic diseases [4, 9, 16]. This finding suggests that patients with respiratory, cardiovascular, and metabolic disorders require more careful monitoring of oxygenation during the postoperative period. The presence of comorbidities should therefore be thoroughly assessed during the preoperative evaluation as a critical risk factor. Such information facilitates the anticipation of potential postoperative oxygenation problems and supports the planning of individualized nursing interventions, including supplemental oxygen therapy, patient positioning, breathing exercises, and close monitoring. Consequently, these measures contribute to the early prevention of complications and the promotion of a safe postoperative care process.

When surgery-related factors were examined, the incidence of desaturation in this study was found to be higher in procedures performed in the supine or reverse Trendelenburg positions, as well as in surgeries involving upper abdominal and thyroid incisions. Although Kaushal *et al*. [16] reported that the type of surgical incision was not a determining risk factor for oxygen saturation, another study, consistent with our findings, demonstrated an association between subcostal incisions and early postoperative hypoxemia [4]. Regarding anesthesia-related and anesthetic risk factors, no significant difference was observed between the groups in terms of type of anesthesia in the present study. In contrast, the literature reports that general anesthesia increases the risk of postoperative hypoxemia [2, 4, 19, 26]. The lack of a significant difference in our study may be attributable to the fact that the vast majority of patients in both groups underwent surgery under general anesthesia. The literature presents inconsistent findings regarding the duration of anesthesia. Some studies have reported that an anesthesia duration longer than 120 minutes increases the risk of hypoxemia [4, 27]. In contrast, Kaushal *et al*. [16] indicated that anesthesia duration is not an independent risk factor, and similarly, the studies by

Maity *et al*. [23] and Taye *et al*. [9] found no significant association between duration and hypoxemia. Consistent with these findings, the present study also identified no significant relationship between anesthesia duration or transfer time and desaturation. The findings of Walker *et al*. [27], which suggest that transport duration is not a determining factor for hypoxemia, are likewise in agreement with the results of our study.

Strengths and Limitations

The assessment of oxygen saturation levels in patients transferred from the operating room to the PACU without oxygen support, under routine clinical conditions, using a prospective design and continuous pulse oximetry monitoring, constitutes a key strength of this study. The limitation of this study is that it was conducted in a single training and research hospital, which may restrict the generalizability of the findings to institutions with different infrastructure, transfer distances, and postoperative care practices.

CONCLUSION

The findings of this study demonstrate that, in the early postoperative period during which all patients were transferred without supplemental oxygen, nearly half of the patients had oxygen saturation levels of SpO_2 ≤94%. Advanced age, higher BMI, preoperative SpO_2 levels close to the lower limit of normal, presence of comorbidities, upper abdominal and thyroid surgical incisions, and the use of respiratory-depressant neuromuscular blocking agents during anesthesia were identified as factors that negatively affect postoperative oxygen saturation. Therefore, it is recommended that operating room nurses and the anesthesia team comprehensively assess patients for these risk factors during the preoperative period and identify patients at increased risk. In addition, continuous monitoring of SpO_2 levels during patient transfer and the provision of supplemental oxygen during transfer for high-risk patients are recommended. This approach, together with effective communication of patient status to the care team upon handover to the PACU, is considered a prophylactic, simple, and effective nursing intervention that contributes to maintaining adequate oxygenation in the early postoperative period.

Ethics Approval and Consent to Participate

The study was approved by the Ankara Medipol University Non-Interventional Clinical Research Ethics Committee (Decision no.: 84, date: 11.07.2023). All procedures performed during data collection, review of patient records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Written informed consent was obtained from all individual participants included in the study. In addition, institutional permission was obtained from the relevant clinic and the hospital where the study was conducted.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: EAP, NI, HA; Study Design: EAP, NI, HA; Supervision: HA, Eİ; Funding: N/A; Materials: EAP; Data Collection and/or Processing: EAP, NI, HA; Statistical Analysis and/or Data Interpretation: EAP, NI, HA, Eİ, Literature Review: EAP, NI, HA; Manuscript Preparation: EAP, NI, HA; and Critical Review: EAP, NI, HA, Eİ.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgment

The authors would like to thank the all participating patiens involved.

Generative Artificial Intelligence Statement

The author(s) declare that artificial intelligence–

based tools were used in accordance with academic ethical standards during the preparation of this manuscript. ChatGPT (version 5.2) was utilized solely to assist with language editing, translation, and grammatical corrections. The authors maintained full responsibility for the content and ensured complete oversight and final control of the manuscript.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Hines R, Barash PG, Watrous G, O'Connor T. Complications occurring in the postanesthesia care unit: a survey. Anesth Analg. 1992;74(4):503-509. doi: 10.1213/00000539-199204000-00006.
2. Siddiqui N, Arzola C, Teresi J, Fox G, Guerina L, Friedman Z. Predictors of desaturation in the postoperative anesthesia care unit: an observational study. J Clin Anesth. 2013;25(8):612-617. doi: 10.1016/j.jclinane.2013.04.018.
3. Jones JG, Sapsford DJ, Wheatley RG. Postoperative hypoxaemia: mechanisms and time course. Anaesthesia. 1990;45(7):566-573. doi: 10.1111/j.1365-2044.1990.tb14833.x.
4. Melesse DY, Denu ZA, Kassahun HG, Agegnehu AF. The incidence of early post-operative hypoxemia and its contributing factors among patients underwent operation under anesthesia at University of Gondar comprehensive and specialized referral hospital, North West Ethiopia: a prospective observational study. Int J Surg Open. 2020;22:38-46. doi: 10.1016/j.ijso.2019.11.011.
5. Habre W, Peták F. Perioperative use of oxygen: variabilities across age. Br J Anaesth. 2014;113 Suppl 2:ii26-36. doi: 10.1093/bja/aeu380.
6. Labaste F, Silva S, Serin-Moulin L, et al. Predictors of desaturation during patient transport to the postoperative anesthesia care unit: an observational study. J Clin Anesth. 2016;35:210-214. doi: 10.1016/j.jclinane.2016.07.018.
7. Ayhan H, Iyigun E, Tastan S, Orhan ME, Ozturk E. Comparison of two different oxygen delivery methods in the early postoperative period: randomized trial. J Adv Nurs. 2009;65(6):1237-1247. doi: 10.1111/j.1365-2648.2009.04984.x.
8. Triyasunant N, Siriussawakul A, Nimmannit A, Kunawudhi A, Jirakulsawat A. Postoperative Supplemental Oxygen Is Unnecessary in Low-Risk Cesarean Delivery Parturients Under Spinal Anesthesia: An Observational Study. AANA J. 2016;84(5):358-361.
9. Taye MG, Molla A, Teshome D, et al. Predictors of hypoxemia after general anesthesia in the early postoperative period in a hospital in Ethiopia: an observational study. Multidiscip Respir Med. 2021;16(1):782. doi: 10.4081/mrm.2021.782.
10. World Health Organization. Pulse Oximetry Training Manual. Geneva: WHO; 2011.
11. Herbst A, Goel S, Beane A, et al. Oxygen saturation targets for adults with acute hypoxemia in low and lower-middle income countries: a scoping review with analysis of contextual factors. Front Med (Lausanne). 2023;10:1148334. doi: 10.3389/fmed.2023.1148334.
12. Siemieniuk RAC, Chu DK, Kim LH, et al. Oxygen therapy for acutely ill medical patients: a clinical practice guideline. BMJ. 2018;363:k4169. doi: 10.1136/bmj.k4169.
13. World Health Organization. WHO Guidelines on the Use of Pulse Oximetry. Geneva: WHO; 2016. Available from: https://www.who.int/publications/i/item/9789241550475.
14. Berríos-Torres SI, Umscheid CA, Bratzler DW, et al; Healthcare Infection Control Practices Advisory Committee. Centers for Disease Control and Prevention Guideline for the Prevention of Surgical Site Infection, 2017. JAMA Surg. 2017;152(8):784-791. doi: 10.1001/jamasurg.2017.0904.
15. Ramachandran SK, Thompson A, Pandit JJ, Devine S, Shanks AM. Retrospective observational evaluation of postoperative oxygen saturation levels and associated postoperative respiratory complications and hospital resource utilization. PLoS One. 2017;12(5):e0175408. doi: 10.1371/journal.pone.0175408.
16. Kaushal A, Goyal P, Dhiraaj S, Agarwal A, Singh PK. Identification of Various Perioperative Risk Factors Responsible for Development of Postoperative Hypoxaemia. Turk J Anaesthesiol Reanim. 2018;46(6):416-423. doi: 10.5152/TJAR.2018.82160.
17. Tyler IL, Tantisira B, Winter PM, Motoyama EK. Continuous monitoring of arterial oxygen saturation during transfer to the recovery room. Anesth Analg. 1985;64(11):1108-1112.
18. Blair I, Holland R, Lau W, McCarthy N, Chiah TS, Ledwidge D. Oxygen saturation during transfer from operating room to recovery after anaesthesia. Anaesth Intensive Care. 1987;15(2):147-150. doi: 10.1177/0310057X8701500205.
19. Moller JT, Wittrup M, Johansen SH. Hypoxemia in the postanesthesia care unit: an observer study. Anesthesiology. 1990;73(5):890-895. doi: 10.1097/00000542-199011000-00016.
20. Gift AG, Stanik J, Karpenick J, Whitmore K, Bolgiano CS. Oxygen saturation in postoperative patients at low risk for hypoxemia: is oxygen therapy needed? Anesth Analg. 1995;80(2):368-372. doi: 10.1097/00000539-199502000-00028.
21. Raosoft Inc. Sample Size Calculator. Available from: http://www.raosoft.com/samplesize.html.
22. O'Driscoll BR, Howard LS, Davison AG; British Thoracic Society. BTS guideline for emergency oxygen use in adult patients. Thorax. 2008;63 Suppl 6:vi1-68. doi: 10.1136/thx.2008.102947.
23. Maity A, Saha D, Swaika S, Maulik SG, Choudhury B, Sutradhar M. Detection of hypoxia in the early postoperative period. Anesth Essays Res. 2012;6(1):34-37. doi: 10.4103/0259-1162.103369.
24. Wolde GD, Awol MA, Obsa MS, Wesene NG, Gemechu AD. Magnitude and Associated Factors of Immediate Postoperative Hypoxemia among Elective Surgical Procedures at Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia. J Anesth

Clin Res 218;9(5):1000821. doi: 10.4172/2155-6148.1000821.
25. Meiklejohn BH, Smith G, Elling AE, Hindocha N. Arterial oxygen desaturation during postoperative transportation: the influence of operation site. Anaesthesia. 1987;42(12):1313-1315. doi: 10.1111/j.1365-2044.1987.tb05282.x.
26. Andualem AA, Yesuf KA. Incidence and associated factors of postoperative hypoxemia among adult elective surgical patients at Dessie Comprehensive Specialized Hospital: An observational study. Ann Med Surg (Lond). 2022;78:103747. doi: 10.1016/j.amsu.2022.103747.
27. Walker M, Farmer RG, Schelew B. Risk factors for oxygen desaturation on arrival in the postanesthesia care unit. Can J Anaesth. 2015;62(9):1019-1020. doi: 10.1007/s12630-015-0371-6.
28. Scholes RL, Browning L, Sztendur EM, Denehy L. Duration of anaesthesia, type of surgery, respiratory co-morbidity, predicted VO2max and smoking predict postoperative pulmonary complications after upper abdominal surgery: an observational study. Aust J Physiother. 2009;55(3):191-198. doi: 10.1016/s0004-9514(09)70081-9.

The European Research Journal 2026;12(3):358-366 Original Article

DOI: https://doi.org/10.18621/eurj.1811898 *Health Care*

Examining the Relationship Between Learning Needs and Quality of Life in Patients with Stoma

Gülistan Yurdagül[1], İsmail Dusak[2]

[1]Department of Nursing, Faculty of Health Sciences, Osmaniye Korkut Ata University, Osmaniye, Türkiye; [2]Şanlıurfa Mehmet Akif İnan Training and Research Hospital, Şanlıurfa, Türkiye

ABSTRACT

Objectives: To examine the relationship between the learning needs and quality of life of patients with a stoma.

Methods: This descriptive and correlational study was conducted with stoma patients receiving care in the surgical and oncology units of a training and research hospital between September 25, 2023 and April 25, 2024. The study was completed with 208 patients who volunteered to participate and met the inclusion criteria. Data were collected through face-to-face interviews using the Patient Information Form, Patient Learning Needs Scale, and Stoma Quality of Life Scale. Data were analyzed with SPSS 23 using Kruskal–Wallis, ANOVA, Student's t-test and Mann–Whitney U test. A significance level of $P<0.05$ was accepted.

Results: There was no significant relationship between the learning needs and quality of life of stoma patients ($P>0.05$). However, patients with rectal cancer had significantly higher scores on the body image/sexuality subscale ($t = -2.902$, $P=0.004$) and total quality of life ($t = -2.264$, $P=0.025$) compared to those with colon cancer.

Conclusions: The higher body image/sexuality and quality of life scores among patients with rectal cancer suggest that their care needs may be more related to social acceptance rather than purely physiological processes.

Keywords: Body Image, Patient Education, Nursing Care, Ostomy, Quality of Life

A stoma is a surgically created opening that connects a hollow organ of the body to the skin surface. Patients with intestinal stomas have various needs that arise from living with this condition. These needs are influenced not only by the type and construction of the stoma, but also by the patient's personality characteristics and demographic factors [1]. Although there are standardized nursing care guidelines for patients with a stoma [2, 3] each individual has different demographic and personality characteristics; therefore, in addition to standard stoma care, a holistic and individualized care plan is recommended [1, 4]. In addition, being left alone with their stoma and its care after discharge may lead to anxiety and stress in patients. Nurses should anticipate these experiences in advance and focus on meeting patient learning needs by considering their personality traits and demographic characteristics. The stress and anxiety levels of patients with an intestinal stoma may affect their readiness and motivation to learn [5]. Nurses should be aware of these psychological and individual differences and identify the learning needs

Received: October 27, 2025 **Accepted:** December 25, 2025 **Available Online:** December 29, 2025 **Published:** March 1, 2025

How to cite this article: Yurdagül G, Dusak İ. Examining the Relationship Between Learning Needs and Quality of Life in Patients with Stoma. Eur Res J. 2026;12(3):358-366. doi: 10.18621/eurj.1811898

Corresponding author: Gülistan Yurdagül, MD., Assist. Prof., Phone: +90 328 827 10 00, E-mail: yurdagulistan@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

of patients. Determining the learning needs of individuals with a stoma may improve their self-management after discharge. To enhance stoma self-management levels, various models such as the transtheoretical model, multimedia-based approaches, telehealth services, and the chronic care model have been implemented in patients with intestinal stomas [6]. However, for these models to be effectively implemented, nurses must determine the specific areas in which patients lack knowledge and assess whether they have the motivation to learn. Studies in the literature indicate that adequately addressing the learning needs of stoma patients significantly improves their self-management skills and quality of life. From a conceptual perspective, meeting patients' learning needs is considered a key component of effective self-care and adaptation to living with a stoma [4, 6]. Self-care ability and adaptation processes have been shown to be closely associated with psychosocial well-being and quality of life in individuals with a stoma [5, 7]. In this context, learning needs may be indirectly related to quality of life through their influence on self-care and adaptation.

In the literature, the quality of life of patients with intestinal stomas has often been reported as low [4, 8, 9]. The quality of life of patients with a stoma increases when they are younger than 60 years of age [7], have effective psychosocial coping skills [10], and possess strong social support systems. Moreover, poor self-care status is closely associated with lower quality of life [5]. In order for patients to perform stoma self-care effectively, they must adapt to living with a stoma. Achieving stoma adaptation contributes to an improvement in quality of life. Factors that support adaptation to the new life with a stoma include partner acceptance, social support from family and friends, returning to work, and maintaining overall health [7]. Therefore, adaptation to a stoma plays a crucial role in improving the quality of life.

Studies conducted in our country have examined factors affecting the quality of life of stoma patients, such as their sleep patterns [11, 12], religious rituals [13], sexuality [14], anxiety and sexuality [15], body image and self-perception, self-efficacy, and stoma adaptation [16, 17]. However, no studies have been found regarding their learning needs. Therefore, the aim of this study was to examine the relationship between the learning needs and quality of life of patients with a stoma.

METHODS

Study Design

The study is cross-sectional, descriptive, and correlational in design.

Population and Sample

The population of the study consisted of intestinal stoma patients who received healthcare services in the general surgery and radiotherapy/chemotherapy units of Şanlıurfa Mehmet Akif İnan Training and Research Hospital between September 25, 2023 and April 25, 2024. The sample size was evaluated in accordance with methodological recommendations for relationship-based analyses, such as correlation and regression. In the literature, it is reported that approximately 80–85 participants are sufficient to detect a moderate effect size ($r \approx 0.30$ for correlation; Cohen's $f^2 = 0.15$ for regression) [18]. In this study, a total of 208 patients who met the inclusion criteria and agreed to participate were reached, providing adequate statistical power for the analyses.

Study Setting and Timeframe

Data were collected between September 25, 2023 and April 25, 2024 through face-to-face interviews with intestinal stoma patients receiving healthcare services in the general surgery and radiotherapy/chemotherapy units of Şanlıurfa Mehmet Akif İnan Training and Research Hospital. All data in the study were collected by the researcher İ.D. using the same questionnaires, with the same questions and explanations provided to each patient. This approach was adopted to ensure standardization of the data collection process. The questionnaires and explanations were administered to all patients in the same order and with the same content, thereby aiming to minimize variations arising from the data collection procedure. Each interview lasted approximately 20–30 minutes depending on the patient's condition. Patients who were unable to continue with the study were withdrawn, and any partially completed questionnaires were excluded from the analysis.

Inclusion Criteria

•Being a stoma patient receiving healthcare services at Şanlıurfa Mehmet Akif İnan Training and

Research Hospital

•Voluntarily agreeing to participate and signing the informed consent form

•Having no psychiatric disorders and being mentally competent

•Being 18 years of age or older.

Exclusion Criteria

This study was planned only with patients with intestinal stomas; therefore, individuals with other stoma types such as tracheostomy, urostomy, and gastrostomy were excluded. Because the measurement tools used in the study were self-report-based, participants needed sufficient cognitive function to understand and carefully answer the questions. Cognitive competence was assessed by observing participants' general communication skills and orientation to person, place, and time during their responses. Considering that languages other than Turkish were used in the hospital where the study was conducted, and also because the measurement tools were administered in Turkish, individuals who did not speak Turkish were excluded from the study. To prevent data duplication, care was taken to ensure that the same participant was not included in the study more than once, and only the first completed datasets were included in the analysis. Furthermore, patients in clinical conditions requiring emergency medical intervention, such as the early postoperative period or while the effects of anesthesia were still present, were excluded from the study to both avoid risking patient safety and to ensure the reliability of the data obtained.

Data Collection Tools and Procedures

In this study, a Patient Information Form, the Stoma Quality of Life Scale, and the Patient Learning Needs Scale were used to collect data.

Patient Information Form

This form was created by researchers through a literature review. This form includes questions related to the demographic and clinical characteristics of the participants, such as age, gender, education level, marital status, number of children, health insurance status, income level, presence of chronic disease, medical diagnosis, type of stoma, status of receiving education about stoma care, and the person responsible for stoma care [11, 14–16, 19]. Since the Patient Information Form only contains descriptive questions, a pilot study was not conducted.

Stoma Quality of Life Scale

The Stoma Quality of Life Scale was developed by Baxter *et al*. in 2006 to evaluate the quality of life of individuals with a stoma [20]. The Turkish validity and reliability study of the scale was conducted by Karadağ *et al*. in 2011 [21]. The scale consists of 19 items in total, and the Cronbach's alpha coefficient of the original scale was reported as 0.87. In the present study, the Cronbach's alpha value was 0.881.

The first two items of the scale (Section 1) assess the individual's overall satisfaction with life. The first item evaluates instant satisfaction with life, while the second item evaluates overall satisfaction during the past month. This section is scored between 0 and 100, where 0 indicates complete dissatisfaction and 100 indicates complete satisfaction.

The second section of the scale consists of 5-point Likert-type items (1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Always) grouped into three sub-dimensions: Work and Social Life (6 items), Sexuality and Body Image (5 items), and Stoma Function (6 items).

Scores for the sub-dimensions are calculated as follows:

a)Work and Social Life: $25 \times (12 + S3 + S4 - S5 - S6 - S18 + S19) / 6$

b)Sexuality and Body Image: $25 \times (1 + S7 + S8 - S9 + S12 + S15) / 5$

c)Stoma Function: $25 \times (24 - S10 - S11 + S13 - S14 - S16 - S17) / 6$

Both sub-dimension scores and total scale scores range from 0 to 100. Higher scores indicate better quality of life. Questionnaires with one missing item were accepted, whereas those with two or more missing items were excluded from the analysis.

Patient Learning Needs Scale (PLNS)

The Patient Learning Needs Scale was developed by Bubela *et al*. in 1990 to determine whether the information needs of patients at the time of discharge are adequately met [22]. The Turkish validity and reliability study of the scale was conducted by Çatal and Dicle in 2008 [23]. The scale consists of 50 items

and seven sub-dimensions, rated on a 5-point Likert scale (1 = not important, 2 = slightly important, 3 = moderately important, 4 = very important, 5 = extremely important).

The seven sub-dimensions of the scale are classified as follows: medication (3, 8, 16, 18, 37, 39, 44, 45), activities of daily living (2, 5, 14, 17, 27, 28, 29, 30, 48), community and follow-up (6, 9, 22, 31, 36, 41), feelings related to condition (7, 24, 32, 35, 42), treatment and complications (1, 4, 10, 19, 20, 23, 26, 38, 47), quality of life (11, 13, 15, 21, 34, 40, 46, 50), and skin care (12, 25, 33, 43, 49). The total score of the scale ranges from 50 to 250, with higher scores indicating greater learning needs. The scale contains no reverse-coded items.

In the original Turkish validity and reliability study, the Cronbach's alpha coefficient was reported as 0.95 for the total scale, and between 0.69 and 0.88 for the sub-dimensions. In the present study, Cronbach's alpha was 0.89 for the total scale, and between 0.71 and 0.85 for the sub-dimensions.

Ethical Considerations

Ethical approval for the study was obtained from the local Ethics Committee of Kilis 7 Aralık University. Written informed consent was obtained from each participant. At the beginning of the interview, participants were informed that their participation was voluntary, that they could withdraw from the study at any time, and that their personal information would remain confidential. There was no conflict of interest between the researchers and the stoma patients. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. The STROBE reporting guidelines were followed in conducting and reporting the study.

Statistical Analysis

Normality distributions of the scales and their sub-dimensions were examined using skewness and kurtosis analyses. For non-normally distributed data, the Mann–Whitney U test was used for comparisons between two groups, and the Kruskal–Wallis test was applied for comparisons among three or more groups. For normally distributed data, ANOVA and Student's t-test were performed. To assess the relationships between scales, Pearson correlation analysis was used for normally distributed data, whereas Spearman correlation analysis was used for non-normally distributed data. A significance level of $P<0.05$ was accepted.

RESULTS

The demographic and stoma-related characteristics of the participants are presented in Table 1. Of the patients included in the study, 53.8% (n=112) were male, with a mean age of 54.38±11.36 years; 95.2% (n=198) were married, 56.3% (n=117) were primary school graduates, 36.5% (n=76) had three to four children, and 53.4% (n=111) reported an income equal to their expenses. A total of 63.9% (n=133) had a chronic disease, 63.5% (n=132) were diagnosed with colon cancer, 57.7% (n=120) had a colostomy, and 81.7% (n=170) had their stoma care performed by a family member. All patients received education regarding stoma care, and 55.2% (n=115) received this education from nurses (Table 1).

Table 2 presents the comparison of the Stoma Quality of Life Scale and its sub-dimensions with demographic variables. The analysis revealed a significant difference between the number of children and both the sexuality and body image subscale and the total quality of life score. Patients with three to four children had higher sexuality and body image scores compared to those with fewer or more children ($F = 3.881$, $P=0.022$). Similarly, patients with three to four children had higher total quality of life scores than the other groups ($F = 3.699$, $P=0.026$). In post hoc analyses performed for variables where significance was found in triple comparisons, it was determined that there was a significant difference in terms of sexuality and body image sub-dimension and total quality of life scores between the 0–2 children (1) and 3–4 children (2) groups in the number of children variable. It was also found that patients diagnosed with rectal cancer had higher sexuality and body image scores ($t = -2.902$, $P=0.004$) and higher overall quality of life scores ($t = -2.264$, $P=0.025$) compared to those diagnosed with colon cancer (Table 2).

The demographic data of stoma patients were compared with the Patient Learning Needs Scale and its sub-dimensions. As a result of this analysis, no

TABLE 1. Socio-Demographic Characteristics of the Participants

Sociodemographic characteristics	Data
Age (years)	54.38±11.36
18-49	71 (34.1)
50-61	71 (34.1)
62-76	66 (31.8)
Gender	
Female	96 (46.2)
Male	112 (53.8)
Marital status	
Married	198 (95.2)
Single	10 (4.8)
Education status	
Primary School	117 (56.3)
Middle school and above	91 (43.7)
Number of children	
0-2 children	63 (30.3)
3-4 children	76 (36.5)
≥5 children	69 (33.2)
Income status	
Income less than expenses	97 (46.6)
Income equal to expenses	111 (53.4)
Presence of chronic disease	
Yes	133 (63.9)
No	75 (36.1)
Medical diagnosis	
Colon cancer	132 (63.5)
Rectal cancer	76 (36.5)
Type of stoma	
Colostomy	120 (57.7)
Ileostomy	88 (42.3)
Who performs stoma care?	
Self	38 (18.3)
Family member	170 (81.7)
Have you received stoma care education?	
Yes	208 (100.0)
No	0 (0.0)
Who provided your stoma care education?	
Nurse	115 (55.2)
Physician	49 (23.6)
Medical supplier	44 (22.2)

Data are shown as mean±standard deviation or n (%) where appropriate.

significant findings were observed between the demographic data and the patient learning needs and its sub-dimensions. The P values obtained in these comparisons ranged between 0.212 and 0.953, and all were above the 0.05 significance level.

As a result of the correlation analysis, no significant relationship was found between the Stoma Quality of Life Scale and its sub-dimensions and the sub-dimensions of the Patient Learning Needs Scale (r=0.042, P=0.549). This finding indicates that the characteristics assessed by the two scales generally represent independent constructs.

DISCUSSION

The results of this study, which was conducted to evaluate the relationship between the learning needs and quality of life of stoma patients, showed no significant relationship between the two scales. However, it was observed that some demographic variables may have different levels of influence on quality of life.

The lack of a significant relationship between learning needs and quality of life constitutes a primary finding of this study and requires careful consideration. Learning needs continue throughout life. However, individuals must have the capacity and competence to recognize, express, and request information regarding their health-related deficiencies [24]. Furthermore, this capacity is closely related to the individual's role and responsibilities within the family, their education level, and economic status. In addition, the individual's emotional state can also affect learning needs. In particular, patients with high anxiety levels may experience negative effects on their attention, focus, and processing capacity for information acquisition. High anxiety levels negatively impact the patient's ability to acquire, identify, and process information that may be necessary after discharge [25]. Additionally, while determining learning needs represents the current situation, the fact that quality of life is influenced by previous experiences and shaped by long-term adjustment processes may be another possible explanation for the lack of a direct relationship between these two variables.

A significant difference was found between the

TABLE 2. Distribution and Comparison of the Sub-Dimension and Total Scores of the Stoma Quality of Life Scale According to the Demographic Characteristics of the Patients with a Stoma

Sociodemographic characteristics	n	%	Satisfaction Mean±SD	Work and social life Mean Rank	Sexuality and body image Mean±SD	Stoma function Mean±SD	Total scale score Mean±SD
Gender							
Female	96	46.2	44.9±1.21	106.33	28.43±18.64	28.29±15.54	21.30±10.90
Male	112	53.8	46.8±0.99	102.93	28.52±16.06	30.35±16.33	21.70±9.57
			T:-1.256	U:5200	T:-0.037	T:-0.926	T:-0.285
			P=0.211	P=0.684	P=0.970	P=0.355	P=0.776
Age (years)							
18-49	71	34.1	91.69±24.619	110.69	31.18±16.83	31.37±12.98	22.74±9.33
50-61	71	34.1	92.81±20.92	104.30	30.59±17.07	35.46±13.53	22.61±10.60
62-76	66	31.88	91.36±20.74	98.05	26.84±14.89	33.27±14.60	20.04±9.94
			F:0.082	X^2: 1.522	F: 1.402	F: 1.583	F: 1.575
			P=0.921	P=0.467	P=0.249	P=0.208	P=0.210
Marital status							
Married	198	95.2	91.66±21.88	103.29	29.40±16.37	33.21±13.74	21.75±9.95
Single	10	4.8	98.00±26.88	128.50	33.60±16.28	36.46±13.93	23.55±11.25
			t:-0.884	U:750	t:-0.791	T:-0.729	T:-0.555
			P=0.378	P=0.194	P=0.430	P=0.467	P=0.580
Education status[a]							
Primary School	117	56.3	97.00±27.35	68.70	31.50±18.85	35.00±11.60	22.84±11.50
Middle school and above	91	43.7	90.77±20.67	75.98	30.53±17.61	33.17±14.89	22.02±10.51
			T:1.196	U:1164	T:0.226	T: 0.523	T: 0.321
			P=0.234	P=0.481	P= 0.822	P=0.602	P=0.749
Number of children							
1: 0-2 children	63	30.3	91.74±22.68	104.04	25.33±10.19	31.51±12.57	20.57±7.74
2: 3-4 children	76	36.5	93.02±21.91	114.92	32.98±18.53	34.92±14.43	24.29±10.72
3: ≥5 children	69	33.2	91.01±22.03	93.44	29.78±17.64	33.36±13.96	20.30±10.59
			F:0.153	X^2: 4.653	F:3.881	F: 1.064	F: 3.699
			P=0.858	P=0.098	**P=0.022**	P=0.347	**P=0.026**
Incomes status							
< Expenses	47	22.5	93.82±17.51	106.47	32.44±19.19	32.63±13.38	23.01±9.99
= Expenses	153	77.5	91.04±23.59	98.67	29.26±15.47	33.98±13.74	21.68±10.06
			T:0.748	U:3315	T:1.037	T:-0.594	T:0.789
			P=0.456	P=0.417	P=0.304	P=0.553	P=0.431
Presence of chronic disease							
Yes	133	64.0	91.50±19.24	103.86	29.42±17.17	32.03±13.38	21.22±9.88
No	75	36.0	92.80±26.53	105.63	29.93±14.91	35.75±14.13	22.93±10.17
			T:-0.371	U:4902	T: -0.216	T: -1.886	T: -1.184
			P=0.711	P=0.838	P=0.829	P=0.061	P=0.238
Medical diagnosis							
Colon cancer	132	63.5	91.43±23.41	102.58	27.05±15.09	32.90±14.75	20.66±10.19
Rectal cancer	76	36.5	92.89±19.71	107.84	34.03±17.59	34.18±11.83	23.89±9.36
			T: -0.456	U: 4762	T: -2.902	T: -0.686	T:-2.264
			P=0.649	P=0.542	**P=0.004**	P=0.493	**P=0.025**
Type of stoma							
Colostomy	120	57.6	91.08±23.03	103.73	29.06±16.20	33.19±13.48	21.25±10.08
Ileostomy	88	42.4	93.18±20.81	105.55	30.34±16.63	33.60±14.16	22.65±9.88
			T: -0.676	U: 5188	T:0.554	T:-0.213	T:0.999
			P=0.500	P=0.829	P=0.580	P=0.832	P=0.319
Who performs stoma care?							
Self	38	18.3	91.31±23.72	114.32	31.71±14.71	34.19±15.23	22.50±10.32
Family member	170	81.7	92.11±21.78	102.31	29.13±16.70	33.18±13.42	21.69±9.95
			T:-0.202	U: -1.117	T: 0.877	T: 0.406	T:0.451
			P=0.840	P=0.264	P=0.382	P= 0.685	P=0.652
Who provided stoma care education?							
Nurse	115	55.2	92.86±20.67	109.02	30.76±16.56	32.71±13.88	22.34±9.82
Physician	49	23.6	90.40±24.99	95.29	29.06±17.13	31.88±14.04	20.43±10.38
Medical supplier	44	22.2	91.36±22.67	102.94	27.18±14.94	36.75±12.74	22.11±10.12
			F: 0.232	X^2: 1.844	F: 0.797	F: 1.767	F: 0.644
			P=0.793	P=0.398	P=0.452	P=0.173	P=0.526

Data are shown as mean±standard deviation or n (%) where appropriate. SD, standard deviation.
T, Independent samples t-test; F, One-way ANOVA; U, Mann–Whitney U test; X^2, Kruskal–Wallis test.
Statistically significant P-values are shown in bold.

number of children and both the sexuality and body image subscale and total quality of life scores of the patients with a stoma. Individuals with stronger social support systems may adapt more easily to life after stoma surgery [26]. It has been reported that the support provided by family caregivers plays an important role in improving the quality of life of patients with a stoma [27]. Individuals who have three to four children may have higher sexuality and body image subscale scores and overall quality of life scores compared to others due to their relatively stronger levels of social support. In this group, emotional interaction between parents and children continues, while the decrease in the children's physical dependence may allow family members to maintain their social roles more balancedly. In addition to this possible interpretation, it should be noted that the result may be coincidental. The ages of the children were not questioned, and the sample distribution was limited, which should not be overlooked. In future studies, questioning the number and age ranges of children and asking open-ended questions about parent–child relationships may help fill this gap in the literature.

Individuals with rectal cancer were found to have significantly higher sexuality/body image subscale scores and total stoma quality of life scores compared to those with colon cancer [11, 28]. Although studies have reported a significant difference between urostomy and ileostomy [29]; no study has been found comparing rectal cancer with sexuality/body image and quality of life. Sexuality and body image, together with quality of life, represent the social aspect of an individual [26, 30]. The social needs of individuals with a stoma (such as acceptance by society and their partner, and a sense of belonging) come after physiological needs according to Maslow's hierarchy of needs [31]. When evaluated from this perspective, it is thought that patients with rectal cancer experience fewer physiological problems due to the proximity of their stoma to the anus and the fact that bowel functions (such as nutrient, vitamin, and mineral absorption) are largely completed, and therefore they may focus more on social and psychological aspects. The higher body image, sexuality, and quality of life scores of patients with rectal cancer compared to those with colon cancer may be explained by the support of physiological well-being for social satisfaction and body image processes.

From this perspective, it may be recommended that nurses focus more on body image, sexuality, and quality of life when planning nursing care for patients with rectal cancer.

Strengths and Limitations

Some significant results were obtained between the demographic data and the quality of life of the patients. However, to discuss the reasons for these significant differences in more depth, the scope of demographic variables needs to be expanded. In particular, more detailed information is needed regarding the ages of the children and the roles within the family. Additionally, since other social support systems apart from the family were not questioned in detail, the explanatory power of these significant findings is limited. The fact that the study was conducted in a single center and the sample size was limited also restricts the generalizability of the results.

The sample size, which includes different diagnostic groups and stoma types, increases the applicability of the results to clinical practice. Furthermore, the study's strengths include its provision of practical recommendations for nursing practices related to patient education and discharge planning. In addition, adherence to STROBE guidelines strengthened the methodological rigor of the study.

CONCLUSION

This study aimed to examine whether there is a relationship between the learning needs and quality of life of stoma patients, and no significant relationship was found between these variables in this sample. However, some significant differences were observed between demographic characteristics and quality of life and its sub-dimensions. A significant difference was found in the sexuality and body image subscale and total quality of life scores between patients with rectal cancer and those with three to four children. These findings should be evaluated within a descriptive framework, not with the aim of establishing a causal relationship, due to the cross-sectional nature of the study. From this perspective, it can be suggested that nurses focus more on body image, sexuality, and quality of life when planning nursing care for patients with rectal cancer. In

addition, it is recommended that the social support systems of stoma patients be evaluated and the role distribution within the family be determined, as these will be important for future studies.

Ethics Approval and Consent to Participate

This study was approved by the Kilis 7 Aralık University Ethics Committee (Decision No: 2023/16, 05; date: 31.08.2023). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Written informed consent was obtained from each participant.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: GY, İD; Study Design: GY; Supervision: GY; Funding: N/A; Materials: GY, İD; Data Collection and/or Processing: İD; Statistical Analysis and/or Data Interpretation: GY; Literature Review: GY; Manuscript Preparation: GY; and Critical Review: GY.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

We would like to thank all the patients who participated in the study and the healthcare professionals who supported us throughout the data collection process.

Generative Artificial Intelligence Statement

After the completion of the study, generative AI tools were used only for language editing and translation purposes. All data analysis, interpretation, and scientific conclusions are solely the responsibility of the authors. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Marinova P, Marinova R. Patient-centred stoma care support: colostomy patients. Br J Community Nurs. 2024;29(10):494-502. doi: 10.12968/bjcn.2024.0088.
2. Forest-Lalande L. Best Practice Guidelines for Ostomy Care in Neonates, Children, and Adolescents: An Executive Summary. J Wound Ostomy Continence Nurs. 2023;50(5):381-385. doi: 10.1097/WON.0000000000001001.
3. Roveron G, Barbierato M, Rizzo G, et al.; Multidisciplinary Italian Study Group for STOmas (MISSTO). Italian Guidelines for the Nursing Management of Enteral and Urinary Stomas in Adults: An Executive Summary. J Wound Ostomy Continence Nurs. 2021;48(2):137-147. doi: 10.1097/WON.0000000000000745.
4. Marcomini I, Iovino P, Rasero L, Manara DF, Vellone E, Villa G. Self-Care and Quality of Life of Ostomy Patients: A Structural Equation Modeling Analysis. Nurs Rep. 2024;14(4):3417-3426. doi: 10.3390/nursrep14040247.
5. Geng Z, Howell D, Xu H, Yuan C. Quality of Life in Chinese Persons Living With an Ostomy: A Multisite Cross-sectional Study. J Wound Ostomy Continence Nurs. 2017;44(3):249-256. doi: 10.1097/WON.0000000000000323.
6. Liu Y, Wang L, Zhu L. The impact of stoma management education on the self-care abilities of individuals with an intestinal stoma. Br J Nurs. 2023;32(6):S28-S33. doi: 10.12968/bjon.2023.32.6.S28.
7. Chrobak-Bień J, Marciniak A, Kozicka I, et al. Quality of Life in Patients over Age 65 after Intestinal Ostomy Creation as Treatment of Large Intestine Disease. Int J Environ Res Public Health. 2023;20(3):1749. doi: 10.3390/ijerph20031749.
8. Alenezi A, Livesay K, McGrath I, Kimpton A. Ostomy-related problems and their impact on quality of life of Saudi ostomate patients: A mixed-methods study. J Clin Nurs. 2023;32(13-14):3707-3719. doi: 10.1111/jocn.16466.
9. Liu H, Zhu X, Yu J, et al. The Quality of Life of Patients with Colorectal Cancer and a Stoma in China: A Quantitative Cross-sectional Study. Adv Skin Wound Care. 2021;34(6):302-307. doi: 10.1097/01.ASW.0000744348.32773.b9.
10. Ketterer SN, Leach MJ, Fraser C. Factors Associated With

Quality of Life Among People Living With a Stoma in Nonmetropolitan Areas. Nurs Res. 2021;70(4):281-288. doi: 10.1097/NNR.0000000000000511.
11. Harputlu D, Öztürk FÖ, Aydın D, Akyol C, Tezel A. Effect of Laughter Yoga on Sleep and Quality of Life in Individuals With Fecal Ostomies: A Randomized Controlled Trial. J Wound Ostomy Continence Nurs. 2023;50(4):307-312. doi: 10.1097/WON.0000000000000988.
12. Harputlu D, Talaz D, Cenan D, Karadağ S, Öğce F. Stomalı Bireylerde Uyku Kalitesinin İncelenmesi [Investigation of Sleep Quality in Individuals with Stoma]. Dokuz Eylül Üniversitesi Hemşirelik Fakültesi Elektronik Dergisi. 2020;13(2):59-66. [Article in Turkish]
13. Gül A, Aygin D. Stomalı Hastalarda Dini Ritüeller [Religious Rituals in Patients with Stoma]. İstanbul Gelişim Üniversitesi Sağlık Bilimleri Dergisi. 2021;14:389-402. doi: 10.38079/igusabder.937111. [Article in Turkish]
14. Kaya Ö, Aygin D. Stomanın cinsel yaşam üzerindeki etkileri [Effects of stoma on sexual life]. Androloji Bülteni. 2020;22(3):194-198. doi: 10.24898/tandro.2020.68984. [Article in Turkish]
15. Ongün P, Yılmaz NB, Kırtıl Yİ. Stomalı Bireylerde Anksiyete, Cinsellik ve Yaşam Kalitesi: Sistematik Derleme [Anxiety, Sexuality and Quality of Life of Individuals with Stoma: A Systematic Review]. Türk Hemşireler Derneği Dergisi. 2021;2(2):41-52. [Article in Turkish]
16. Yaşar Z, Üstündağ H. Stomalı Hastaların Yaşam Kalitesi ve Öz Yeterlilik Düzeylerinin Belirlenmesi [The Determination of Self-Efficacy and Quality of Life in Patients with Stoma]. Celal Bayar Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi. 2020;8(1):107-115. doi: 10.34087/cbusbed.813080. [Article in Turkish]
17. Üstündağ H, Demir N, Zengin N, Gül A. Stomalı Hastalarda Beden İmajı ve Benlik Saygısı [Body Image and Self-Esteem in Patients with Stoma]. Turkiye Klinikleri J Med Sci. 2007;27(4):522-527. [Article in Turkish]
18. Althubaiti A. Sample size determination: A practical guide for health researchers. J Gen Fam Med. 2022;24(2):72-78. doi: 10.1002/jgf2.600.
19. Toprak TN, Tuna A. Ostomili Hastaların Bilgi Gereksinimlerinin Ve Ostomiye Uyumlarının Belirlenmesi [Determination of Informatıon Requirements and Complince of Ostomy Patients]. Paramedik ve Acil Sağlık Hizmetleri Dergisi. 2022;3(1):46-62. doi: 10.54862/pashid.909588. [Article in Turkish]
20. Baxter NN, Novotny PJ, Jacobson T, Maidl LJ, Sloan J, Young-Fadok TM. A stoma quality of life scale. Dis Colon Rectum. 2006;49(2):205-212. doi: 10.1007/s10350-005-0275-6.
21. Karadağ A, Öztürk D, Çelik B. Ostomili Bireylere Yönelik Yaşam Kalitesi Ölçeğinin Türkçeye Uyarlanması [Adaptation of Quality Life Scale for the Ostomy Patients into Turkish]. Kolon Rektum Hast Derg. 2011;21(4):173-181. doi: 10.5505/KRHD.2011.66588. [Article in Turkish]
22. Bubela N, Galloway S, McCay E, et al. The Patient Learning Needs Scale: reliability and validity. J Adv Nurs. 1990;15(10):1181-1187. doi: 10.1111/j.1365-2648.1990.tb01711.x.
23. Çatal E, Dicle A. Hasta Öğrenim Gereksinimleri Ölçeği'nin Türkiye'de Geçerlik ve Güvenirlik Çalışması [A Validity and Reliability Study of the Patient Learning Needs Scale in Türkiye]. Dokuz Eylül Üniversitesi Hemşirelik Yüksekokulu Elektronik Dergisi. 2008;1(1):19-32. [Article in Turkish]
24. Pehlïvan Z, Güner SG, Nural N. Kemoterapiye Bağlı Gelişen Bulantı Kusma Yönetimine İlişkin Hemşirelik Alanında Yapılmış Deneysel Tezlerin İncelenmesi [Investigation of Experimental Theses in Nursing on the Management of Nausea and Vomiting Related to Chemotherapy]. Ordu Üniversitesi Hemşirelik Çalışmaları Dergisi. 2023;6(1):193-204. doi: 10.38108/ouhcd.1070519. [Article in Turkish]
25. Kazanç Ş, Koçyiğit H, Guler N. Mastektomi Hastalarının Öğrenim Gereksinimleri ve Kaygı Düzeyleri Arasındaki İlişki [The Relationship between Learning Needs and Anxiety Levels of Patients with Mastectomy]. Ege Üniversitesi Hemşirelik Fakültesi Dergisi. 2023;39(2):193-202. doi: 10.53490/egehemsire.1149013. [Article in Turkish]
26. Davis D, Ramamoorthy L, Pottakkat B. Impact of stoma on lifestyle and health-related quality of life in patients living with stoma: A cross-sectional study. J Educ Health Promot. 2020;9:328. doi: 10.4103/jehp.jehp_256_20.
27. Golpazir-Sorkheh A, Ghaderi T, Mahmoudi S, Moradi K, Jalali A. Family-Centered Interventions and Quality of Life of Clients with Ostomy. Nurs Res Pract. 2022;2022:9426560. doi: 10.1155/2022/9426560.
28. Werther C. Predictors of Health-Related Quality of Life in Younger Persons With a Fecal Ostomy and Inflammatory Bowel Disease. J Wound Ostomy Continence Nurs. 2024;51(4):297-302. doi: 10.1097/WON.0000000000001087.
29. Zhu X, Chen Y, Tang X, Chen Y, Liu Y, Guo W, Liu A. Sexual Experiences of Chinese Patients Living With an Ostomy. J Wound Ostomy Continence Nurs. 2017;44(5):469-474. doi: 10.1097/WON.0000000000000357.
30. Muzii B, Di Bello F, Collà Ruvolo C, et al. Stoma Acceptance Mediates Body Image Distress and Mental Health-Related Quality of Life: A Single-Center Study on Radical Cystectomy Patients with Ureterostomy. J Clin Med. 2024;13(24):7682. doi: 10.3390/jcm13247682.
31. Rojas M, Méndez A, Watkins-Fassler K. The hierarchy of needs empirical examination of Maslow's theory and lessons for development. World Development. 2023;165:106185. doi: 10.1016/j.worlddev.2023.106185.

The European Research Journal 2026;12(3):367-375 Original Article

DOI: https://doi.org/10.18621/eurj.1849380 *Rheumatology*

Is Seronegative Rheumatoid Arthritis Far From Being a Benign Subtype of Rheumatoid Arthritis? A Retrospective Comparative Study of Seronegative Rheumatoid Arthritis and Seropositive Rheumatoid Arthritis

Salim Mısırcı[1], Mustafa Çağatay Büyükuysal[2]

[1]Department of Rheumatology, Amasya Sabuncuoğlu Şerefeddin Training and Research Hospital, Amasya, Türkiye; [2]Department of Biostatistics, Zonguldak Bülent Ecevit University, Faculty of Medicine, Zonguldak, Türkiye

ABSTRACT

Objectives: Our study aimed to compare patients with seronegative rheumatoid arthritis (SNRA) and seropositive rheumatoid arthritis (SPRA) to determine whether they represent favourable and unfavourable subtypes of the same disease, or whether SNRA patients follow an aggressive course similar to SPRA patients.

Methods: Patients diagnosed with rheumatoid arthritis (RA) according to the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria (n=334) were evaluated retrospectively. Patients were classified into SNRA (n=108) and SPRA (n=226) groups based on their seropositivity status. The SNRA and SPRA groups were compared regarding clinical, laboratory, imaging, and medical treatment characteristics.

Results: Gender, age at diagnosis, comorbidities, acute phase reactants, medical therapies (excluding leflunomide), anaemia, cancer, and mortality did not differ significantly between the two groups (P>0.05). There was also no significant difference between the groups in the total number of joints where erosion was detected (SNRA: n=9 (8.3%), SPRA: n=25 (11.1%), P=0.563). Interstitial lung disease (ILD) was the most common extra-articular involvement and was detected only in the SPRA group (n=9 (4.0%), P=0.034).

Conclusions: It should be noted that SNRA patients, like SPRA patients, may experience an aggressive disease course and develop erosive joint damage. Seropositivity is important regarding ILD, and necessary follow-up and treatment plans should be implemented without delay for both SNRA and SPRA patients.

Keywords: Interstitial Lung Disease, Joint Erosion, Seronegative Rheumatoid Arthritis, Seropositive Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a chronic autoimmune inflammatory joint disease that can also present with numerous extra-articular symptoms [1, 2]. Lung involvement is the most commonly detected extra-articular manifestation [3]. The pathophysiology of RA is complex, with genetic and environmental factors as the main contributors. Its prevalence in the white race is 0.5-1%, and it is more common in women. It usually appears in middle age in women, but may begin later in men. In untreated patients,

Received: December 26, 2025 **Accepted:** February 12, 2026 **Available Online:** February 13, 2026 **Published:** March 1, 2026

How to cite this article: Mısırcı S, Büyükuysal MÇ. Is Seronegative Rheumatoid Arthritis Far From Being a Benign Subtype of Rheumatoid Arthritis? A Retrospective Comparative Study of Seronegative Rheumatoid Arthritis and Seropositive Rheumatoid Arthritis. Eur Res J. 2026;12(3):367-375. doi: 10.18621/eurj.1849380

Corresponding author: Salim Mısırcı, MD., Phone: +90 358 218 40 00, E-mail: dr.salim-misirci@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

joint damage and other systemic involvement of RA can lead to poor quality of life and disability [1, 2].

RA patients are classified as seropositive RA (SPRA) when anti-cyclic citrullinated peptide (anti-CCP) and/or rheumatoid factor (RF) antibodies are present, while those with both negative RF and anti-CCP antibodies are classified as seronegative RA (SNRA) [4]. Although it is stated that SNRA patients constitute approximately 20–40% of RA patients, it has been reported that the incidence has increased in recent years [5, 6]. It is thought that SPRA, which is more common, has a more severe course than SNRA in terms of erosive joint damage, extra-articular involvement, comorbidities, and mortality [7]. Therefore, SNRA patients are considered a well-defined subtype of RA patients. However, recent studies have emphasised that, due to the scoring of anti-CCP and RF antibodies in the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria [8], SNRA patients require more joint involvement and higher levels of inflammatory markers to meet these criteria [9].

Early diagnosis and treatment planning in RA patients can reduce disability by preventing the missed opportunity window for both SNRA and SPRA. Furthermore, approximately half of SNRA patients may later become seropositive. The absence of antibodies, delayed diagnosis of SNRA in the early stages, and delayed initiation of necessary treatment may contribute to an aggressive course in SNRA patients, similar to that in SPRA patients [10].

Due to these conflicting results, the clinical course and disease severity of SPRA and SNRA patients remain areas that require further research.

This study aimed to compare SNRA and SPRA patients in terms of clinical and demographic characteristics, erosive joint involvement, extra-articular involvement, and medical treatments used, to assess whether they represent good and bad subtypes of the same disease, or whether SNRA patients follow an aggressive course similar to SPRA patients.

METHODS

Patients over 18 years of age who met the 2010 ACR/EULAR classification criteria for RA [8] were retrospectively screened from the records of a tertiary teaching and research hospital (Ethics Committee Approval: 26 November 2025; protocol code: 2025/18-183). Patients were required to have been followed up for at least six months at the centre, and the treatments they received at previous centres had to be recorded in the system. Patients who could not be accessed retrospectively from the hospital record system for clinical, laboratory, imaging, and medical treatment records, as well as those with other concomitant inflammatory rheumatic diseases, were excluded from the study.

A total of 533 RA patients were assessed for eligibility. Of these, 334 were deemed eligible to participate in the study and were assigned to the SNRA (n=108) and SPRA (n=226) groups according to their seropositivity status. Patients who tested positive for either or both anti-CCP and RF antibodies were included in the SPRA group. Patients with results ≥ 3 times the accepted cut-off value for RF or anti-CCP were considered high positive. Patients were further classified as double high positive (both RF and anti-CCP values ≥ 3 times the cut-off) or single high positive (only one of RF or anti-CCP values ≥ 3 times the cut-off).

The patients' demographic data, joint involvement pattern (monoarthritis, oligoarthritis, polyarthritis) based on the number of joints involved at diagnosis, specific joints involved (large or small), and disease duration (months) were recorded. The presence of bone erosion in the joint regions during follow-up was determined by a specialist rheumatologist based on the evaluation of the most recent X-ray images.

The presence of extra-articular involvement, comorbidities, malignancy, acute phase reactants (C-reactive protein (CRP), erythrocyte sedimentation rate (ESR)) at the last check-up, anaemia, and exitus were assessed and recorded. Interstitial lung disease (ILD), as a component of extra-articular involvement, was assessed using high-resolution computed tomography (HRCT).

Conventional disease-modifying antirheumatic drugs (csDMARDs), biologic disease-modifying antirheumatic drugs (bDMARDs), and targeted synthetic disease-modifying antirheumatic drugs (tsDMARDs) that were actively or previously used during the follow-up period after diagnosis were assessed and recorded. For the study, glucocorticoids

(GCs) and their doses (prednisolone ≤5 mg equivalent or prednisolone >5 mg equivalent) in active use at the time of patient assessment were recorded.

Statistical Analysis

Statistical analyses were conducted using the SPSS 29.0 software package. Descriptive statistics for quantitative variables included mean, standard deviation, median, minimum, and maximum values; qualitative variables were presented as frequency and percentage. The normality of quantitative variables was assessed with the Shapiro-Wilk test, and the Mann-Whitney U test was used to compare two independent groups. Pearson's chi-square and Fisher's exact chi-square tests were used for intergroup comparisons of qualitative variables. In all statistical analyses in the study, results with a P-value below 0.05 were considered statistically significant.

RESULTS

The study included 334 RA patients, most of whom were SPRA patients (SNRA: n=108, 32.3%; SPRA: n=226, 67.7%). The mean age in the SNRA group was 57.52±14.13 years, and in the SPRA group, 60.20 ± 11.50 years; there was no significant difference between the groups (P=0.211). The mean age at diagnosis was 47.01±15.33 years in the SNRA group and 50.58±12.30 years in the SPRA group, with no significant difference between the groups (P=0.084). The proportion of female patients was higher in both groups (SNRA: 75.9%; SPRA: 78.3%) (P=0.624).
When patients in the SPRA group were examined for seropositivity, RF positivity was found in 184 (81.4%) patients, while anti-CCP positivity was found in 191 (84.5%) patients. When evaluated in terms of titre, 124 (54.9%) patients had high RF positivity, 155 (68.6%) patients had high anti-CCP positivity, 104 (46.0%) patients had high positivity for both RF and anti-CCP, and 70 (31.0%) patients had high positivity for either RF or anti-CCP.

When joint involvement at diagnosis was assessed, almost all patients in both groups had at least one small joint affected (SNRA: 99.1%, SPRA: 98.7%). Although polyarthritis was the most common pattern of joint involvement in both groups, it was more prevalent in the SPRA group (SNRA: 63.0%, SPRA: 88.9%, P<0.001). The rates of oligoarthritis (SNRA: 33.3%, SPRA: 9.7%, P<0.001) and large joint involvement (SNRA: 57.4%, SPRA: 36.3%) were higher in the SNRA group. There was no significant difference between the groups in the total joint area affected by erosion (SNRA: n=9 (8.3%), SPRA: n=25 (11.1%), P=0.563), and erosion was most frequently detected in the metacarpophalangeal (MCP) joint region (SNRA: n=3 (2.8%), SPRA:n=16 (7.1%), P=0.182). The proximal interphalangeal (PIP) joint and wrist were the second most common sites of erosion, and the rate of erosion in the PIP region was statistically significantly higher in the SPRA group than in the SNRA group (SNRA: n=1 (0.9%), SPRA: n=14 (6.2%), P=0.044).

When assessed for extra-articular involvement, ILD was the most common and was observed only in the SPRA group (n=9, p=0.034). No extra-articular involvement was detected except for ILD and pleural effusion (SNRA: n=1 (0.9%), SPRA: n=1 (0.4%), P=0.543). Acute phase reactants (ESR, CRP), comorbidities, anaemia, cancer, and exitus did not differ significantly between the two groups (P>0.05 for all). Table 1 presents the clinical, demographic, and laboratory characteristics of SNRA and SPRA patients.

There was no statistically significant difference between the two groups in the rates of csDMARDs (excluding leflunomide), bDMARDs, tsDMARDs, and mycophenolate mofetil previously or currently used by patients for RA treatment. Similarly, there was no significant difference in the rates of patients actively using GCs. The comparison of medical treatments used in SNRA and SPRA patients is shown in Table 2.

DISCUSSION

When comparing SNRA and SPRA patients in our study, we found no significant differences in age, age at diagnosis, gender, comorbidities, anaemia, malignancy, exitus, or rates of bDMARDs and tsDMARDs used for treatment. Again, there was no significant difference in the proportion of patients with total erosion in the affected joint areas, indicating that SNRA patients can also have an aggressive course

TABLE 1. Comparison of Clinical, Demographic, and Laboratory Characteristics of Seronegative Rheumatoid Arthritis (SNRA) and Seropositive Rheumatoid Arthritis (SPRA) Patients

	SNRA n=108 (32.3%)	SPRA n=226 (67.7%)	P-value
Age (years)	57.52±14.13	60.20±11.50	0.211 [m]
Age at Diagnosis	47.01±15.33	50.58±12.30	0.084 [m]
Gender			
Female	82 (75.9%)	177 (78.3%)	0.624 [χ2]
Male	26 (24.1%)	49 (21.7%)	
Duration of Disease (month)	108 (12, 456)	96 (6, 480)	0.375 [m]
Pattern and number of joint involvement at diagnosis			
Monoarthritis	4 (3.7%)	3 (1.3%)	0.219 [f]
Oligoarthritis	36 (33.3%)	22 (9.7%)	**<0.001** [χ2]
Polyarthritis	68 (63.0%)	201 (88.9%)	**<0.001** [χ2]
Large joint	62 (57.4%)	82 (36.3%)	**<0.001** [χ2]
Small joint	107 (99.1%)	223 (98.7%)	1.000 [f]
ESR (mm/saat)	30 (2-106)	29 (3-95)	0.383 [m]
CRP (mg/L)	5 (0.1-37)	4 (0.19-102)	0.177 [m]
Erosion	9 (8.3%)	25 (11.1%)	0.563 [χ2]
MCP	3 (2.8%)	16 (7.1%)	0.182 [χ2]
PIP	1 (0.9%)	14 (6.2%)	**0.044** [f]
Wrist	4 (3.7%)	11 (4.9%)	0.781 [f]
Elbow	2 (1.9%)	0 (0.0%)	0.104 [f]
Hip	1 (0.9%)	1 (0.4%)	0.543 [f]
Knee	3 (2.8%)	4 (1.8%)	0.686 [f]
MTP	0 (0.0%)	4 (1.8%)	0.309 [f]
Ankle	0 (0.0%)	1 (0.4%)	1.000 [f]
Carpal fusion	0 (0.0%)	1 (0.4%)	1.000 [f]
Extra-articular involvement	1 (0.9%)	10 (4.4%)	0.112 [f]
ILD	0 (0.0%)	9 (4.0%)	**0.034** [f]
Pleural effusion	1 (0.9%)	1 (0.4%)	0.543 [f]
Comorbidity	58 (53.7%)	139 (61.5%)	0.175 [χ2]
Diabetes mellitus	23 (21.3%)	58 (25.7%)	0.384 [χ2]
Hypertension	39 (36.1%)	95 (42.0%)	0.301 [χ2]
Coronary artery disease	5 (4.6%)	17 (7.5%)	0.319 [χ2]
Chronic kidney disease	3 (2.8%)	8 (3.5%)	1.000 [f]
Hyperlipidaemia	7 (6.5%)	18 (8.0%)	0.824 [f]
Congestive heart failure	2 (1.9%)	6 (2.7%)	1.000 [f]
COPD	2 (1.9%)	7 (3.1%)	0.724 [f]
Hypothyroidism	9 (8.3%)	15 (6.6%)	0.651 [f]
Asthma	2 (1.9%)	12 (5.3%)	0.241 [f]
Malignancy	1 (0.9%)	6 (2.7%)	0.435 [f]
Anemia	19 (17.6%)	34 (15.0%)	0.551 [χ2]
Exitus	0 (0.0%)	2 (0.9%)	1.000 [f]

Data are shown as mean±standard deviation or median (minimum-maximum) or n (%) where appropriate. COPD, Chronic obstructive pulmonary disease; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ILD, Interstitial lung disease; MCP, Metacarpophalangeal; MTP, Metatarsophalangeal; PIP, Proximal Interphalangeal. [f]Fisher chi-square test; [m]Mann-Whitney U test; [χ2]Pearson chi-square test. Statistically significant P-values are shown in bold.

TABLE 2. Comparison of Medical Treatments Used in Seronegative Rheumatoid Arthritis (SNRA) and Seropositive Rheumatoid Arthritis (SPRA) Patients

	SNRA n=108 (32.3%)	SPRA n=226 (67.7%)	P-value
bDMARDs, n (%)	36 (33.3)	86 (38.1)	0.402 $^{\chi 2}$
Infliximab	4 (3.7)	5 (2.2)	0.478 ^{f}
Abatacept	1 (0.9)	4 (1.8)	1.000 ^{f}
Adalimumab	19 (17.6)	41 (18.1)	0.903 $^{\chi 2}$
Certolizumab	9 (8.3)	11 (4.9)	0.212 $^{\chi 2}$
Etanercept	10 (9.3)	23 (10.2)	0.793 $^{\chi 2}$
Golimumab	3 (2.8)	6 (2.7)	1.000 ^{f}
Rituximab	4 (3.7)	18 (8.0)	0.142 $^{\chi 2}$
Tocilizumab	3 (2.8)	12 (5.3)	0.402 ^{f}
tsDMARDs, n (%)	8 (7.4)	21 (9.3)	0.567 $^{\chi 2}$
Tofacitinib	5 (4.6)	16 (7.1)	0.388 $^{\chi 2}$
Baricitinib	0 (0.0)	1 (0.4)	1.000 ^{f}
Upadacitinib	3 (2.8)	5 (2.2)	0.717 ^{f}
csDMARDs, n (%)	108 (100.0)	224 (99.1)	1.000 ^{f}
Methotrexate	71 (65.7)	165 (73.0)	0.172 $^{\chi 2}$
Leflunomide	42 (38.9)	126 (55.8)	**0.004** $^{\chi 2}$
Sulfasalazine	24 (22.2)	38 (16.8)	0.234 $^{\chi 2}$
Hydroxychloroquine	57 (52.8)	116 (51.3)	0.804 $^{\chi 2}$
Mycophenolate mofetil, n (%)	0 (0.0)	3 (1.3)	0.554 ^{f}
Glucocorticoids, n (%)	72 (66.7)	148 (65.5)	0.832 $^{\chi 2}$
Prednisolone ≤5 mg equivalent dose	69 (63.9)	141 (61.4)	0.791 $^{\chi 2}$
Prednisolone >5 mg equivalent dose	3 (2.8)	7 (3.1)	1.000 ^{f}

Data are shown as n (%).bDMARDs, biologic disease-modifying antirheumatic drugs; csDMARDs, conventional synthetic disease-modifying antirheumatic drugs; tsDMARDs, targeted synthetic disease-modifying antirheumatic drugs.
^{f}Fisher chi-square test; ^{m}Mann-Whitney U test; $^{\chi 2}$Pearson chi-square test. Statistically significant P-values is shown in bold.

similar to SPRA patients. On the other hand, ILD was detected only in the SPRA group. Furthermore, polyarthritis was the most common pattern of joint involvement in both groups; however, it was more prevalent in the SPRA group. Again, the higher rates of oligoarthritis and large joint involvement in the SNRA group raise the question of whether the diagnoses of some patients in this group will change to a spondyloarthropathy (SpA) or another inflammatory rheumatic disease in the course of the disease.

The genders and mean ages at diagnosis of the patients in our study (SNRA: 47.01±15.33; SPRA: 50.58±12.30) showed that women in the perimenopausal period (SNRA: 75.9%, SPRA: 78.3%) were affected more than men (approximately 3:1) in both groups. Our results were consistent with the literature [1, 11, 12]. Again, the literature states that approximately 20–40% of RA patients are SNRA patients, but that these rates are increasing due to factors such as longer life expectancy and lower smoking rates [5, 6]. In our study, similar to the literature, the SNRA patient rate was 32.3%.

RA, particularly in SPRA patients, typically presents as symmetrical polyarthritis affecting small joints such as the wrist, MCP, PIP, and metatarsophalangeal joints [13]. On the other hand,

involvement of large joints such as the shoulders and knees may also be detected, and patients in this group may initially have higher disease activity than those without large joint involvement [14]. Again, it is generally thought that SPRA patients have a more aggressive course than SNRA patients in many parameters, such as radiographic progression of the joints, extra-articular involvement, comorbidities, and mortality [7]. In a study conducted by Gadeholt *et al*. [15] evaluating the X-ray findings of a total of 56 SNRA and 57 SPRA patients, more erosion and joint space narrowing were detected in SPRA patients. The erosion burden was significantly higher in the SPRA group, particularly in the feet and MCP joints; however, there was no significant difference between the groups in erosion burden in the PIP and wrist joints. Conversely, a study by Choi *et al*. [16] indicated that SNRA patients initially showed a more active disease course. In the ACR/EULAR classification criteria for RA [8], the inclusion of anti-CCP antibodies and RF in the scoring system means that SNRA patients require more inflammatory activity and joint involvement for classification, supporting the view that SNRA patients may also show severe inflammatory findings. In our study, the incidence of polyarthritis was higher in SPRA patients, while oligoarthritis and large joint involvement were more common in the SNRA group. In a study, changes in International Classification of Diseases (ICD) codes were evaluated over a 15-year follow-up period for 9,784 patients with SNRA. It was determined that the ICD codes of 546 of the evaluated patients changed to one of the SpA ICD codes (275 (48.7%) psoriatic arthritis, 245 (43.4%) axial SpA, 44 (7.8%) inflammatory bowel disease) [17]. One reason the rates of oligoarthritis and large joint involvement were higher in patients in the SNRA group in our study may be this situation. Because our study was retrospective and included patients with a short follow-up period, such an assessment was not possible. Clinicians should bear in mind that the diagnoses of SNRA patients may change during the follow-up period, particularly to diseases in the SpA group.

On the other hand, the absence of significant differences between the SNRA and SPRA groups in total patient rates of joint erosions, mortality, and comorbidities contradicts literature indicating that SPRA patients generally have a worse prognosis than SNRA patients. Furthermore, some studies emphasise that SNRA patients may initially have higher inflammatory activity and that radiographic progression is similar during follow-up [9, 16]. In our study, the most common area of erosion was the MCP, and no significant difference was found between the SNRA and SPRA groups. Although the PIP and wrist were also common sites of erosion, only the erosion rate in the PIP joint was statistically significantly higher in the SPRA group than in the SNRA group. The absence of a statistically significant difference in the total joint area affected by erosion suggests that SNRA patients, like SPRA patients, may also experience a poor disease course due to erosive joint damage.

Extra-articular involvement most frequently affects the lungs, which can increase morbidity and mortality. With the increased use of HRCT, ILD has become more prominent than pleural effusion, which was previously considered the most common lung involvement in RA. Clinically significant ILD is detected in approximately 2–10% of RA patients. Seropositivity is known to be an important risk factor for ILD [3, 18]. In a study by Kim *et al*. [19], 52,325 RA patients and 261,625 non-RA controls were evaluated. During an average follow-up of 4.4 years, ILD developed in 3.7% of the RA cohort, compared with 0.5% of the control group. When the RA group was subdivided into SNRA and SPRA and analysed separately, the risk of ILD was higher in the SPRA group. However, the risk of ILD was also higher in the SNRA group than in the control group. In our study, ILD was detected in 9 (2.7%) of 334 RA patients, all of whom were SPRA patients (4% of SPRA patients). Additionally, double (RF and anti-CCP) high positivity was detected in six patients with ILD, while single (RF or anti-CCP) high positivity was found in the other three patients with ILD. Although our patient numbers are small, our findings suggest that the seropositivity status and titre of RA patients should be considered a warning sign for ILD. Based on symptom assessment and physical examination findings, imaging with HRCT should be performed when indicated.

In our study, there was no significant difference between the SNRA and SPRA groups in terms of csDMARDs (except leflunomide), bDMARDs, and tsDMARDs used. Again, although there was no significant difference in terms of GC use, the proportion of patients using prednisolone ≤5 mg

equivalent dose was quite high in both groups. In the two-year data from the ARCTIC study, which compared the clinical course and treatment methods of RA patients classified as SNRA and SPRA according to the ACR/EULAR classification criteria, it was noted that the rates of methotrexate monotherapy, triple therapy, and biological therapy were similar [9]. The study conducted by Choi *et al.* [16] also found no difference between the SNRA and SPRA groups regarding the initial dose of prednisolone, the total number of DMARDs used, the methotrexate dose required to achieve remission, or the proportion of patients using bDMARDs. The results of our study are consistent with this recent literature. Furthermore, the similarity of medical treatments used in both groups may be significant in that it shows clinicians focus on suppressing inflammation rather than patients' seropositive status in their clinical practice. On the other hand, the higher rates of large joint involvement in the SNRA group may have led to lower rates of leflunomide use and higher rates of sulfasalazine use.

Strengths and Limitations

The strongest aspect of this article is that it demonstrates SNRA patients can also experience an aggressive course, contrary to the belief that the clinical course of SPRA patients is more aggressive than that of SNRA patients. The most significant limitation of our study is its design as a single-centre, retrospective study with a small patient cohort. The limited number of cases has not permitted advanced statistical analysis of ILD risk factors, defined as extra-articular involvement. Another limitation is that disease severity has mostly been assessed based on joint involvement and medical treatments, due to the absence of extra-articular involvement other than ILD and pleural effusion. Furthermore, it is stated that the diagnoses of SNRA patients may change during follow-up, particularly to diseases within the SpA group [17]. Again, it is stated that chondrocalcinosis and calcium pyrophosphate deposition disease are more common in SNRA patients than in SPRA patients [20]. Although our study included patients with long follow-up periods, some had follow-up of only six months. Due to the retrospective design and short follow-up period, it was not possible to evaluate potential changes in the diagnoses of SNRA patients. Other limitations include the inability to evaluate the Disease Activity Score (DAS)-28 because it was not documented for every patient due to the retrospective design, and difficulties in adequately assessing erosion in patients with a short follow-up period. Again, due to the retrospective design, the time from symptom onset to diagnosis could not be included in the article. Erosions that may develop as a result of possible diagnostic delays may be misleading.

CONCLUSION

In conclusion, as patients classified as SNRA according to the 2010 RA ACR/EULAR classification criteria may experience an aggressive course with erosive joint damage similar to SPRA patients, the appropriate treatment plan should be implemented without delay. Seropositivity is important in relation to ILD, and necessary follow-up and investigations should be planned, especially for patients with high-titre RF and/or anti-CCP positivity.

Ethics Approval and Consent to Participate

This study was approved by the Amasya University Non-Interventional Clinical Research Ethics Committee (Decision No: 2025/18-183; date: 26.11.2025). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Informed consent was waived because of the retrospective nature of the study and the analysis used anonymous clinical data.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: SM; Study Design: SM; Supervision: SM; Funding: N/A; Materials: SM; Data Collection and/or Processing: SM; Statistical Analysis

and/or Data Interpretation: MÇB; Literature Review: SM; Writer: SM, MÇB; and Critical Review: SM.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Smolen JS, Aletaha D, Barton A, et al. Rheumatoid arthritis. Nat Rev Dis Primers. 2018;4:18001. doi: 10.1038/nrdp.2018.1.
2. Krajewska-Włodarczyk M, Szeląg M, Batko B, et al. Rheumatoid arthritis epidemiology: a nationwide study in Poland. Rheumatol Int. 2024;44(6):1155-1163. doi: 10.1007/s00296-024-05591-8.
3. Bendstrup E, Møller J, Kronborg-White S, Prior TS, Hyldgaard C. Interstitial Lung Disease in Rheumatoid Arthritis Remains a Challenge for Clinicians. J Clin Med. 2019;8(12):2038. doi: 10.3390/jcm8122038.
4. Mo Q, Wang M, Cai S, Zhong J, Dong L. A comparative study on the clinical and magnetic resonance imaging features between seronegative and seropositive rheumatoid arthritis. Clin Exp Rheumatol. 2024;42(1):77-85. doi: 10.55563/clinexprheumatol/9lfvfy.
5. Paroli M, Sirinian MI. When Autoantibodies Are Missing: The Challenge of Seronegative Rheumatoid Arthritis. Antibodies (Basel). 2023;12(4):69. doi: 10.3390/antib12040069.
6. Sokolova MV, Schett G, Steffen U. Autoantibodies in Rheumatoid Arthritis: Historical Background and Novel Findings. Clin Rev Allergy Immunol. 2022;63(2):138-151. doi: 10.1007/s12016-021-08890-1.
7. Angelozzi L, Gillibert A, Brevet P, et al. Comparative analysis of clinical profile, therapeutic management, and clinical prognosis of patients with seropositive or seronegative rheumatoid arthritis following the introduction of a first targeted therapy in a real-life setting. Clin Rheumatol. 2025;44(4):1527-1536. doi: 10.1007/s10067-025-07390-3.
8. Aletaha D, Neogi T, Silman AJ, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010;62(9):2569-2581. doi: 10.1002/art.27584.
9. Nordberg LB, Lillegraven S, Aga AB, et al. Comparing the disease course of patients with seronegative and seropositive rheumatoid arthritis fulfilling the 2010 ACR/EULAR classification criteria in a treat-to-target setting: 2-year data from the ARCTIC trial. RMD Open. 2018;4(2):e000752. doi: 10.1136/rmdopen-2018-000752.
10. Coffey CM, Crowson CS, Myasoedova E, Matteson EL, Davis JM 3rd. Evidence of Diagnostic and Treatment Delay in Seronegative Rheumatoid Arthritis: Missing the Window of Opportunity. Mayo Clin Proc. 2019;94(11):2241-2248. doi: 10.1016/j.mayocp.2019.05.023.
11. Finckh A, Gilbert B, Hodkinson B, et al. Global epidemiology of rheumatoid arthritis. Nat Rev Rheumatol. 2022;18(10):591-602. doi: 10.1038/s41584-022-00827-y.
12. Perera J, Delrosso CA, Nerviani A, Pitzalis C. Clinical Phenotypes, Serological Biomarkers, and Synovial Features Defining Seropositive and Seronegative Rheumatoid Arthritis: A Literature Review. Cells. 2024;13(9):743. doi: 10.3390/cells13090743.
13. Żelnio E, Taljanovic M, Mańczak M, Sudoł-Szopińska I. Hand and Wrist Involvement in Seropositive Rheumatoid Arthritis, Seronegative Rheumatoid Arthritis, and Psoriatic Arthritis-The Value of Classic Radiography. J Clin Med. 2023;12(7):2622. doi: 10.3390/jcm12072622.
14. Shirasugi I, Onishi A, Nishimura K, et al. Association of large joint involvement at the start of biological disease-modifying antirheumatic drugs and Janus kinase inhibitors with disease activity and drug retention in patients with rheumatoid arthritis: The ANSWER cohort study. Int J Rheum Dis. 2024;27(3):e15097. doi: 10.1111/1756-185X.15097.
15. Gadeholt O, Hausotter K, Eberle H, Klink T, Pfeil A. Differing X-ray patterns in seronegative and seropositive rheumatoid arthritis. Clin Rheumatol. 2019;38(9):2403-2410. doi: 10.1007/s10067-019-04602-5.
16. Choi S, Lee KH. Clinical management of seronegative and seropositive rheumatoid arthritis: A comparative study. PLoS One. 2018;13(4):e0195550. doi: 10.1371/journal.pone.0195550.
17. Paalanen K, Puolakka K, Nikiphorou E, Hannonen P, Sokka T. Is seronegative rheumatoid arthritis true rheumatoid arthritis? A nationwide cohort study. Rheumatology (Oxford). 2021;60(5):2391-2395. doi: 10.1093/rheumatology/keaa623.
18. Kadura S, Raghu G. Rheumatoid arthritis-interstitial lung

disease: manifestations and current concepts in pathogenesis and management. Eur Respir Rev. 2021;30(160):210011. doi: 10.1183/16000617.0011-2021.
19. Kim BG, Lee H, Eun Y, et al. Association between rheumatoid arthritis and interstitial lung disease and impact of serologic status: a large-scale longitudinal study. Sci Rep. 2025;15(1):4885. doi: 10.1038/s41598-025-88323-7.
20. Krekeler M, Baraliakos X, Tsiami S, Braun J. High prevalence of chondrocalcinosis and frequent comorbidity with calcium pyrophosphate deposition disease in patients with seronegative rheumatoid arthritis. RMD Open. 2022;8(2):e002383. doi: 10.1136/rmdopen-2022-002383.

The European Research Journal 2026;12(3):376-386 Original Article

DOI: https://doi.org/10.18621/eurj.1847969 *Otorhinolaryngology-Head and Neck Surgery*

Outcomes of Teletherapy and Face-to-Face Voice Therapy for Vocal Fold Nodules: A First-Line Treatment Comparison

Esma Altan[1], Elife Barmak[2], Zeynep Yılmaz[2], Dilara Söylemez[1], Tuğçe Pütürgeli Özer[3], Emel Çadallı Tatar[3]

[1]Department of Otorhinolaryngology-Head and Neck Surgery, Etlik City Hospital, Ankara, Türkiye; [2]Department of Speech and Language, Yıldırım Beyazıt University, Faculty of Health Sciences, Ankara, Türkiye; [3]Private Practice Doctor, Ankara, Türkiye

ABSTRACT

Objectives: This study aimed to compare the effectiveness of teletherapy and face-to-face voice therapy in patients diagnosed with vocal fold nodules by evaluating acoustic, aerodynamic, perceptual, and patient-reported voice outcomes.

Methods: A total of 40 patients with vocal fold nodules received either teletherapy or face-to-face voice therapy for a period of six to eight weeks. Voice assessments were performed before and after therapy and included videolaryngoscopic examination, acoustic voice analysis, aerodynamic voice measurements, auditory-perceptual voice evaluation, and patient-reported outcome measures related to voice handicap, voice-related quality of life, and reflux-related symptoms.

Results: Both therapy groups demonstrated significant improvements in acoustic voice stability, perceptual voice quality, and patient-reported voice outcomes after treatment. No significant differences were observed between teletherapy and face-to-face therapy for most outcome measures. Patients receiving face-to-face therapy showed a greater change in fundamental frequency compared with those receiving teletherapy. Both therapy modalities were associated with significant reductions in voice-related handicap, improvements in voice-related quality of life, and decreased reflux-related symptoms.

Conclusions: Teletherapy appears to be an effective and accessible alternative to face-to-face voice therapy for patients with vocal fold nodules. Further studies are warranted to investigate long-term outcomes and to explore integrated and hybrid treatment models in voice rehabilitation.

Keywords: Dysphonia, Vocal Fold Nodules, Voice Therapy, Teletherapy, Tele-Rehabilitation, Non-Surgical Management

Vocal fold nodules are benign lesions of various sizes located bilaterally in the membranous middle part of the vocal folds [1]. They are fundamentally characterized by thickening of the epithelium along with varying degrees of inflammatory reactions in the superficial lamina propria [2]. The primary cause of vocal fold nodules is phonotrauma. In particular, excessive and improper

Received: December 24, 2025 **Accepted:** February 5, 2026 **Available Online:** February 10, 2026 **Published:** March 1, 2026

How to cite this article: Altan E, Barmak E, Yılmaz Z, Söylemez D, Pütürgeli Özer T, Çadallı Tatar E. Outcomes of Teletherapy and Face-to-Face Voice Therapy for Vocal Fold Nodules: A First-Line Treatment Comparison. Eur Res J. 2026;12(3):376-386. doi: 10.18621/eurj.1847969

Corresponding author: Esma Altan, MD., Phone: +90 312 797 00 00, E-mail: esmaaltan@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

voice use is crucial in the formation of vocal fold nodules. Vibrations during phonation lead to increased forces and maximum impact stress and trauma in the middle membranous part of the vocal fold, resulting in scarring and subsequent nodule development [3]. In the literature, the only significant vocal fold nodule prevalence data (n=19,636 according to research data) was reported to be 1.31% [4]. Perceptual features such as roughness, low pitch, breathy voice, strained voice quality, imbalance and vocal fry are observed in vocal fold nodules [5].

Pharmacological treatments, phonosurgery, and voice therapy are among the intervention options in the treatment of voice disorders. For patients with vocal fold nodules, voice therapy and laryngeal microsurgical methods are generally preferred [5]. This highlights the necessity of effective non-surgical rehabilitation approaches that may reduce or delay the need for operative intervention in appropriate patients. Regardless of therapy approach differences, most therapy protocols should include three basic components: (a) educating the patient on behaviors contributing to nodule formation, (b) eliminating maladaptive behaviors exacerbating dysphonia, and (c) modifying individual and situation-specific behaviors that worsen the condition [1]. Cognitive processes aimed at motor learning and maintenance/transfer of new voice behaviors are also crucial. In addition, motor learning principles, including skill acquisition, retention, and transfer of newly learned vocal behaviors, play a central role in achieving long-term therapeutic success. Although studies have shown positive effects of vocal therapies on the healing process of vocal fold nodules, it has been reported that there are differences in the duration and intensity of therapy. On an international level, individuals with voice disorders often work in long-duration, inflexible job conditions, which hinder regular participation in voice therapy [6]. Additionally, they often must travel long distances to access experienced therapists, resulting in transportation costs and limited participation (typically 2 weeks) [7]. The COVID-19 pandemic introduced remote interventions due to social distancing and led to increased teletherapy applications thanks to technological advancements such as faster internet and wider access [8, 9].

Teletherapy offers several benefits, including increased accessibility for individuals with geographic or physical limitations, cost-effectiveness, ease of use, and increased adherence to behavioral interventions [9, 10]. Studies on teletherapy for various voice disorders have shown positive outcomes, including in Parkinson's disease [6, 11]. It is found to be a versatile and effective approach for dysphonic patients across age groups and regions [12]. Intensive teletherapy programs are potentially beneficial for patients with vocal fold nodules. Given that successful rehabilitation may prevent escalation to surgical management, establishing the effectiveness of teletherapy for vocal fold nodules has direct implications for clinical decision-making.

Recent studies have demonstrated that teletherapy is an effective and feasible approach for the management of various voice disorders, offering comparable outcomes to face-to-face interventions in selected patient populations [13-15]. Most of the available evidence, however, focuses on heterogeneous groups of voice disorders or functional dysphonia, while data specifically addressing vocal fold nodules remain limited [15]. This gap in the literature underscores the need for studies directly comparing teletherapy and conventional face-to-face voice therapy in patients with vocal fold nodules.

This study aims to compare the effectiveness of face-to-face and teletherapy interventions as first-line, non-surgical treatment options in patients with vocal fold nodules. The research questions are as follows:

1.Is face-to-face therapy more effective in patients with vocal fold nodules?

2.Are teletherapy applications more effective?

3.Are both therapy modalities equally effective?

4.Could teletherapy reduce the clinical need for phonosurgery by providing an effective first-line treatment option?

METHODS

Study Place and Design

This study was designed as a retrospective observational cohort study, conducted by reviewing medical records and voice therapy charts of patients diagnosed with bilateral vocal fold nodules. Data were collected from the University of Health Sciences, Dışkapı Yıldırım Beyazıt Training and Research

Hospital, Department of Otorhinolaryngology. Medical records and voice therapy charts of patients diagnosed with bilateral vocal fold nodules who received routine clinical care between May 2020 and March 2021 were retrospectively reviewed.

This study was approved by the Clinical Research Ethics Committee of the Ministry of Health, University of Health Sciences Dışkapı Yıldırım Beyazıt Training and Research Hospital (17.05.2021 / No: 111/01). Since the design was retrospective, no intervention was performed for research purposes, and only existing clinical data were evaluated. The study was conducted in accordance with the principles of the Declaration of Helsinki.

In our tertiary laryngology clinic, a standardized comprehensive voice assessment protocol is routinely applied to all patients presenting with voice complaints, both at initial evaluation and after completion of voice therapy. This routine protocol includes videolaryngostroboscopic examination, acoustic and aerodynamic measurements, auditory-perceptual evaluation, and patient-reported outcome measures. Therefore, these assessments are part of standard clinical care and were naturally available in the medical records for both pre- and post-therapy evaluations.

Participants

Patients diagnosed with bilateral vocal fold nodules by videolaryngoscopic examination were included. The teletherapy and face-to-face therapy groups were created retrospectively based on the therapy modality the patients had received during routine clinical practice. No randomization was performed.

All patients who had complete baseline and post-therapy data were included (complete-case approach). No patients were lost to follow-up during the therapy period.

All individuals were diagnosed with bilateral vocal fold nodules through videolaryngoscopic examination performed by an experienced laryngologist and referred for voice therapy. In this study, a consent form was filled out for all patients and necessary consents were obtained.

Inclusion criteria are the following:

a)Diagnosed with vocal fold nodules,

b)Aged between 18–55 years,

c)No other organic or functional cause for voice disorder,

d)No history of neck or laryngeal surgery,

e)Agreed to participate in voice therapy.

Teletherapy group-specific inclusion criteria are included as follows:

a)Use of secure, low-cost web systems with video capabilities,

b)No hearing or visual impairments,

c)Physical endurance to sit and use a keyboard/mouse,

d)Ability to stay seated and minimize movement,

e)Adequate cognitive function,

f)Willingness to receive teletherapy,

g)Functional communication abilities (hearing, literacy, speech intelligibility).

Voice disorder assessment and therapy were conducted by three experienced speech-language therapists. All sessions were conducted in quiet rooms to minimize distractions.

Evaluations

All patients were evaluated before and after therapy with videolaryngoscopic examination, acoustic and aerodynamic measurements, auditory-perceptual evaluations, and patient self-report instruments [Grade, Roughness, Breathiness, Asthenia, and Strain scale (GRBAS), Voice Handicap Index (VHI), Voice-Related Quality of Life (V-RQOL), Reflux Symptom Index (RSI)].

Videolaryngoscopic Examination

Conducted using a rigid endoscope (XION, Berlin, Germany), capturing /i/ vowel production at comfortable pitch and loudness. Nodules were typically located at the midpoint of the membranous vocal folds.

Acoustic and Aerodynamic Measurements

Voice samples were recorded in a sound-treated environment (background noise level < 30 dB SPL) using the KayPENTAX Model 4500 CSL® (Computerized Speech Lab) system and a Shure SM48-LC microphone, positioned approximately 10 cm from the mouth at a 45° angle. The sustained vowel /a/ was recorded at a sampling rate of 44,100

Hz. Sound pressure level (SPL) calibration was performed prior to each session using a Brüel & Kjær 4231 calibrator. Acoustic analyses were performed with the Multi-Dimensional Voice Program (MDVP; KayPENTAX, version 5.1), yielding the following parameters: mean fundamental frequency (F0), jitter (%), shimmer (%), noise-to-harmonic ratio (NHR), relative average perturbation (RAP), and Soft Phonation Index (SPI). To ensure measurement reliability, each sample was recorded twice, and the mean value was used for analysis. Aerodynamic measures included maximum phonation time and s-to-z ratio (s/z ratio).

Auditory-Perceptual Evaluation and Patient Self-Report Instruments

GRBAS scale was used by three experienced clinicians, who rated each dimension [Grade (G), Roughness (R), Breathiness (B), Asthenia (A), and Strain (S)] on a 0–3 scale based on a standard reading passage. Voice Handicap Index, a self-report questionnaire where patients rated their perception of voice problems on a 5-point Likert scale (0 = never to 4 = always). Patient-reported outcomes were evaluated using the Turkish validated versions of the Voice Handicap Index-10 (VHI-10), the Voice-Related Quality of Life and Reflux Symptom Index questionnaires. Both tools have been previously adapted to Turkish with demonstrated reliability and validity [16-18].

Procedure

Each patient underwent detailed voice assessment in the laryngology clinic as part of routine clinical care. Each group received a 6–8-week voice therapy program, with sessions lasting 30–45 minutes. The initial sessions included indirect voice therapy (education on voice production and vocal hygiene), followed by individualized indirect and direct voice therapy. Voice reassessment was conducted one week after the final therapy session.

The therapy modality (teletherapy or face-to-face therapy) was determined based on routine clinical practice and patient-related or logistical factors, particularly during and after the COVID-19 pandemic. No randomization was performed.

he teletherapy sessions were conducted using the Zoom® platform (version 5.15) with a minimum internet bandwidth of 10 Mbps. The Opus audio codec (48 kHz) was used to ensure high-fidelity audio transmission. The mean attendance rate was 95%, and no participants dropped out of the study.

Both teletherapy and face-to-face interventions were delivered by the same speech-language pathologists following an identical therapy framework. Each program included indirect voice therapy (education on vocal hygiene and efficient voice use) and direct exercises (resonant voice, easy onset, flow phonation, and breathing coordination). Home practice assignments were standardized across groups and reviewed during weekly sessions.

Statistical Analysis

The data were first processed in Excel and then transferred to the IBM SPSS version 26.0 statistical software (IBM Corp., USA). Descriptive statistics were presented as mean±standard deviation and frequencies (percentages) for demographic characteristics, acoustic and aerodynamic voice parameters, auditory-perceptual ratings, and patient-reported outcome measures. Group differences in categorical variables were analyzed using the chi-square test. Data distribution was assessed using the Kolmogorov–Smirnov test, along with evaluation of skewness and kurtosis values (acceptable range: −1.5 to +1.5). Between-group comparisons were performed using the independent samples t-test for normally distributed variables and the Mann–Whitney U test for non-normally distributed variables. Within-group pre- and post-therapy comparisons were conducted using the paired samples t-test for normally distributed

TABLE 1. Gender Distribution of the Study Groups

		Male	Female	Total
Teletherapy	n	2	18	20
	%	10.0	90.0	100.0
Face-to-face	n	5	15	20
	%	25.0	75.0	100.0
Total	n	7	33	40
	%	17.5	82.5	100.0

n, number of participants. Statistical analysis: Chi-square test.

variables and the Wilcoxon signed-rank test for non-normally distributed variables. Standardized effect sizes (Cohen's d) with 95% confidence intervals were calculated for within-group pre- and post-therapy changes in both the teletherapy and face-to-face therapy groups. Given the retrospective design and fixed sample size, reporting effect sizes was intended to facilitate interpretation of the magnitude and clinical relevance of treatment effects. The level of statistical significance was set at P<0.05.

RESULTS

The study included 40 patients: 20 (50%) in the teletherapy group and 20 (50%) in the face-to-face group. In the teletherapy group, 18 (90%) were women and 2 (10%) were men, while in the face-to-face group, 15 (75%) were women and , 5 (25%) were men. The average age of the teletherapy group is 32.45±8.7 and the other group is 36.90± 9.4. Both groups are similar in age and gender (Tables 1 and 2).

Within-group analyses demonstrated significant improvements in several acoustic and aerodynamic parameters following therapy in both the teletherapy and face-to-face groups (Table 3). Upon analysis of Table 3, it was observed that the mean scores for pre-therapy jitter, shimmer, RAP, NHR, and SPI in the teletherapy group were statistically significantly different from the post-therapy mean values (P=0.002, P=0.004, P=0.002, P=0.042 and P=0.033, respectively). The mean values of F0 (196.13 Hz), jitter (1.76), shimmer (5.69), RAP (1.05), and SPI (17.72) prior to therapy in the face-to-face therapy group were statistically significantly higher than the mean values of F0 (214.29 Hz), jitter (0.872), shimmer (3.21), RAP (0.52), and SPI (10.56) following therapy (P=0.003, P=0.004, P=0.023, P=0.038 and P=0.001, respectively).

In the comparison of the teletherapy and face-to-face therapy groups before and after therapy, only the mean MPT scores were statistically significant (P=0.028 and P=0.021, respectively), while there were no statistically significant differences in other parameters (P>0.05). When comparing the pre- and post-therapy changes in acoustic and aerodynamic parameters between teletherapy and face-to-face voice therapy, a statistically significant difference was found only in the F0 parameter for both groups (P=0.005). The mean F0 change value of the face-to-face group (mean: 18.15) was higher than the mean change value of the teletherapy group (mean: -1.53).

Auditory-perceptual ratings and patient-reported outcome measures showed significant pre- to post-therapy improvement in both treatment modalities (Table 4). Table 4 shows that the mean pre-therapy G, R, B, S, GRBAS, VHI, V-RQOL, and RSI scores of the teletherapy and face-to-face voice therapy groups were statistically significant compared to the post-therapy results (P<0.05). In the comparison of the pre- and post-therapy changes within teletherapy and face-to-face voice therapy groups, no statistically significant difference was found in terms of mean G, R, B, A, S, GRBAS, VHI, V-RQOL, and RSI scores (P>0.05).

Notably, the VHI, V-RQOL, and RSI scores also reflected meaningful improvements, reinforcing the therapeutic efficacy of both intervention methods. In general, voice therapy resulted in improvement in subjective and objective voice parameters in both groups. When we compare the improvement between online or face-to-face therapy groups, there is no statistical difference between the groups.

DISCUSSION

This study compared the effectiveness of teletherapy

TABLE 2. Age Distribution of the Study Groups

	n	Mean±SD	Min-Max	df	Value (t)	P-value
Teletherapy	20	32.45±8.768	21.0-49.0	38	-1.544	0.131
Face-to-face	20	36.90±9.447	23.0-55.0			

SD, standard deviation; df, degrees of freedom; Min, minimum; Max, maximum
Statistical analysis: Independent samples t-test.

TABLE 3. Pre- and Post-Therapy Acoustic and Aerodynamic Outcomes

Parameter	Therapy	Teletherapy			Face-to-Face			Teletherapy and Face-to-Face Comparison (Before and After)		Comparison of teletherapy and face-to-face therapy in terms of exchange value						
		Mean±SD	Value	P-value	Mean±SD	Value	P-value	Value	P-value		Mean Change Value±SD	Value	P-value	95% Convidence Internal		Cohen d
														Lower	Upper	
F0	Before	212.51±31.62	-.485[a]	0.627	196.139±44.60	-2.949[a]	**0.003**	-1.217[b]	0.224	Teletherapy	-1.53±17.11					
	After	210.98±23.32			214.297±37.56			-0.920[b]	0.369	Face to Face	18.15±22.14	-2.840[b]	**0.005***	-34.41	-7.571	0.525
Jitter (%)	Before	1.729±0.887	-3.173[a]	**0.002**	1.765±2.198	-2.053[a]	**0.040**	-1.691[b]	0.91	Teletherapy	-0.78±0.85					
	After	0.943± 0.407			0.872±0.645			-1.163[b]	0.253	Face to Face	-0.89±2.18	-1.204[b]	0.229	-.811	.183	0.225
RAP	Before	1.044±0.549	-3.173	**0.002**	1.052±1.288	-2.072[a]	**0.038**	-1.623[b]	0.108	Online	-0.47±0.53					
	After	0.571±0.251			0.525±0.388			-1.136[b]	0.265	Face to Face	-0.52±1.27	-1.109[b]	0.267	-.484	.125	0.205
Shimmer	Before	4.788±1.735	-2.912	**0.004**	5.693±5.546	-2.277[a]	**0.023**	-1.055[b]	0.331	Online	-1.07±1.59					
	After	3.712±1.360			3.214±1.458			-1.461[b]	0.149	Face to Face	-2.47±5.12	-0.203[b]	0.839	-.992	.930	0.035
NHR	Before	0.136±0.028	-2,035	**0.042**	0.169±0.106	-1.120[a]	0.263	-0.298[b]	0.0779	Teletherapy	-0.01±0.02					
	After	0.123±0.018			0.131±0.023			-0.609[b]	0.547	Face to Face	-0.03±0.10	-0.122[b]	0.903	-.022	.025	0.027
VTI	Before	0.048±0.013	-.153	0.879	0.0450±0.016	-1.374[a]	0.170	-0.582[b]	0.565	Teletherapy	-0.00±0.01					
	After	0.048±0.014			0.0487± 0.018			0.000[b]	1.000	Face to Face	0.00±0.01	-0.962[b]	0.336	-.009	.003	0.175
SPI	Before	16.802±7.332	-2.128	**0.033**	17.726±9.050	-3.248[a]	**0.001**	-0.135[b]	0.904	Teletherapy	-3.74±7.82					
	After	13.055±5.447			10.569±4.863			-1.623[b]	0.108	Face to Face	-7.15±9.03	-0.703[b]	0.482	-3.093	7.698	0.130
MPT	Before	8.70±3.420	-1,470	0.082	11.05±4.071	-1.714[a]	0.087	-2.203[b]	**0.028**	Teletherapy	1.05±2.56					
	After	9.75±2.552			12.45±4.466			-2.314[b]	**0.021**	Face to Face	1.40±3.73	-0.095[b]	0.924	-2.000	1.999	0.017
s/z	Before	1.15±0.182	-1.020	0.308	1.17±0.209	-0.882[a]	0.378	-0.345[b]	0.738	Teletherapy	-0.05±0.19					
	After	1.10±0.217			1.10±0.355			-1.019[b]	0.322	Face to Face	-0.06±0.37	-0.096[b]	0.923	-.199	.199	0.0175

F0, fundamental frequency; RAP, relative average perturbation; NHR, noise-to-harmonic ratio; VT, voice turbulence index; SPI, soft phonation index; MPT, maximum phonation time; SD, standard deviation.

[a]Wilcoxon Signed Rank Test, [b]Mann-Whitney U Test. Statistically significant P-values are shown in bold.

TABLE 4. Auditory-perceptual and Patient-Reported Outcome Measures

Parameter	Teletherapy Before Mean±SD	Teletherapy After Mean±SD	Within-group P-value (Teletherapy)	Face-to-face Before Mean±SD	Face-to-face After Mean±SD	Within-group P-value (Face-to-face)	Change (Δ) Teletherapy Mean±SD	Change (Δ) Face-to-face Mean±SD	Between-group P value (Δ)	Cohen's d	95% CI Lower	95% CI Upper
G	1.45±0.510	0.25±0.444	**<0.001**	1.30±0.470	0.30±0.571	**<0.001**	-1.2±0.410	-1.0±0.458	0.162	0.180	-1.58e-5	5.89e-5
R	1.10±0.553	0.40±0.503	**<0.001**	0.95±0.394	0.15±0.366	**<0.001**	-0.7±0.470	-0.8±0.523	0.564	0.085	-1.79e-5	3.47e-5
B	0.75±0.444	0.15±0.366	**0.001**	0.70±0.571	0.15±0.366	**0.001**	-0.6±0.502	-0.5±0.510	0.752	0.050	-7.39e-5	3.06e-5
A	0.05±0.224	0.00±0.00	0.317	0.00±0.000	0.00±0.000	1.000	-0.05±0.223	0.0±0.000	0.317	0.050	0.000	0.000
S	0.50±0.513	0.00±0.00	**0.002**	0.70±0.657	0.00±0.000	0.001	-0.5±0.512	-0.7±0.656	-0.472	0.150	-3.30e-5	1.000
GRBAS	3.80±1.240	0.80±1.00	**<0.001**	3.65±1.496	0.75±1.33	**<0.001**	-3.0±1.169	-2.9±1.518	0.637	0.085	-1.00	1.000
VHI-10	16.05±10.08	4.95±6.245	**0.003**	18.50±9.04	7.50±8.389	**0.001**	-11.1±13.353	-11.0±10.228	0.903	0.022	-8.00	9.000
V-RQOL	57.70±28.20	91.87±11.58	**<0.001**	61.37±20.9	86.75±15.7	**<0.001**	34.1±30.948	25.3±23.581	0.394	0.157	-10.00	25.000
RSI	16.60±8.53	6.75±6.828	**0.001**	15.30±9.75	8.50±8.5	**0.003**	-9.8±9.499	-6.4±7.487	0.489	0.127	-8.00	3.000

Abbreviations: G, grade; R, roughness; B, breathiness; A, asthenia; S, strain; GRBAS, perceptual rating scale; VHI, Voice Handicap Index; V-RQOL, Voice-Related Quality of Life; RSI, Reflux Symptom Index; SD, standard deviation.
Within-group p-values were obtained using the Wilcoxon signed-rank test, as indicated by "a" in the original table. Between-group comparisons of change scores were performed using the Mann–Whitney U test, as indicated by "b" in the original table. Statistically significant P-values are shown in bold.

and face-to-face voice therapy in individuals diagnosed with vocal fold nodules and evaluated various acoustic, aerodynamic, and perceptual voice parameters, as well as patient self-reported voice perceptions. Our findings indicate that both therapy modalities led to significant improvements in voice quality, with no statistically significant differences between the two approaches for most outcome measures. This suggests that teletherapy is a viable and effective alternative to traditional in-person therapy, particularly for patients facing barriers such as geographic distance, time constraints, or limited access to specialized care.

In our study, improvements were observed in the patients' acoustic and aerodynamic sound parameters in the comparative effectiveness of teletherapy and face-to-face therapy. In particular, both groups showed significant decreases in jitter, shimmer, RAP, and SPI values after therapy, indicating that patients had improved voice quality and voice stability. These findings align with prior research demonstrating that structured voice therapy - regardless of delivery mode - can mitigate phonotrauma and promote healthier vocal fold vibration [19, 20]. The face-to-face group showed a trend toward greater improvement in fundamental frequency (F0). This significant change in F0 may reflect the general severity of voice impairment in the 20 adult individuals, which was evident before starting face-to-face therapy and the potential impact of real-time tactile and auditory feedback during face-to-face sessions [21]. The teletherapy group also demonstrated a significant decrease in the signal-to-noise ratio, which is evidence that therapy results in improvements in the values of noise perturbation.

It should be noted that the available literature specifically comparing teletherapy and face-to-face voice therapy in patients with vocal fold nodules remains limited. Therefore, some supportive references are derived from studies on structured voice therapy or other dysphonia populations, and the present findings should be interpreted within this context.

Individuals in both the in-person and teletherapy groups significantly improved in voice grade, roughness, breathiness, and strain based on GRBAS assessments after completing voice therapy sessions. These results indicated that after receiving voice

therapy, individuals experienced more regular vocal fold vibration, significantly improved glottal closure, increased vocal clarity, more stable breath support, and improved overall voice quality. No individuals in our study exhibited voice asthenia. Furthermore, individuals receiving face-to-face therapy experienced slightly stronger perceptual gains in the roughness (R) parameter. This is consistent with studies emphasizing the role of immediate clinician feedback in refining vocal technique [22]. Both groups reported significant improvements in Voice Handicap Index, and Voice-Related Quality of Life scores, reinforcing that patient-reported outcomes are comparable across modalities. This echoes findings by Portone-Maira *et al*. [23], who noted that teletherapy can enhance adherence and satisfaction by reducing logistical burdens. The significant decrease in RSI scores, post-therapy in both groups, suggests that voice therapy may indirectly alleviate laryngopharyngeal reflux (LPR) symptoms. This could be attributed to improved vocal hygiene and reduced mechanical irritation from phonotrauma [24].

The comparable clinical outcomes observed in this study support the growing consideration of teletherapy in speech-language pathology practice, particularly in post-pandemic healthcare settings [25].

Although both teletherapy and face-to-face voice therapy resulted in significant clinical improvements, certain modality-specific considerations should be acknowledged. Face-to-face therapy may offer advantages in terms of immediate visual, auditory, and tactile feedback, particularly during the early stages of therapy or in patients with more severe dysphonia. In contrast, teletherapy provides greater accessibility, flexibility, and continuity of care, especially for individuals facing geographical, occupational, or logistical barriers. These complementary characteristics suggest that teletherapy can serve as an effective alternative or adjunct to in-person therapy in appropriately selected patients.

The predominance of female participants (82.5%) aligns with the higher prevalence of vocal nodules in women [26], but future studies should include more male participants. Differences in participants' socio-cultural and educational levels were not controlled, which may have affected individual adaptation and participation in teletherapy. Future studies should consider stratifying samples according to these variables to minimize bias. The sample size was adequate to detect medium within-group improvements but was likely underpowered for small Group × Time interaction effects. Accordingly, we report effect sizes with 95% CIs to complement p-values, and we recommend a larger, prospectively powered RCT to confirm interaction effects. The study did not include a non-therapy control group, which limits the ability to distinguish treatment effects from spontaneous recovery or placebo responses. Future controlled trials including a no-treatment arm are warranted to strengthen the causal interpretation of these findings. Longer-term studies are needed to assess the durability of improvements and relapse rates. Variability in platforms and exercises may influence outcomes [27]. All participants received counseling on diet and lifestyle modifications to minimize laryngopharyngeal reflux symptoms (e.g., avoiding acidic and spicy foods, limiting caffeine, maintaining hydration, and not eating before bedtime). In symptomatic cases, proton pump inhibitors were prescribed as part of routine clinical management. However, adherence to these measures was not systematically tracked; therefore, LPR management was not analyzed as a separate variable.

Although voice therapy remains the primary treatment for vocal fold nodules, phonosurgery or microlaryngeal excision may be considered when conservative treatment fails or when nodules become fibrotic and persist over time. Previous studies have shown that surgical removal can improve voice quality, particularly in treatment-resistant lesions. Yet, recurrence and postoperative dysphonia may still occur, emphasizing the need for continued behavioral management after surgery [28-30]. From this perspective, the comparable therapeutic outcomes observed between teletherapy and face-to-face therapy in our study highlight teletherapy as a clinically meaningful option that may reduce or delay the necessity for surgical intervention in suitable patients. Establishing non-surgical treatment efficacy is particularly important given that nodule recurrence and poor postoperative vocal habits remain risk factors for re-operation. This reinforces the concept of teletherapy as not only an equivalent modality to in-person rehabilitation but also a potential tool for decreasing surgical demand in laryngology practice. However, these implications should be interpreted

cautiously, given the retrospective design of the study.

From a practical perspective, teletherapy-based voice therapy has been seen as an alternative treatment approach to face-to-face therapy for individuals with vocal fold nodules. It can help individuals receive the treatment they need when they are unable to attend face-to-face therapy sessions due to factors such as transportation and financial constraints (individuals living in different cities from their therapist), busy work schedules, or a pandemic. Furthermore, combining teletherapy with occasional face-to-face sessions during patients' therapy processes can optimize outcomes [31].

Strengths and Limitations

This study has several limitations that should be acknowledged. First, the retrospective design limits causal inference and increases susceptibility to selection bias. Second, although the sample size was sufficient to detect medium within-group effects, it was likely underpowered to identify small Group × Time interaction effects between teletherapy and face-to-face therapy modalities. Third, the predominance of female participants may limit the generalizability of the findings to male patients with vocal fold nodules. In addition, socio-cultural and educational differences among participants were not controlled and may have influenced engagement with teletherapy. The absence of a non-therapy control group precludes differentiation between treatment effects and spontaneous recovery. Furthermore, adherence to laryngopharyngeal reflux–related lifestyle modifications and pharmacological management was not systematically monitored and therefore could not be analyzed as an independent variable. Finally, the use of different telecommunication platforms and exercise delivery methods may have introduced variability in treatment outcomes.

Despite these limitations, this study has notable strengths. It provides one of the relatively few direct clinical comparisons between teletherapy and face-to-face voice therapy specifically in patients with vocal fold nodules. The inclusion of multidimensional outcome measures—acoustic, aerodynamic, perceptual, and patient-reported—offers a comprehensive evaluation of treatment effectiveness. Reporting effect sizes with confidence intervals strengthens the interpretation of results beyond P-values. Importantly, the findings support teletherapy as a clinically effective and accessible alternative to in-person voice therapy, with potential implications for reducing barriers to care and optimizing long-term behavioral management in laryngology practice.

CONCLUSION

This study reinforces voice therapy as the primary conservative management approach for vocal fold nodules and suggests that teletherapy represents a clinically effective and accessible alternative for patients who are unable to attend in-person sessions due to geographical, economic, or time-related limitations. Comparable improvements observed across acoustic, aerodynamic, perceptual, and patient-reported voice outcomes indicate that teletherapy can achieve similar therapeutic benefits to face-to-face voice therapy in appropriately selected patients. While these findings support the clinical feasibility of teletherapy as part of conservative voice management, they should be interpreted in light of the retrospective design, limited sample size, and lack of randomization. Further prospective, randomized studies with larger cohorts are warranted to confirm these results and to better define the role of teletherapy within standard voice rehabilitation pathways.

Ethics Approval and Consent to Participate

This study was approved by the University of Health Sciences Dışkapı Yıldırım Beyazıt Training and Research Hospital Clinical Research Ethics Committee (Decision No: 111/01; date: 17.05.2021. All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Informed consent was not required in this study because this is a retrospective study.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: EA, EB; Study Design: EA, EB, DS; Supervision: EA, EÇT; Funding: EA; Materials: EA, DS; Data Collection and/or Processing: EA, DS, ZY; Statistical Analysis and/or Data Interpretation: EA, EB; Literature Review: EA; Manuscript Preparation: EA, EB; and Critical Review: EA, TPÖ.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

This study was conducted at Dışkapı Yıldırım Beyazıt Training and Research Hospital, Department of Otorhinolaryngology, and Etlik City Hospital, Department of Otorhinolaryngology.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Birchall MA, Carding P. Vocal nodules management. Clin Otolaryngol. 2019;44(4):497-501. doi: 10.1111/coa.13324.
2. Pedersen M, McGlashan J. Surgical versus non-surgical interventions for vocal cord nodules. Cochrane Database Syst Rev. 2012;2012(6):CD001934. doi: 10.1002/14651858.CD001934.pub2.
3. Karkos PD, McCormick M. The etiology of vocal fold nodules in adults. Curr Opin Otolaryngol Head Neck Surg. 2009;17(6):420-423. doi: 10.1097/MOO.0b013e328331a7f8.
4. Woo SH, Kim RB, Choi SH, Lee SW, Won SJ. Prevalence of laryngeal disease in South Korea: data from the Korea National Health and Nutrition Examination Survey from 2008 to 2011. Yonsei Med J. 2014;55(2):499-507. doi: 10.3349/ymj.2014.55.2.499.
5. Mansuri B, Tohidast SA, Soltaninejad N, Kamali M, Ghelichi L, Azimi H. Nonmedical Treatments of Vocal Fold Nodules: A Systematic Review. J Voice. 2018;32(5):609-620. doi: 10.1016/j.jvoice.2017.08.023.
6. Fu S, Theodoros DG, Ward EC. Intensive versus traditional voice therapy for vocal nodules: perceptual, physiological, acoustic and aerodynamic changes. J Voice. 2015;29(2):260.e31-44. doi: 10.1016/j.jvoice.2014.06.005.
7. Mashima PA, Doarn CR. Overview of telehealth activities in speech-language pathology. Telemed J E Health. 2008;14(10):1101-1117. doi: 10.1089/tmj.2008.0080.
8. Saki N, Nasiri R, Bayat A, Nikakhlagh S, Salmanzadeh S, Khoramshahi H. Relationship Between Vocal Fatigue Index and Acoustic Voice Scales in Patients With Coronavirus Infection. J Voice. 2025;39(5):1415.e1-1415.e5. doi: 10.1016/j.jvoice.2023.04.023.
9. Gherson S, Tripp R, Goudelias D, Johnson AM. Rapid Implementation of Teletherapy for Voice Disorders: Challenges and Opportunities for Speech-Language Pathologists. J Voice. 2025;39(6):1535-1543. doi: 10.1016/j.jvoice.2023.06.024.
10. Myers B, Hary E, Ellerston J, Barkmeier-Kraemer JM. Telepractice Considerations for Evaluation and Treatment of Voice Disorders: Tailoring to Specific Populations. Am J Speech Lang Pathol. 2022;31(2):678-688. doi: 10.1044/2021_AJSLP-21-00206.
11. Constantinescu G, Theodoros D, Russell T, Ward E, Wilson S, Wootton R. Assessing disordered speech and voice in Parkinson's disease: a telerehabilitation application. Int J Lang Commun Disord. 2010;45(6):630-644. doi: 10.3109/13682820903470569.
12. Smith JA, Ryder C, Uink B, et al. Celebrating Aboriginal and Torres Strait Islander voices in Australian health promotion. Health Promot J Austr. 2023;34(4):728-730. doi: 10.1002/hpja.775.
13. Scott AM, Clark J, Cardona M, et al. Telehealth versus face-to-face delivery of speech language pathology services: A systematic review and meta-analysis. J Telemed Telecare. 2025;31(9):1203-1215. doi: 10.1177/1357633X241272976.
14. Varma M, Nayak S, Rao RR, Devadas U. Effectiveness of Teletherapy in Managing Voice Disorders: A Systematic Review. Indian J Otolaryngol Head Neck Surg. 2025;77(8):3320-3336. doi: 10.1007/s12070-025-05775-z.
15. Salehi S, Ghelichi L, Tarameshlu M. Tele-practice Versus In-Person Voice Therapy: Impacts on Vocal Characteristics of Voice Actors. Med J Islam Repub Iran. 2025;39:54. doi: 10.47176/mjiri.39.54.
16. Kılıç MA, Okur E, Yıldırım İ, et al. Ses Handikap Endeksi (Voice Handicap Index) Türkçe versiyonunun güvenilirliği ve geçerliliği [Reliability and validity of the Turkish version of the Voice Handicap Index]. Kulak Burun Bogaz Ihtis Derg. 2008;18(3):139–147. [Article in Turkish]
17. Tezcaner ZÇ, Aksoy S. Reliability and Validity of the Turkish Version of the Voice-Related Quality of Life Measure. J Voice. 2017;31(2):262.e7-262.e11. doi: 10.1016/j.jvoice.2016.04.012.

18. Akbulut S, Aydinli FE, Kuşçu O, et al. Reliability and Validity of the Turkish Reflux Symptom Index. J Voice. 2020;34(6):965.e23-965.e28. doi: 10.1016/j.jvoice.2019.05.015.
19. Roy N, Barkmeier-Kraemer J, Eadie T, et al. Evidence-based clinical voice assessment: a systematic review. Am J Speech Lang Pathol. 2013;22(2):212-26. doi: 10.1044/1058-0360(2012/12-0014).
20. Van Stan JH, Roy N, Awan S, Stemple J, Hillman RE. A taxonomy of voice therapy. Am J Speech Lang Pathol. 2015;24(2):101-25. doi: 10.1044/2015_AJSLP-14-0030.
21. Chen SH, Hsiao TY, Hsiao LC, Chung YM, Chiang SC. Outcome of resonant voice therapy for female teachers with voice disorders: perceptual, physiological, acoustic, aerodynamic, and functional measurements. J Voice. 2007;21(4):415-425. doi: 10.1016/j.jvoice.2006.02.001.
22. Behrman A, Rutledge J, Hembree A, Sheridan S. Vocal hygiene education, voice production therapy, and the role of patient adherence: a treatment effectiveness study in women with phonotrauma. J Speech Lang Hear Res. 2008;51(2):350-366. doi: 10.1044/1092-4388(2008/026).
23. Farsi S, Brown MC, Alsup N, et al. The Role of Telemedicine in Voice Therapy. Ear Nose Throat J. 2024 Jun 10:1455613241258646. doi: 10.1177/01455613241258646.
24. Koufman JA, Aviv JE, Casiano RR, Shaw GY. Laryngopharyngeal reflux: position statement of the committee on speech, voice, and swallowing disorders of the American Academy of Otolaryngology-Head and Neck Surgery. Otolaryngol Head Neck Surg. 2002;127(1):32-35. doi: 10.1067/mhn.2002.125760.
25. Kratz E, Scheffer J, Volc D, Seebacher B. Comparing face-to-face and online LSVT®LOUD speech training using LSVT®Coach in patients with Parkinson's disease: a pilot randomised controlled trial. BMC Neurol. 2025;25(1):183. doi: 10.1186/s12883-025-04161-0.
26. Cohen SM, Kim J, Roy N, Asche C, Courey M. Prevalence and causes of dysphonia in a large treatment-seeking population. Laryngoscope. 2012;122(2):343-348. doi: 10.1002/lary.22426.
27. Cartmill B, Wall LR, Ward EC, Hill AJ, Porceddu SV. Computer Literacy and Health Locus of Control as Determinants for Readiness and Acceptability of Telepractice in a Head and Neck Cancer Population. Int J Telerehabil. 2016;8(2):49-60. doi: 10.5195/ijt.2016.6203.
28. Béquignon E, Bach C, Fugain C, et al. Long-term results of surgical treatment of vocal fold nodules. Laryngoscope. 2013;123(8):1926-1930. doi: 10.1002/lary.23768.
29. Mobarsa V, Samdani SK, Gurjar VS. Outcome Analysis of Microlaryngeal Surgery for Benign Lesions of Vocal Cord Using Videostroboscopy and Voice Handicap Index. Indian J Otolaryngol Head Neck Surg. 2019;71(Suppl 1):327-332. doi: 10.1007/s12070-018-1300-0.
30. Martins RHG, Siqueira DB, Dias NH, Gramuglia ACJ. Laryngeal Microsurgery for the Treatment of Vocal Nodules and Cysts in Dysphonic Children. Folia Phoniatr Logop. 2020;72(4):325-330. doi: 10.1159/000502477.
31. Guglani I, Sanskriti S, Joshi SH, Anjankar A. Speech-Language Therapy Through Telepractice During COVID-19 and Its Way Forward: A Scoping Review. Cureus. 2023;15(9):e44808. doi: 10.7759/cureus.44808.

The European Research Journal 2026;12(3):387-396 Original Article

DOI: https://doi.org/10.18621/eurj.1746220 *Public Health*

Validity and Reliability of the Turkish Version of the Mental Health Literacy Scale in a General Population

Buğra Taygun Gülle[1], Selma Karabey[2]

[1]Republic of Turkey Ministry of Health, İzmir Provincial Health Directorate, İzmir, Türkiye; [2]Department of Public Health, İstanbul University, İstanbul Faculty of Medicine, İstanbul, Türkiye

ABSTRACT

Objectives: This study examines the validity and reliability of the Turkish version of the Mental Health Literacy Scale (T-MHLS) in a general population sample. Mental health literacy (MHL) is essential for recognizing and addressing mental health disorders; however, limited knowledge in this area contributes to underreported cases and untreated conditions.

Methods: The study was conducted using face-to-face interviews, adapting the T-MHLS from the original MHLS. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were conducted to validate the structure of the scale. The sample size and data collection method ensured robust data, while validity and reliability analyses followed established psychometric protocols.

Results: Data were collected from 401 participants, and items with low factor loadings were removed, resulting in a final five-factor structure with 30 items, explaining 58.7% of the variance. Known-group validity analysis revealed that factors such as gender, marital status, education level, income perception, and personal or familial experiences with mental illness significantly influenced MHL scores. The Cronbach's alpha for the final scale was 0.861.

Conclusions: The T-MHLS is shown to be a valid and reliable tool for assessing mental health literacy in Turkey. This scale is recommended for use in future studies involving various demographic groups to effectively measure MHL levels.

Keywords: Mental Health Literacy, Turkish Version, Scale Validation, Validity

Globally, 1.1 billion individuals are affected by mental illnesses or substance abuse disorders, collectively representing a substantial proportion of the non-fatal disease burden [1]. Untreated mental disorders give rise to a number of adverse outcomes, including difficulties in work settings, the formation of problematic interpersonal relationships, and impaired family functioning. Furthermore, they are associated with an increased risk of developing physical conditions such as diabetes mellitus, heart disease, and stroke, which ultimately reduce life expectancy [2-4].

Despite the significance of mental health, the proportion of individuals with mental health disorders who seek and obtain appropriate assistance remains relatively low [5]. Mental disorders affect one-third of the global population each year, but only two-thirds receive treatment. In countries like India and China,

Received: July 19, 2025 **Accepted:** August 30, 2025 **Available Online:** October 1, 2025 **Published:** March 1, 2026

How to cite this article: Gülle BT, Karabey S. Validity and Reliability of the Turkish Version of the Mental Health Literacy Scale in a General Population. Eur Res J. 2026;12(3):387-396. doi: 10.18621/eurj.1746220

Corresponding author: Buğra Taygun Gülle, MD., Phone: +90 232 441 81 11, E-mail: bugrataygun.gulle@deu.edu.tr

Available Online at https://dergipark.org.tr/en/pub/eurj

up to 80% do not seek medical assistance [5]. In Europe, 27% of the population is affected annually, with 74% not receiving treatment [6]. Factors contributing to these low treatment rates include stigma, distrust of those who can offer help, and difficulty in recognizing mental health disorders [7].

The concept of health literacy, which has been in use for over four decades, has recently gained significant attention, resulting in a notable increase in research publications on the subject [1]. As research on health literacy has expanded, it has provided a more nuanced understanding of the concept, which has further allowed for the study of health literacy in specific subcategories. Mental health literacy (MHL) is one of that subcategories.

In 1997, Jorm *et al*. [8] defined MHL as "knowledge and beliefs about mental disorders which aid their recognition, management, or prevention". MLH's scope was defined as the ability to recognize specific disorders; knowing how to seek mental health information; knowledge of risk factors and causes, of self-treatments, and of professional help available; and attitudes that promote recognition and appropriate help-seeking in the same article. Low MHL rates contribute to the underreporting and undiagnosis of mental disorders, which, in turn, result in untreated illnesses, unemployment, higher morbidity and mortality rates, increased stress for caregivers, and lost years of life due to suicide and homicide [9].

A variety of methods and scales are employed to assess MHL. Some of these scales concentrate on particular disorders (e.g., depression or bipolar disorder) and may consequently fail to include essential elements of MHL. Conversely, other scales include questions that extend beyond the boundaries of MHL. Moreover, some scales are unable to generate quantitative results. This underscores the necessity for an updated, quantitative scale to assess MHL in a reliable and consistent manner [10].

O'Connor and Casey [11] developed Mental Health Literacy Scale (MHLS) in 2015 to address these issues. In a systematic review, the MHLS was identified as the optimal scale in accordance with the COSMIN (Consensus-Based Standards for the Selection of Health Status Measurement Instruments) criteria for measuring MHL [12]. Since its publication, the MHLS has been translated into numerous languages and has been employed extensively.

Despite previous studies examining the validity and reliability of the MHLS in Turkish, discrepancies in sampled populations and results have led to variations in item counts and the distribution of sub-factors in the Turkish version of the scale [13, 14]. This study aimed to evaluate the validity and reliability of the Turkish version of the MHLS within a population representative of the broader community and to examine factors associated with MHLS scores.

METHODS

Sampling and Data Collection

This study was designed as a methodological study. Data were collected from the waiting rooms of surgery and internal medicine outpatient clinics from the relatives of patients who were over 18 years old and volunteered to participate in the study. To determine the sample size, a variable-to-participant ratio of at least 10:1 was targeted. The scale, consisting of 35 items, was administered to 401 participants. Data collection was carried out through face-to-face interviews, which ensured no missing data.

Data Collection Tools

Personal Information Form

A personal data form, developed by the researchers based on a literature review, was used to gather information on participants' age, gender, marital status, education level, income perception, occupation, any diagnosed chronic illnesses (and specific illnesses, if any), any diagnosed mental disorders, current treatment for mental disorders, and presence of family members with a mental disorder diagnosis.

Turkish Version of Mental Health Literacy Scale

The MHLS is a 35-item scale developed by O'Connor and Casey [11] in 2015 for the purpose of measuring six key characteristics of mental health literacy. The MHLS employs a 4- or 5-point Likert scale, with a potential range of scores from 35 to 160. It incorporates both positively and negatively worded items.

Turkish version of MHLS (T-MHLS) was formed by Akdoğan [13] in 2018. A group translation method (Hambleton, Merenda, and Spielberger) was used to

facilitate linguistic adaptation. In the initial phase, the scale was translated into Turkish by seven experts proficient in English and native Turkish speakers. In the second phase, a distinct group of translators revised the translated text independently, without reference to the original, and subsequently conducted a comparison with the source text. In the third phase, the revised scale was back-translated into English. Three linguistic experts compared this version with the original, made the necessary adjustments, and obtained feedback from the original author. Finally, the scale was tested on a group of 20 individuals, and adjustments were made based on their feedback. The content reliability index was calculated, with the lowest score determined to be 0.890 [13].

Ethical Considerations

Ethical approval was obtained from the Istanbul University Istanbul Faculty of Medicine Clinical Research Ethics Committee (decision date: 09.08.2019; decision number: 2019-13/1007). Permission to employ the MHLS was obtained from Matt O'Connor, the scale's original developer, and Emine Akdoğan, the creator of the Turkish version.

Statistical Analysis

The reliability of the Turkish version of the MHLS (T-MHLS) was evaluated through the implementation of Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). In order for an item to be included in the item analysis, the item-total score correlation must be at least 0.20 for all items. The adequacy of the sample was evaluated through the application of the Kaiser-Meyer-Olkin (KMO) coefficient and Bartlett's Sphericity Test, with the established criteria being a KMO value exceeding 0.70 and a highly significant Bartlett's X^2 [15]. Principal component analysis with Varimax rotation was employed in the EFA. Factors that contributed more than 5% of the variance and had eigenvalues above 1 were retained. It was anticipated that each item would demonstrate a minimum shared variance and factor loading of 0.40 in both the factor and pattern matrices, with a minimum difference of 0.10 to control for double loading. CFA was conducted to validate the model generated by EFA. Error adjustments were applied based on model recommendations: adjustments decreasing χ^2 by at least 20 points if $\chi^2 <$ 1000, and by at least 30 points if $\chi^2 \geq 1000$, within the same factor's variables.

The model fit was evaluated using several indices, including the ratio of χ^2 to df, the root mean square error of approximation (RMSEA), the 90% confidence interval for RMSEA, the standardized root mean square residual (SRMR), the Tucker–Lewis index (TLI), and the comparative fit index (CFI). The acceptable thresholds for these indices are outlined in Table 1. The objective was to achieve a difference of no more than 0.10 between the lower and upper bounds of the 90% confidence interval (CI) for the RMSEA. Cronbach's alpha was calculated to assess the internal consistency reliability of the scale, while the Guttman split-half coefficient was employed to evaluate the correlation between the first and second halves of the scale.To assess known-group validity, the Mann-Whitney U test, Kruskal-Wallis test, or Spearman correlation were applied to group-based questions from the personal information form, and effect sizes were calculated for group differences. CFA was conducted using IBM SPSS AMOS 16, and other statistical analyses were performed with SPSS version 21.0.

RESULTS

Sample characteristics

The mean age of the participants was 38.4±13.8 years; 55.1% were men. Most participants were married (63.3%). Other characteristics of the participants are shown in Table 2.

Construct Validty

Item Analysis and EFA

Following the initial item-total score correlation analysis, four items (questions 9, 10, 12, and 15) with correlation coefficients below 0.2 were identified, but they were retained for the initial factor analysis phase. The KMO coefficient for the T-MHLS was calculated as 0.864, and Bartlett's Sphericity Test yielded a χ^2 value of 6476.476 ($P<0.001$), indicating that the sample size was adequate for factor analysis.

A principal component analysis with Varimax rotation was conducted to identify the underlying factors. This initial analysis revealed eight factors with eigenvalues greater than one. Since five of these

TABLE 1. Model Fit Indices and Acceptable Thresholds for Goodness-Of-Fit

Fit indices	Good fit criteria	Acceptable fit criteria
χ^2/df[1]	$0 \leq \chi^2/df \leq 2$	$2 \leq \chi^2/df \leq 5$
RMSEA[2]	0≤RMSEA≤0.05	0.05≤RMSEA≤0.08
SRMR[3]	0≤SRMR≤0.05	0.05≤SRMR≤0.08
TLI[4]	0.95≤TLI≤1.00	0.90≤TLI≤0.95
CFI[5]	0.95≤CFI≤1.00	0.90≤CFI≤0.95

[1]χ^2/df, the ratio of chi-square to degrees of freedom; [2]RMSEA, root mean square error of approximation; [3]SRMR, standardized root mean square residual, [4]TLI, Turker-Lewis index; [5]CFI, comparative fit index.

TABLE 2. General Characteristics of the Participants

Descriptive Characteristics	Number	Percentage (%)
Gender		
Male	221	55.1
Female	180	44.9
Marital status		
Married	254	63.3
Single	147	36.7
Education level		
Primary school or below	49	12.2
Middle school	30	7.5
High school	115	28.7
University or above	186	51.6
Perceived income		
High income	63	15.7
Average	230	57.4
Low income	108	26.9
Chronic illness		
Yes	69	17.2
No	332	82.8
Presence of mental illness		
Yes	27	6.7
No	374	93.3
Relatives with mental illness		
Yes	21	5.2
No	380	94.8

factors explained approximately 5% or more of the variance, the analysis was repeated with a five-factor model. In this analysis, the same four items exhibited factor loadings below 0.4. In light of their low item-total correlations and factor loadings, these items were excluded from the scale.

In the final analysis, all remaining items exhibited extraction values above 0.30 prior to rotation, and no item demonstrated a factor loading below 0.40 on its primary factor within the rotated factor matrix. Furthermore, no items demonstrated cross-loading. Following the exclusion of the four items previously mentioned, a five-factor structure was identified, which explained 57.3% of the total variance in the 31-item version of the scale (Table 3).

Confirmatory Factor Analysis

The five-factor, 31-item scale was found to be a valid and reliable measurement tool for MHL following the EFA. CFA was conducted to assess the model's fit and to identify any potentially problematic items. In the initial model, only one item (item 20) had a factor loading below 0.3 (factor loading = 0.298). As this item was also the only one with a factor loading below 0.5 in the EFA (0.441), it was excluded, and the model was re-evaluated (Figure 1).

CFA Model Fit Indices

The fit index values of the model before setting the covariance of errors are shown in Table 4. Values other than TLI were found to be close to acceptable or good fit levels. The 90% confidence interval bounds for RMSEA were found to have a difference of 0.10 units between them. Corrections were applied to the model by adjusting three error covariances as presented in Figure 1, and the fit index values obtained after setting the covariance are presented in Table 4.

Known-group Validity

Table 5 shows the results of the known-group validity analysis. Higher mental health literacy scores were significantly associated with being female (P=0.018), single (P=0.016), having a higher level of education (P<0.001), perceiving high income (P=0.003), having a mental illness (P<0.001), receiving treatment for a mental illness (P<0.001), or having a relative with a mental illness (P<0.001).

TABLE 3. Factor Structure, Communalities, and Item-Total Correlations of the T-MHLS

Factor names, item numbers, and content	Factor loadings	Communalities	Item-total correlation
Attitude toward individuals with mental disorders			
Eigenvalue= 6.17, Variance= 19.91%, α= 0.930			
29. Being a neighbor with them	0.843	0.735	0.409
30. Participating in social activities with them	0.861	0.775	0.413
31. Being friends with them	0.885	0.802	0.416
32. Working with them at the workplace	0.900	0.824	0.374
33. Marrying someone with a mental disorder (within family)	0.822	0.709	0.370
34. Voting for them	0.663	0.672	0.407
35. Hiring them for a job	0.853	0.747	0.389
Attitudes toward mental disorders			
Eigenvalue= 5.13, Variance= 16.54%, α= 0.875			
20. They can recover if they want*	0.441	0.491	0.225
21. It is a sign of weakness	0.683	0.632	0.546
22. It is not a disease	0.746	0.600	0.582
23. They are dangerous	0.747	0.660	0.600
24. Avoiding them is the best solution	0.744	0.633	0.609
25. I wouldn't tell anyone if I had one	0.778	0.695	0.595
26. Seeking professional help is a sign of weakness	0.777	0.695	0.632
27. I wouldn't seek help from a professional	0.716	0.634	0.605
28. Treatment is ineffective	0.660	0.560	0.500
Recognition of mental disorders			
Eigenvalue= 3.07, Variance= 9.89%, α= 0.809			
1. Recognizing social phobia	0.624	0.532	0.221
2. Recognizing generalized anxiety disorder	0.726	0.586	0.338
3. Recognizing major depressive disorder	0.684	0.551	0.338
4. Recognizing personality disorders	0.623	0.489	0.285
5. Recognizing dysthymia	0.723	0.589	0.355
6. Recognizing agoraphobia symptoms	0.580	0.513	0.277
7. Recognizing bipolar disorder symptoms	0.645	0.488	0.365
8. Recognizing substance dependency symptoms	0.504	0.444	0.315
Accessing mental health information			
Eigenvalue=1.88, Variance=6.06%, α=0.79			
16. Obtaining information	0.753	0.615	0.449
17. Using computers and phones to access information	0.715	0.643	0.317
18. Getting information from a professional	0.786	0.718	0.497
19. Accessing resources	0.734	0.673	0.530
Knowledge about mental disorder treatments			
Eigenvalue=1.51, Variance=4.86%, α=0.66			
11. Quality of sleep	0.643	0.670	0.284
13. Cognitive behavioral therapy	0.814	0.690	0.313
14. Confidentiality principle of professionals	0.772	0.682	0.269

T-MHLS, Turkish version of the mental health literacy scale.
*Removed from the scale after confirmatory factor analysis.

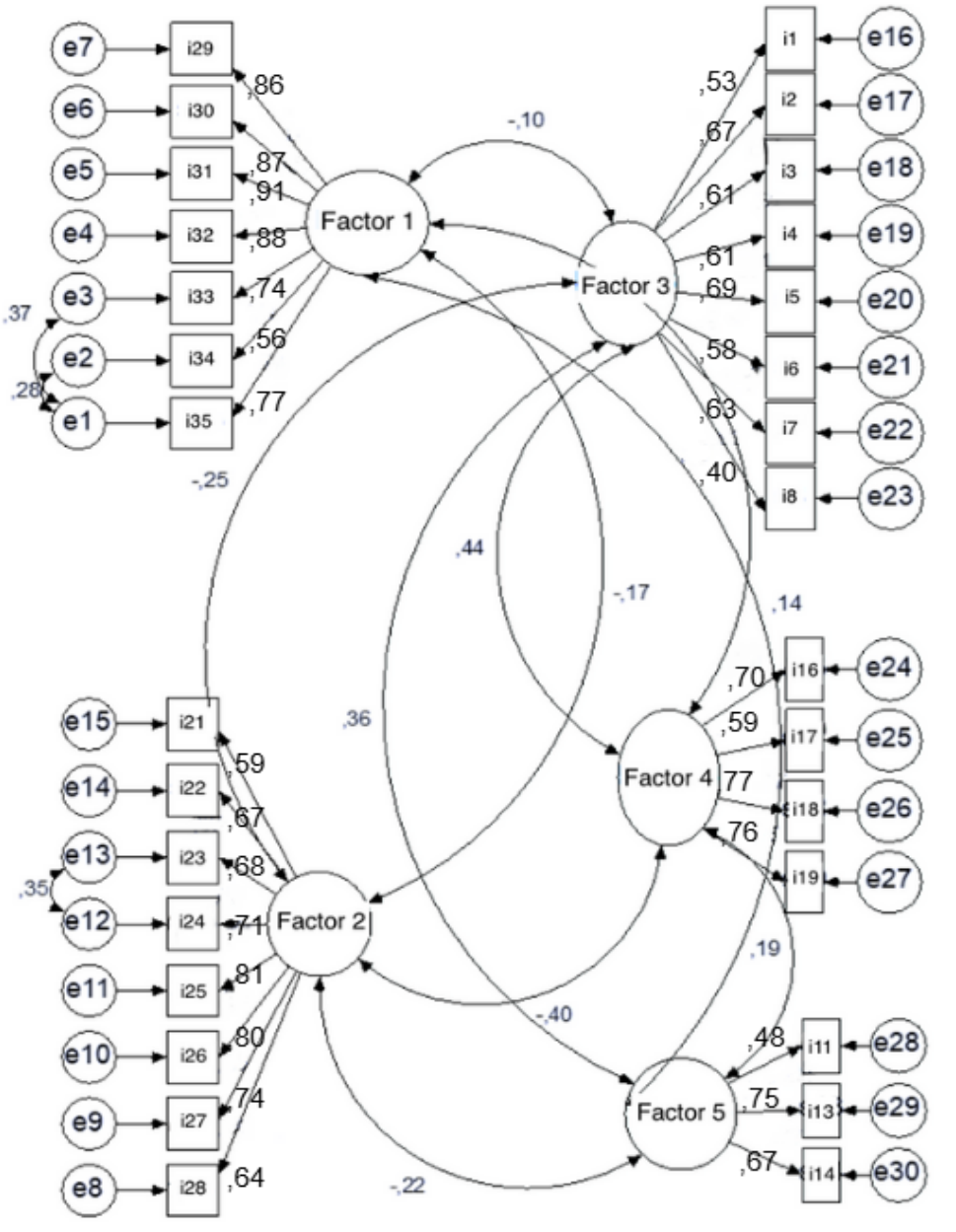

FIGURE 1. **Confirmatory factor analysis model of the 30-item mental health literacy scale.**

TABLE 4. **Model Fit indices Before and After Corrections**

Fit indices	Before correction	After correction
χ^2/df[1]	2.4	2.06
CFI[2]	0.901	0.926
TLI[3]	0.891	0.918
90% CI RMSEA[4]	0.054-0.064	0.046-0.056
RMSEA[5]	0.059	0.051
SRMR[6]	0.054	0.054

[1]Chi-square divided by degrees of freedom, [2]Comparative Fit Index, [3]Turker-Lewis Index, [4]Confidence Interval, [5]Root Mean Square Error of Approximation, [6]Standardized Root Mean Square Residual.

TABLE 5. Comparison of Mental Health Literacy Across Demographic Groups

Variable		Attitude toward individuals with mental disorders	Attitudes toward mental disorders	Recognition of mental disorders	Accessing mental health information	Knowledge about mental disorder treatments	Total MHL scale score
Gender							
Male (n=221)		13.8±5.9	30.1±8.1	24.8±3.74	15.8±3.1	9±2	93.5±13.7
Female (n=180)		15.3±5.8	31.2±6.8	25.2±3.93	15.8±3.4	9.3±1.9	96.8±13
	P value	**0.006**	0.381	0.214	0.811	0.164	**0.018**
Marital status (P value)							
Single (n=147)		14.8±6.3	31.8±7.4,	25.4±3.8	16.3±3.1	9.2±2	97.6±13.1
Married (n=254)		14.3±5.7	29.9±7.6	24.8±3.8	15.5±3.3	9.1±1.9	93.5±13.5
	P value	0.435	**0.008**	0.331	**0.033**	0.221	**0.016**
Education level							
High school or below (n=215)		14.2±6.1	28.6±8	24.4±4	15.3±3.3	8.7±2.1	91±14.1
University or above (n=186)		14.8±5.8	33±6.2	25.7±3.5	16.4±3	9.6±1.6	99.5±10
	P value	0.179	**<0.001**	**0.001**	**0.001**	**<0.001**	**<0.001**
Income perception							
Low income (n=108)		13.7±5.8	28.3±8.5	24.6±4.2	15.3±3.2	8.8±2.1	90.7±14.3
Average income (n=230)		14.8±6.1	31.5±7.1	25.2±3.4	16.1±3.1	9.2±1.8	96.8±12.1
High income (n=63)		14.4±5.3	31.5±6.7	25.1±4.6	15.5±3.6	9.3±2.1	95.7±12.1
	P value	0.262	**0.004**	0.381	0.069	0.166	**0.003**
Chronic illness							
Yes (n=69)		15.7±6.8	30.4±8	25.2±3.7	15.2±3.2	9.1±1.9	95.5±16.2
No (n=332)		14.3±5.7	30.7±7.5	24.9±3.9	15.9±3.2	9.1±2	94.9±12.9
	P value	0.174	0.967	0.493	0.116	0.636	0.665
Mental illness							
Yes (n=27)		20.8±6.7	32.6±7.9	26.6±3.9	15.8±4.1	10.1±1.5	105.8±17.3
No (n=374)		14±5.6	30.5±7.5	24.9±3.8	15.8±3.2	9.1±1.9	94.2±12.8
	P value	**<0.001**	0.075	**0.015**	0.501	**0.006**	**<0.001**
Mental health treatment							
Receiving treatment (n=21)		21.1±7.2	33.3±7.9	27±4.2	15.9±4	10.1±1.5	107.3±17.8
Not receiving treatment (n=380)		14.1±5.7	30.5±7.5	24.9±3.8	15.8±3.2	9.1±1.9	94.3±12.9
	P value	**<0.001**	**0.027**	**0.012**	0.462	**0.029**	**<0.001**
Relative with mental illness							
Yes (n=39)		21.1±5.5	32.6±6.7	25.6±3.4	16±3.3	10.1±1.4	105.4±12
No (n=362)		13.8±5.5	30.4±7.6	24.9±3.9	15.8±3.2	9±2	93.3±13.2
	P value	**<0.001**	0.060	0.077	0.488	**0.001**	**<0.001**

Data are shown as mean±standard deviation. MHL, mental health literacy.

Reliability

The Cronbach's alpha value for the final 30-item version of the scale was found to be 0.861. Cronbach's alpha values for the individual factors were 0.930, 0.890, 0.809, 0.792, and 0.656, respectively. Each item was most strongly correlated with its own factor score, with the lowest item-factor correlation being 0.555. For item-total correlations, no item had a correlation below 0.2, with the lowest being 0.232.

The Guttman Split-Half coefficient for the model was calculated as 0.875. The average T-MHLS total score was 95±13.5, with a minimum score of 55 and a maximum score of 132 out of a possible 139.

DISCUSSION

This study assesses the validity and reliability of the MHLS by applying it to a sample that is more representative of the general population. Furthermore, it compares the final version of the scale with other adaptations conducted in various countries.

The original MHLS was applied to a group of psychology students with an average age of 21.1±6.3 years [11]. In the Turkish adaptation by Akdoğan, the scale was administered to a sample in which 98% were single and 38.8% had taken a psychology course, with a mean age of 22±2.2 years [13]. Akdoğan's study, conducted among university students, suggested reducing the T-MHLS to a 22-item version by excluding 13 questions. However, this shortened version covers only three of the MHLS's six main aspects, omitting important dimensions such as risk factors, information about causes, self-treatment knowledge, and professional help information [13].In comparison, the study by Kesgin *et al*. included a smaller sample (282 vs. 401 in our study) with a demographic distribution less representative of the general population. Additionally, Kesgin *et al*. did not apply EFA, whereas our study employed both EFA and CFA to establish a more robust factor structure. In the final version of our T-MHLS, no items had factor loadings below 0.5 in EFA or below 0.4 in CFA, indicating that each item strongly aligns with the scale's purpose. Moreover, our scale does not contain factors with only two items, unlike Kesgin *et al*.'s version, which included two such factors and several items with factor loadings below 0.4. These aspects underscore the enhanced structural integrity and reliability of our version of the scale [14].

Although the original 35-item scale comprises six main aspects of MHLS, the T-MHLS is a 30-item, five-factor scale that explains 58.7% of the variance and demonstrates acceptable to good fit indices. It covers five main aspects of MHLS, excluding risk factors and information about causes. Table 6 shows the distribution of T-MHLS questions and factors across the MHL aspects.

Questions 9, 10, 12, 15, and 20 were excluded from the Turkish version of the scale due to insufficient factor loadings. Nonetheless, the factor

TABLE 6. T-MHLS Factors and MHL Characteristics

Question numbers	T-MHLS factors	Characteristics comprising MHL scope
29-30-31-32-33-34-35	Attitudes toward individuals with mental disorders	Attitudes that support recognizing mental disorders and seeking appropriate help
21-22-23-24-25-26-27-28	Attitudes toward mental disorders	
1-2-3-4-5-6-7-8	Recognition of mental disorders	Ability to recognize specific disorders
16-17-18-19	Accessing information on mental disorders	Knowledge on how to access mental health information
11-13-14	Knowledge of mental disorder treatments	Self-treatment knowledge
		Knowledge of professional help and how to access help
		Knowledge of risk factors and causes

T-MHLS, Turkish version of the mental health literacy Scale; MHL, mental health literacy.

structure of the T-MHLS showed a distribution similar to the original MHLS, suggesting that differences between versions may stem from cultural variations inherent to the distinct countries and populations in which they were applied.

The results of the known-group validity test revealed differences in MHL levels according to gender, marital status, education, income perception, having a diagnosed mental disorder, undergoing mental disorder treatment, and having a relative with a mental disorder. These findings are consistent with previous studies on MHL levels conducted in Australia, Japan, and England [11, 16-18]. Other adaptations of the MHLS show some structural variations as well. For example, the Vietnamese version consists of 31 questions across 4 factors [19] while the Persian version has 29 questions and 5 factors [20]. Another Persian adaptation includes 30 questions with a 5-factor structure [21], and the Pakistani version has 34 questions and 6 factors [22]. When considering the number of items and factor distribution, the T-MHLS aligns with these international adaptations. This suggests that, despite minor cultural differences in each adaptation, the MHLS retains a consistent and broadly applicable structure across diverse countries and cultures.

The Cronbach's alpha value of the T-MHLS was found to be 0.861. The Cronbach's alpha values for the five factors were 0.930, 0.890, 0.809, 0.792, and 0.656, respectively. The T-MHLS and four of its factors demonstrated high internal consistency. The relatively lower Cronbach's alpha value of 0.656 for one factor may be attributed to the fact that this factor comprises only three items.

Strengths and Limitations

The strengths of our study include the use of the MHLS, a new, widely adopted, and high-quality scale, to measure MHL; a study sample with a strong representation of the general population; and the face-to-face data collection method, which resulted in no missing data. However, due to the inability to access the original sample group, test-retest reliability could not be assessed. Additionally, it was not possible to compare the T-MHLS with another scale, as no similar scale exists in Turkish.

CONCLUSION

The T-MHLS is a valid and reliable tool for assessing mental health literacy in Turkish-speaking populations. Five items were removed, resulting in a 30-item scale with five subscales. The T-MHLS can be effectively used to measure mental health literacy levels in Turkey. We recommend testing and applying it in studies involving various demographic groups

Ethics Approval and Consent to Participate

This study was approved by the Istanbul University Istanbul Faculty of Medicine Clinical Research Ethics Committee (Decision No.: 2019-13/1007; date: 09.08.2019). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Permission to employ the MHLS was obtained from Matt O'Connor, the scale's original developer, and Emine Akdoğan, the creator of the Turkish version. Written informed consent was obtained from all individual participants included in the study.

Authors' Contribution

Study Conception: BTG, SK; Study Design: BTG; Supervision: SK; Funding: BTG, SK; Materials: BTG; Data Collection and/or Processing: BTG; Statistical Analysis and/or Data Interpretation: BTG; Literature Review: BTG; Manuscript Preparation: BTG; and Critical Review: SK.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

We would like to thank Hanife Özlem for her valuable assistance in statistical analysis.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390(10100):1211-1259. doi: 10.1016/S0140-6736(17)32154-2.
2. Wei Y, McGrath PJ, Hayden J, Kutcher S. Mental health literacy measures evaluating knowledge, attitudes and help-seeking: a scoping review. BMC Psychiatry. 2015;15:291. doi: 10.1186/s12888-015-0681-9.
3. Mezuk B, Eaton WW, Albrecht S, Golden SH. Depression and type 2 diabetes over the lifespan: a meta-analysis. Diabetes Care. 2008;31(12):2383-2390. doi: 10.2337/dc08-0985.
4. Rugulies R. Depression as a predictor for coronary heart disease. a review and meta-analysis. Am J Prev Med. 2002;23(1):51-61. doi: 10.1016/s0749-3797(02)00439-7.
5. Lancet Global Mental Health Group; Chisholm D, Flisher AJ, Lund C, Patel V, Saxena S, Thornicroft G, Tomlinson M. Scale up services for mental disorders: a call for action. Lancet. 2007 6;370(9594):1241-1252. doi: 10.1016/S0140-6736(07)61242-2.
6. Wittchen HU, Jacobi F. Size and burden of mental disorders in Europe--a critical review and appraisal of 27 studies. Eur Neuropsychopharmacol. 2005;15(4):357-376. doi: 10.1016/j.euroneuro.2005.04.012.
7. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry. 2010;10:113. doi: 10.1186/1471-244X-10-113.
8. Jorm AF, Korten AE, Jacomb PA, Christensen H, Rodgers B, Pollitt P. "Mental health literacy": a survey of the public's ability to recognise mental disorders and their beliefs about the effectiveness of treatment. Med J Aust. 1997;166(4):182-186. doi: 10.5694/j.1326-5377.1997.tb140071.x.
9. Tay JL, Tay YF, Klainin-Yobas P. Mental health literacy levels. Arch Psychiatr Nurs. 2018;32(5):757-763. doi: 10.1016/j.apnu.2018.04.007.
10. O'Connor M, Casey L, Clough B. Measuring mental health literacy--a review of scale-based measures. J Ment Health. 2014;23(4):197-204. doi: 10.3109/09638237.2014.910646.
11. O'Connor M, Casey L. The Mental Health Literacy Scale (MHLS): A new scale-based measure of mental health literacy. Psychiatry Res. 2015;229(1-2):511-516. doi: 10.1016/j.psychres.2015.05.064.
12. Wei Y, McGrath PJ, Hayden J, Kutcher S. Measurement properties of tools measuring mental health knowledge: a systematic review. BMC Psychiatry. 2016;16(1):297. doi: 10.1186/s12888-016-1012-5.
13. Akdoğan E. Ruh Sağlığı Okuryazarlığı Ölçeğinin Türkçe Uyarlaması, Geçerlik ve Güvenirlik Çalışması [Adaptation, validity and reliability study of the Mental Health Literacy Scale into Turkish]. Master's thesis. Karabuk: Karabuk University, Institute of Health Sciences; 2018. [Article in Turkish]
14. Kesgin M, Pehlivan S, Uymaz P. Ruh Sağlığı Okuryazarlığı Ölçeğinin Türkçe uyarlaması: kalıcılık ve yayımlanması çalışması [Study of validity and reliability of the Mental Health Literacy Scale in Turkish]. Anadolu Psikiyatri Derg. 2020;21(Suppl 1):5-13. doi: 10.5455/apd.102104. [Article in Turkish]
15. Şencan H. Reliability and validity in social and behavioral measurements. 1st ed. Ankara: Seçkin Publishing; 2005.
16. Gibbons RJ, Thorsteinsson EB, Loi NM. Beliefs and attitudes towards mental illness: an examination of the sex differences in mental health literacy in a community sample. PeerJ. 2015;3:e1004. doi: 10.7717/peerj.1004.
17. Kaneko Y, Motohashi Y. Male gender and low education with poor mental health literacy: a population-based study. J Epidemiol. 2007;17(4):114-119. doi: 10.2188/jea.17.114.
18. Furnham A, Cook R, Martin N, Batey M. Mental health literacy among university students. J Public Ment Health. 2011;10(4):198-210. doi: 10.1108/17465721111188223.
19. Dang HM, Weiss B, Trung L, Ho H. Mental health literacy and intervention program adaptation in the internationalization of school psychology for Vietnam. Psychol Sch. 2018;55(8):941-954. doi: 10.1002/pits.22156.
20. Nejatian M, Tehrani H, Momeniyan V, Jafari A. A modified version of the mental health literacy scale (MHLS) in Iranian people. BMC Psychiatry. 2021;21(1):53. doi: 10.1186/s12888-021-03050-3.
21. Heizomi H, Kouzekanani K, Asghari Jafarabadi M, Allahverdipour H. Psychometric Properties of the Persian Version of Mental Health Literacy Scale. Int J Womens Health. 2020;12:513-520. doi: 10.2147/IJWH.S252348.
22. Akhter IN, Ahmed A, Khalid S. Mental Health Literacy Scale: translation and validation in Pakistani context. Pak Soc Sci Rev. 2020;4(2):722-735. doi: 10.35484/pssr.2020(4-ii)59.

The European Research Journal 2026;12(3):397-406 Original Article

DOI: https://doi.org/10.18621/eurj.1860097 *Medical Oncology*

Early On-Treatment Nutritional Change and Baseline Inflammation Stratify Outcomes in Patients Receiving Immune Checkpoint Inhibitors

Tolga Doğan[1], Atike Gökçen Demiray[2], Burcu Yapar Taşköylü[2], Emre Hafızoğlu[3], Arzu Yaren[2], Gamze Gököz Doğu[2]

[1]Department of Medical Oncology, Denizli State Hospital, Denizli, Türkiye; [2]Department of Medical Oncology, Pamukkale University, Faculty of Medicine, Denizli, Türkiye; [3]Department of Medical Oncology, Afyonkarahisar State Hospital, Afyonkarahisar, Türkiye

ABSTRACT

Objectives: In patients receiving immune checkpoint inhibitors (ICIs), baseline nutritional status is frequently recorded, but what happens during treatment may be at least as informative. We examined whether the 3-month change in Mini Nutritional Assessment (ΔMNA)–Short Form (SF) is associated with survival, and how this relates to baseline systemic inflammation measured by modified Glasgow Prognostic Score (mGPS).

Methods: We retrospectively screened institutional records and included 54 adults with advanced/metastatic solid tumors treated with ICIs who had MNA-SF documented at treatment start and again at approximately 3 months. ΔMNA was defined as MNA-SF (~3 months) minus baseline MNA-SF (higher values indicate improvement) and was categorized for Kaplan–Meier analyses as worsened (≤ −1), stable (= 0), or improved (≥ +1). Baseline mGPS was derived using CRP >5 mg/L and albumin <35 g/L. Overall survival (OS) and progression-free survival (PFS) were measured from ICI initiation. Cox regression assessed ΔMNA as a continuous variable, with prespecified adjustment for baseline mGPS, age, and sex.

Results: Median follow-up was 317 days (IQR 170.5–457.3). Death occurred in 16/54 patients and a PFS event in 27/54. Median OS was 706 days (1-year OS 71.3%; 2-year OS 49.4%), and median PFS was 337 days (1-year PFS 48.5%; 2-year PFS 15.1%). Nutritional categories shifted toward better status at 3 months (baseline normal/risk/malnutrition 29/23/2 vs 35/17/2 at follow-up). In univariable Cox models, higher ΔMNA was associated with longer survival (PFS: HR 0.72, 95% CI 0.56–0.94; P=0.016; OS: HR 0.56, 95% CI 0.38–0.82; P=0.003). In prespecified multivariable models, baseline mGPS was independently associated with OS (HR 1.86, 95% CI 1.00–3.45; P=0.048). After adjustment, the association between ΔMNA and outcomes remained in the protective direction but did not meet conventional statistical significance (OS: HR 0.67, 95% CI 0.42–1.06; P=0.086; PFS: HR 0.72, 95% CI 0.56–1.05; P=0.094). In an exploratory lung cancer subgroup (n=36), ΔMNA was associated with improved PFS (HR 0.62; P=0.008) and OS (HR 0.59; P=0.018). These subgroup findings should be interpreted as hypothesis-generating.

Conclusions: Baseline mGPS captured inflammatory risk at ICI start, while ΔMNA captured what happened to nutritional status over the first 3 months. In this cohort, nutritional improvement showed a consistent protective direction for OS and PFS even after accounting for baseline inflammation, and mGPS retained an

Received: January 10, 2026 **Accepted:** February 15, 2026 **Available Online:** February 18, 2026 **Published:** March 1, 2026

How to cite this article: Doğan T, Demiray AG, Yapar Taşköylü B, Hafızoğlu E, Yaren A, Gököz Doğu G. Early On-Treatment Nutritional Change and Baseline Inflammation Stratify Outcomes in Patients Receiving Immune Checkpoint Inhibitors. Eur Res J. 2026;12(3):397-406. doi: 10.18621/eurj.1860097

Corresponding author: Tolga Doğan, MD., Phone: +90 258 263 93 11, E-mail: dr_tolgadogan94@yahoo.com

Available Online at https://dergipark.org.tr/en/pub/eurj

independent association with OS. These findings support prospective studies that reassess nutrition during immunotherapy rather than relying on a single baseline measurement.
Keywords: Immune Checkpoint Inhibitor, Mini Nutritional Assessment–Short Form, Nutritional Status, Modified Glasgow Prognostic Score, Systemic Inflammation, Survival

Immune checkpoint inhibitors (ICIs) have reshaped the management of many advanced solid tumors; however, a substantial proportion of patients experience early progression or limited durable benefit despite appropriate indication and contemporary treatment algorithms. Beyond tumor-intrinsic factors, accumulating evidence highlights the relevance of the host milieu - particularly nutritional reserve, systemic inflammation, and functional status - in shaping tolerance, immune competence, and clinical outcomes during systemic therapy. Recent supportive-care guidance also recommends routine nutrition screening and timely management of malnutrition and cachexia as part of standard oncology practice [1].

In routine practice, nutritional risk must be captured with tools that are feasible and reproducible in real-world settings. The Mini Nutritional Assessment–Short Form (MNA-SF) is a brief, validated screening instrument that integrates appetite and weight loss history, mobility, acute disease/stress, neuropsychological problems, and anthropometrics, and has been widely adopted across oncology and geriatric-oncology settings [2]. While body mass index (BMI) remains a readily available measure, it can underestimate clinically meaningful nutritional decline and does not fully reflect metabolic stress or inflammation-driven catabolism. Therefore, coupling nutritional screening with inflammation-related biomarkers may provide a more clinically informative representation of the patient's trajectory under treatment.

Systemic inflammation is a well-established hallmark of cancer progression and treatment vulnerability. Simple C-reactive protein (CRP)–albumin–based scores such as the modified Glasgow Prognostic Score (mGPS) have shown robust, tumor-site–independent prognostic value and are frequently interpreted as surrogates of cancer-related inflammation and cachexia biology [3, 4]. Importantly, contemporary immuno-oncology literature increasingly supports the notion that on-treatment dynamics of these host-related markers may carry incremental information beyond baseline assessment, potentially complementing imaging-based response evaluation in selected settings [5].

Within the ICI era, inflammation–nutrition composite indices have also gained attention. The Prognostic Nutritional Index (PNI), derived from albumin and lymphocyte count, has been associated with outcomes in ICI-treated populations, and systematic reviews/meta-analyses in lung cancer cohorts similarly suggest that poorer pretreatment PNI relates to inferior survival with ICIs [6, 7]. However, most prior work has focused on single time-point measurements. From a clinical standpoint, early within-patient changes may better reflect treatment-period physiologic stress, evolving systemic inflammation, and nutritional resilience than baseline values alone. Whether early within-patient change in nutritional status during immune checkpoint inhibitor therapy provides prognostic information beyond baseline inflammation remains insufficiently defined.

Based on this rationale, we investigated whether early change in nutritional status during the first 3 months of ICI therapy, captured by ΔMNA (month-3 MNA minus baseline MNA), is associated with progression-free and overall survival in a real-world solid tumor cohort. In addition, we evaluated baseline systemic inflammation using mGPS as a pragmatic risk adjustment variable, aiming to contextualize ΔMNA within the inflammation–nutrition framework that underpins outcomes in advanced cancer.

METHODS

Study design and setting

This was a retrospective, single-center cohort study of adult patients with advanced or metastatic solid tumors treated with immune checkpoint inhibitors at our institution. Data were obtained from routine electronic medical records. For all time-to-event analyses, the index date ("time zero") was defined as

the date of immunotherapy initiation. The study was approved by the Pamukkale University Non-Interventional Clinical Research Ethics Committee (approval date: 30 December 2025; document number: E-60116787-020-806466). Due to the retrospective design and use of routinely collected clinical data, the requirement for written informed consent was waived in accordance with institutional regulations.

Cohort Assembly and Data Collection

Patients were identified by retrospective screening of institutional records. Because nutritional screening was not available for all immunotherapy-treated patients in routine practice, the analytic cohort was restricted to patients with documented MNA-SF at two prespecified time points: baseline at immunotherapy initiation and early follow-up at approximately 3 months during routine outpatient visits. Baseline and approximately 3-month laboratory parameters were extracted from the same clinical records (complete blood count and standard biochemistry panels, including CRP and albumin). For the present manuscript, analyses were prespecified to focus on early nutritional change (ΔMNA) and baseline systemic inflammation captured by mGPS; other available laboratory parameters were not included in the primary analytical models and are therefore not presented. No protocolized nutritional interventions such as enteral or parenteral nutrition were initiated as part of the study; patients received standard oncologic care and general dietary counseling. The final analytic cohort comprised 54 patients with MNA-SF recorded at both time points.

Nutritional assessment (MNA-SF) and definition of ΔMNA

Nutritional status was assessed using the MNA-SF; total score 0–14), a validated nutritional screening tool widely used in oncology and geriatric practice. MNA-SF consists of six domains evaluating recent food intake/appetite, involuntary weight loss, mobility, recent acute disease or psychological stress, neuropsychological problems, and BMI. In this cohort, BMI was available for all patients and was used for MNA-SF scoring. The total MNA-SF score was categorized as 0–7 (malnutrition), 8–11 (risk of malnutrition), and 12–14 (normal nutritional status). MNA-SF was recorded at baseline (at immunotherapy initiation) and at early follow-up approximately 3 months after treatment start, according to routine care. The primary exposure variable was early on-treatment change in nutritional status, defined as:

ΔMNA = MNA-SF (approximately 3 months) − MNA-SF (baseline)

Higher ΔMNA values indicate improvement in nutritional status during the early treatment period. For Kaplan–Meier analyses, ΔMNA was categorized as worsened (≤ -1), stable ($= 0$), or improved ($\geq +1$). Because ΔMNA requires a follow-up assessment, results were interpreted within a prespecified 3-month landmark analysis framework.

Systemic inflammation (modified Glasgow Prognostic Score)

Baseline systemic inflammation was quantified using the mGPS, calculated at immunotherapy initiation from CRP and albumin using predefined thresholds (CRP >5 mg/L; albumin <35 g/L). mGPS was coded as 0–2 according to standard criteria. Baseline mGPS was prespecified as the baseline risk covariate and was included in the primary multivariable models as an ordinal variable (per 1-point increase from 0 to 2).

Outcomes

Overall survival (OS) was defined as the time from immunotherapy initiation to death from any cause. Progression-free survival (PFS) was defined as the time from immunotherapy initiation to the first documented disease progression or death, whichever occurred first. Disease progression was determined based on routine radiologic assessment using cross-sectional imaging interpreted according to standard institutional practice, and/or documented clinical progression as recorded by the treating oncologist in the electronic medical record. Patients without an event were censored at the date of last clinical follow-up.

Statistical Analysis

Continuous variables are reported as median (interquartile range [IQR]) and categorical variables as number (%). OS and PFS were estimated using the Kaplan–Meier method and compared across ΔMNA

categories using the log-rank test. Associations between ΔMNA and outcomes were evaluated using Cox proportional hazards regression. ΔMNA was modeled as a continuous variable (per 1-point increase). The prespecified primary multivariable models adjusted for baseline mGPS, age, and sex. Exploratory subgroup analyses were performed in the lung cancer subgroup and were interpreted as hypothesis-generating. Missing data were handled using complete-case analysis for each model. All tests were two-sided and $P<0.05$ was considered statistically significant. Analyses were performed using IBM SPSS Statistics for Windows, version 26.0 (IBM Corp., Armonk, NY, USA).

RESULTS

Patient Characteristics

A total of 54 patients treated with immune checkpoint inhibitors were included in the analysis. Baseline demographic, clinical, and laboratory characteristics of the study population are summarized in Table 1. The median age at immunotherapy initiation was 64.0 years (IQR 59.0–69.8), and the majority of patients were male (83.3%). Lung cancer was the most common primary tumor (66.7%), followed by melanoma (16.7%) and renal cell carcinoma (7.4%). Other tumor types included head and neck cancer (n=2), esophageal cancer (n=1), cervical cancer (n=1), and mesothelioma (n=1). Overall, 37 (68.5%) patients had stage IV disease at the time of immunotherapy initiation.

Baseline inflammatory and nutritional parameters demonstrated substantial inter-individual variability. Median baseline C-reactive protein was 9.85 mg/L (IQR 3.01–36.04), and median serum albumin was 40.1 g/L (IQR 35.4–43.4). The baseline modified Glasgow Prognostic Score distribution was mGPS 0 in 28 (51.9%) patients, mGPS 1 in 16 (29.6%), and mGPS 2 in 10 (18.5%).

Nutritional Status and Early Change in MNA

Baseline nutritional status assessed by the MNA-SF showed that 29 (53.7%) patients were classified as nutritionally normal, 23 (42.6%) were at risk of malnutrition, and 2 (3.7%) were malnourished. At the

TABLE 1. Baseline Characteristics of Patients Receiving Immunotherapy

Variable	Overall cohort (n=54)
Age (years)	64.0 (59.0–69.8)
Sex	
Male	45 (83.3%)
Female	9 (16.7%)
Primary tumor	
Lung cancer	36 (66.7%)
Melanoma	9 (16.7%)
Renal cell carcinoma	4 (7.4%)
Head and neck cancer	2 (3.7%)
Esophageal cancer	1 (1.9%)
Cervical cancer	1 (1.9%)
Mesothelioma	1 (1.9%)
Stage IV disease	37 (68.5%)
Baseline CRP (mg/L)	9.85 (3.01–36.04)
Baseline albumin (g/L)	40.1 (35.4–43.4)
Baseline mGPS	
0	28 (51.9%)
1	16 (29.6%)
2	10 (18.5%)
Baseline MNA-SF category	
Normal (12–14)	29 (53.7%)
At risk (8–11)	23 (42.6%)
Malnourished (0–7)	2 (3.7%)
3-month MNA-SF category	
Normal (12–14)	35 (64.8%)
At risk (8–11)	17 (31.5%)
Malnourished (0–7)	2 (3.7%)
ΔMNA (MNA2 − MNA1)	0 (−1 to 1.75)
ΔMNA category	
Improved (≥ +1)	25 (46.3%)
Stable (0)	13 (24.1%)
Worsened (≤ −1)	16 (29.6%)

Data are shown as median (interquartile [IQR]) or n (%) where appropriate. CRP, C-reactive protein; mGPS, modified Glasgow Prognostic Score; MNA-SF, Mini Nutritional Assessment–Short Form; ΔMNA, change in MNA-SF between baseline and 3 months.

3-month follow-up assessment, the proportion of patients with normal nutritional status increased to 35 (64.8%), whereas the proportion at risk of malnutrition decreased to 17 (31.5%); the number of malnourished patients remained unchanged (n=2, 3.7%).

The early change in nutritional status, defined as ΔMNA (MNA at 3 months minus baseline MNA), demonstrated marked heterogeneity across the cohort. The median ΔMNA was 0 (IQR −1 to 1.75). Overall, 25 patients (46.3%) experienced an improvement in nutritional status (ΔMNA ≥ +1), 13 (24.1%) remained stable (ΔMNA = 0), and 16 (29.6%) showed a deterioration in nutritional status (ΔMNA ≤ −1).

Survival Outcomes

During a median follow-up of 317 days (IQR 170.5–457.3), 16 (29.6%) patients died and 27 (50.0%) patients experienced disease progression or death. Kaplan–Meier analysis demonstrated a median overall survival of 706 days, with estimated 1-year and 2-year overall survival rates of 71.3% and 49.4%, respectively. Median progression-free survival was 337 days, with corresponding 1-year and 2-year progression-free survival rates of 48.5% and 15.1%.

Kaplan–Meier Survival According to ΔMNA Categories

Kaplan–Meier analyses stratified by early change in nutritional status suggested a visual separation of survival curves across ΔMNA categories. Patients who experienced an improvement in nutritional status during the first 3 months of immunotherapy appeared to have longer overall survival compared with those with stable or worsened nutritional status (Figure 1). In parallel, progression-free survival also appeared to differ among ΔMNA groups, with the most favorable outcomes observed in patients with improved nutritional status and the least favorable outcomes in those with nutritional deterioration (Figure 2).

Association of ΔMNA with Progression-free and Overall Survival

In univariable Cox regression analyses, improvement in nutritional status over the first 3 months of immunotherapy was associated with favorable clinical outcomes. Each 1-point increase in ΔMNA was associated with longer progression-free survival (HR 0.72, 95% CI 0.56–0.94, P=0.016) and longer overall survival (HR 0.56, 95% CI 0.38–0.82, P=0.003) (Table 2).

After adjustment for age and sex, ΔMNA remained independently associated with improved overall survival (HR 0.63, 95% CI 0.41–0.97, P=0.038), whereas the association with progression-free survival showed a consistent direction but did not reach statistical significance (HR 0.76, 95% CI 0.56–1.02, P=0.071) (Table 2).

In the prespecified primary multivariable models incorporating baseline systemic inflammatory risk, ΔMNA remained directionally associated with

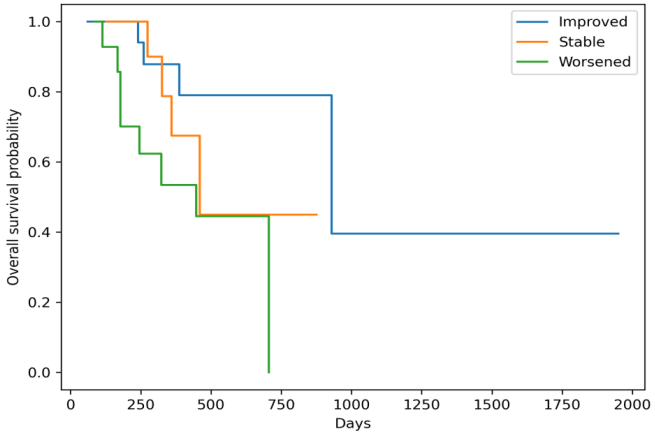

FIGURE 1. Kaplan–Meier curves showing overall survival stratified by early change in nutritional status (ΔMNA) categories (improved, stable, worsened).

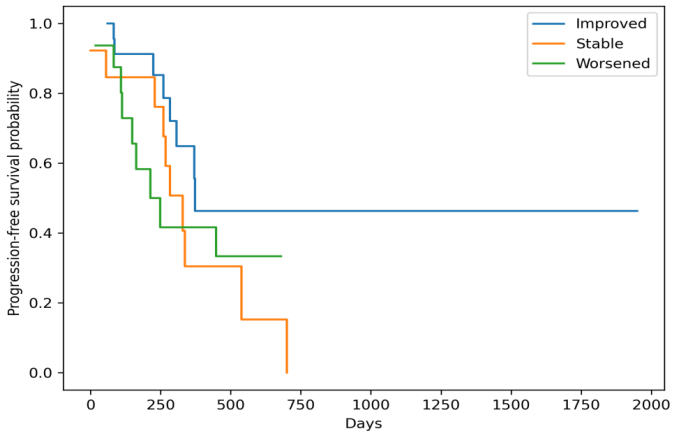

FIGURE 2. Kaplan–Meier curves illustrating progression-free survival according to early change in nutritional status (ΔMNA).

improved outcomes after adjustment for baseline mGPS, age, and sex. For progression-free survival, ΔMNA was associated with a trend toward improved outcomes (HR 0.77, 95% CI 0.56–1.05, P=0.094), while baseline mGPS was not significantly associated with progression-free survival (HR 1.10, 95% CI 0.67–1.81, P=0.702). In contrast, baseline mGPS was independently associated with higher mortality risk in overall survival analyses (HR 1.86, 95% CI 1.00–3.43, P=0.048) (Table 2).

Exploratory Lung Cancer Subgroup Analysis

In the lung cancer subgroup (n=36), ΔMNA remained associated with clinical outcomes in univariable analyses. Each 1-point increase in ΔMNA was associated with improved progression-free survival (HR 0.62, 95% CI 0.43–0.88, P=0.008) and improved overall survival (HR 0.59, 95% CI 0.38–0.91, P=0.018) (exploratory, univariable analyses). These subgroup analyses were exploratory and not powered for definitive inference.

DISCUSSION

In this retrospective cohort of patients treated with immune checkpoint inhibitors, early improvement in nutritional status during the first 3 months - captured by ΔMNA - was consistently associated with more favorable survival outcomes. Consistent with this, each 1-point increase in ΔMNA was associated with longer PFS (HR 0.72, 95% CI 0.56–0.94; P=0.016) and longer OS (HR 0.56, 95% CI 0.38–0.82; P=0.003) in univariable analyses. This observation is consistent with broader oncology evidence indicating that MNA-based nutritional assessment relates to clinically meaningful endpoints, including overall survival and disease progression, although most prior work has focused on single time-point evaluations rather than early trajectories [8]. In the immunotherapy setting, accumulating data support the concept that host immune–nutritional reserve influences treatment benefit: pretreatment PNI has been associated with response and survival in NSCLC patients receiving ICIs [9]. Moreover, worsening nutritional status before immunotherapy initiation - reflected by low PNI and declining pretreatment BMI - may correlate more closely with immunotherapy outcomes than baseline BMI alone, reinforcing the clinical relevance of dynamic nutritional change [10]. Overall, our results are consistent with prior reports and support ΔMNA as an early on-treatment marker in patients receiving ICIs [8–10].

A key element of our prespecified approach was

TABLE 2. Cox Proportional Hazards Analyses for Progression-Free and Overall Survival

A. Progression-free survival (PFS)						
Variable	**HR (Uni)**	**95% CI**	**P-value**	**HR (Multi)**	**95% CI**	**P-value**
ΔMNA (per 1-point)	0.72	0.56–0.94	**0.016**	0.77	0.56–1.05	0.094
Baseline mGPS (0–2)	1.10	0.67–1.81	0.702	1.10	0.67–1.81	0.702
Age (per year)	—	—	—	1.01	0.98–1.04	0.42
Sex (male vs female)	—	—	—	1.12	0.54–2.33	0.76
B. Overall survival (OS)						
Variable	**HR (Uni)**	**95% CI**	**P-value**	**HR (Multi)**	**95% CI**	**P-value**
ΔMNA (per 1-point)	0.56	0.38–0.82	**0.003**	0.67	0.42–1.06	0.086
Baseline mGPS (0–2)	1.97	1.12–3.46	**0.018**	1.86	1.00–3.43	**0.048**
Age (per year)	—	—	—	1.04	1.00–1.09	0.06
Sex (male vs female)	—	—	—	1.28	0.51–3.18	0.59

HR, hazard ratio; CI, confidence interval; ΔMNA, change in Mini Nutritional Assessment–Short Form; mGPS, modified Glasgow Prognostic Score; PFS, progression-free survival; OS, overall survival, Uni, univariable.

Multivariable (multi) models adjusted for age and sex, with baseline mGPS prespecified as a baseline risk covariate. Baseline mGPS was modeled as an ordinal variable (per 1-point increase from 0 to 2). Statistically significant P-values are shown in bold.

to interpret ΔMNA in the context of baseline systemic inflammation, using the mGPS as a pragmatic risk-adjustment variable. Because mGPS integrates CRP and albumin, it reflects the inflammatory–nutritional milieu that can influence both tumor biology and host resilience during immunotherapy. In ICI-treated populations, baseline mGPS has repeatedly demonstrated prognostic value across tumor types, including advanced NSCLC treated with anti–PD-1 agents and metastatic renal cell carcinoma treated with ICI-based regimens [11,12]. Importantly, a recent systematic review and meta-analysis of patients with advanced cancers receiving ICIs further supported the overall prognostic utility of GPS/mGPS, with higher scores associated with worse survival outcomes [13]. Clinically, the message is straightforward: baseline mGPS captured baseline risk - particularly for overall survival - whereas ΔMNA added dynamic prognostic information during the first months of treatment. In the prespecified multivariable model including age and sex, baseline mGPS was independently associated with OS (HR 1.86, 95% CI 1.00–3.43; P=0.048), whereas ΔMNA showed a directionally protective association that did not reach conventional statistical significance (HR 0.67, 95% CI 0.42–1.06; P=0.086).

This pattern is biologically plausible. Cancer-related malnutrition and cachexia are increasingly conceptualized not merely as inadequate intake but as a systemic inflammatory syndrome with downstream metabolic and functional consequences, which may shape both treatment tolerance and long-term outcomes [14]. In this context, reduced nutritional and muscle reserve has been linked to inferior efficacy of immune checkpoint inhibitors: a meta-analysis in NSCLC reported that pre-immunotherapy sarcopenia was associated with poorer survival and less favorable response-related outcomes in patients receiving ICIs [15]. Moreover, dynamic deterioration during ICI therapy appears clinically meaningful; in a multicenter cohort of recurrent/metastatic head and neck cancer, the combination of baseline cachexia and early weight loss during immune checkpoint inhibition was associated with worse overall survival, independent of PD-L1 expression.[16] Collectively, these data provide biological and clinical context for interpreting ΔMNA as an accessible on-treatment marker in our cohort, while acknowledging that - given the retrospective design - ΔMNA may partly reflect underlying disease trajectory and residual confounding rather than a purely causal pathway [14–16].

In our cohort, the prognostic separation by early nutritional trajectory was clinically meaningful: patients with improved ΔMNA during the first 3 months showed better survival patterns on Kaplan–Meier analyses, and ΔMNA remained directionally consistent in Cox models after adjustment for baseline inflammatory risk, with effect estimates favoring improved survival. Notably, in the prespecified multivariable model, baseline mGPS was independently associated with OS (HR 1.86, P=0.048), whereas ΔMNA showed a protective trend (HR 0.67, P=0.086). This supports a practical message for routine care—nutritional status during immunotherapy should be monitored as a dynamic process rather than a single baseline attribute. Current guidance is aligned with this approach: ESPEN recommendations emphasize routine screening with repeated reassessment and stepwise, individualized nutritional interventions (from counseling and oral supplementation to enteral/parenteral strategies when indicated), explicitly recognizing that nutritional risk and needs evolve during active treatment [17]. Likewise, ESMO cachexia guidance highlights early identification of nutrition-impact symptoms and functional decline, and advocates timely, multimodal supportive management tailored to disease trajectory and expected survival [1]. Interpreted alongside these frameworks, ΔMNA offers an accessible way to operationalize “early trajectory” in daily practice—baseline assessment followed by a planned reassessment around 3 months—to identify patients at risk of early nutritional deterioration who may benefit from prioritized supportive interventions. While it remains uncertain whether nutritional optimization directly augments antitumor efficacy of immune checkpoint inhibitors, nutritional care before and during treatment is consistently regarded as clinically relevant for maintaining functional reserve, improving tolerance, and minimizing avoidable treatment interruptions, which plausibly contributes to better real-world outcomes—consistent with the patterns observed in our cohort [18].

In our cohort, the association of ΔMNA with overall survival appeared more robust than with progression-free survival in the fully adjusted models, and this pattern is not unexpected. In contemporary

oncology trials and real-world datasets, early endpoints such as response rate or PFS often show an imperfect and sometimes discordant relationship with OS, particularly when post-progression management, treatment sequencing, and supportive care influence survival after progression [19, 20]. OS captures not only the timing of progression but also post-progression survival, which can be shaped by subsequent therapies, functional reserve, and the ability to continue treatment - domains plausibly linked to nutritional trajectory and symptom burden. Early nutritional improvement (higher ΔMNA) may reflect better host resilience and treatment sustainability, which can translate more clearly into OS than into PFS when imaging intervals, progression ascertainment, and early-event dynamics dilute the PFS signal. Finally, because ΔMNA is a post-baseline measure, analyses that compare groups defined by on-treatment change are vulnerable to guarantee-time (immortal-time) bias unless handled appropriately; using a prespecified landmark framework (e.g., at 3 months) is a methodologically sound approach to mitigate this bias and support interpretable survival comparisons [21].

Strengths and Limitations

Several strengths of this study deserve mention. We evaluated nutritional status longitudinally using a standardized instrument and focused on an early, clinically actionable window (baseline to 3 months), which aligns with routine follow-up schedules in immunotherapy practice. We also prespecified a conservative analytic framework that accounted for baseline systemic inflammation (mGPS), and we explicitly considered guarantee-time bias and prespecified a 3-month landmark framework for interpreting analyses of an on-treatment change variable. At the same time, important limitations should be acknowledged. The retrospective, single-center design and modest sample size increase the risk of selection bias and residual confounding, and the cohort's heterogeneous tumor composition limits tumor-specific inference despite the consistent signal observed in the lung cancer subgroup. In addition, the heterogeneous tumor composition and the limited number of outcome events preclude robust tumor-specific conclusions. Accordingly, the observed associations should be interpreted as reflecting overall host-related dynamics during immunotherapy rather than definitive tumor-type–specific effects. ΔMNA may partly capture underlying disease trajectory and symptom burden (e.g., early progression-related anorexia or treatment-limiting toxicity) rather than a purely modifiable pathway, and we did not have a standardized nutritional intervention protocol to determine whether proactive nutritional optimization can improve outcomes. These findings should be considered hypothesis-generating and require prospective validation. Future multicenter studies should confirm the prognostic utility of ΔMNA within more homogeneous tumor cohorts, integrate objective body-composition and functional measures, and test whether structured early nutritional interventions can translate into improved treatment tolerance and survival. In summary, early nutritional trajectory (ΔMNA), considered alongside baseline inflammatory risk (mGPS), may offer a practical framework for risk stratification during immunotherapy and highlights the value of reassessing nutrition early in treatment [22].

CONCLUSION

In this real-world cohort of patients treated with immune checkpoint inhibitors, baseline mGPS reflected inflammatory risk at treatment initiation, whereas early change in nutritional status (ΔMNA) provided additional insight into the patient's trajectory during the first months of therapy. Patients who experienced nutritional improvement within the first three months tended to have more favorable overall and progression-free survival patterns, even after accounting for baseline inflammatory burden. These findings highlight the clinical importance of reassessing nutritional status during immunotherapy rather than relying solely on a single baseline measurement. Larger prospective studies are needed to confirm these observations and to determine whether early nutritional optimization can improve long-term outcomes.

Ethics Approval and Consent to Participate

This study was approved by the Pamukkale

University Non-Interventional Clinical Research Ethics Committee (meeting date: 30 December 2025; decision no: 24; document no: E-60116787-020-806466). The Committee determined that the study was ethically appropriate. All procedures were conducted in accordance with the Declaration of Helsinki and its later amendments. Given the retrospective design and the use of de-identified routinely collected data, the requirement for written informed consent was waived in accordance with institutional regulations.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author,

Authors' Contribution

Study Conception: TD, AY, GGD; Study Design TD, EH; Supervision: AY, GGD; Funding: N/A; Materials: TD, AGD, BYT; Data Collection and/or Processing: TD, AGD, BYT; Statistical Analysis and/or Data Interpretation: TD, AGD, BYT; Literature Review: TD, EH; Manuscript Preparation: TD; and Critical Review: AY, GGD, EH.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors sincerely thank all patients who participated in this study and the clinical staff involved in their care.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Arends J, Strasser F, Gonella S, et al; ESMO Guidelines Committee. Electronic address: clinicalguidelines@esmo.org. Cancer cachexia in adult patients: ESMO Clinical Practice Guidelines. ESMO Open. 2021;6(3):100092. doi: 10.1016/j.esmoop.2021.100092.
2. Kaiser MJ, Bauer JM, Ramsch C, et al; MNA-International Group. Validation of the Mini Nutritional Assessment short-form (MNA-SF): a practical tool for identification of nutritional status. J Nutr Health Aging. 2009;13(9):782-788. doi: 10.1007/s12603-009-0214-7.
3. Proctor MJ, Morrison DS, Talwar D, et al. A comparison of inflammation-based prognostic scores in patients with cancer. A Glasgow Inflammation Outcome Study. Eur J Cancer. 2011;47(17):2633-2641. doi: 10.1016/j.ejca.2011.03.028.
4. McMillan DC. The systemic inflammation-based Glasgow Prognostic Score: a decade of experience in patients with cancer. Cancer Treat Rev. 2013;39(5):534–540. doi: 10.1016/j.ctrv.2012.08.003.
5. Saal J, Bald T, Eckstein M, et al. Integrating On-Treatment Modified Glasgow Prognostic Score and Imaging to Predict Response and Outcomes in Metastatic Renal Cell Carcinoma. JAMA Oncol. 2023 Aug 1;9(8):1048–1055. doi: 10.1001/jamaoncol.2023.1822.
6. Ni L, Huang J, Ding J, et al. Prognostic Nutritional Index Predicts Response and Prognosis in Cancer Patients Treated With Immune Checkpoint Inhibitors: A Systematic Review and Meta-Analysis. Front Nutr. 2022;9:823087. doi: 10.3389/fnut.2022.823087.
7. Oku Y, Toyokawa G, Wakasu S, et al. Impact of the pretreatment prognostic nutritional index on the survival after first-line immunotherapy in non-small-cell lung cancer patients. Cancer Med. 2023;12(13):14327-14336. doi: 10.1002/cam4.6110.
8. Torbahn G, Strauss T, Sieber CC, Kiesswetter E, Volkert D. Nutritional status according to the mini nutritional assessment (MNA)® as potential prognostic factor for health and treatment outcomes in patients with cancer - a systematic review. BMC Cancer. 2020;20(1):594. doi: 10.1186/s12885-020-07052-4.
9. Shoji F, Takeoka H, Kozuma Y, et al. Pretreatment prognostic nutritional index as a novel biomarker in non-small cell lung cancer patients treated with immune checkpoint inhibitors. Lung Cancer. 2019;136:45–51. doi: 10.1016/j.lungcan.2019.08.006.
10. Johannet P, Sawyers A, Qian Y, et al. Baseline prognostic nutritional index and changes in pretreatment body mass index associate with immunotherapy response in patients with

advanced cancer. J Immunother Cancer. 2020;8(2):e001674. doi: 10.1136/jitc-2020-001674.
11. Freitas C, Jacob M, Tavares N, et al. Modified Glasgow Prognostic Score predicts survival among advanced non-small cell lung carcinoma patients treated with anti-PD1 agents. Anticancer Drugs. 2021;32(5):567–574. doi: 10.1097/CAD.0000000000001060.
12. Brown JT, Liu Y, Shabto JM, et al. Modified Glasgow Prognostic Score associated with survival in metastatic renal cell carcinoma treated with immune checkpoint inhibitors. J Immunother Cancer. 2021;9(7):e002851. doi: 10.1136/jitc-2021-002851.
13. Zhang L, Lin L, Ni J, Ling T, Huang L. Prognostic value of the pretreatment Glasgow prognostic score or modified Glasgow prognostic score in patients with advanced cancer receiving immune checkpoint inhibitors: a systematic review and meta-analysis. Oncol Lett. 2025;30(1):323. doi: 10.3892/ol.2025.15069.
14. McGovern J, Dolan RD, Skipworth RJE, Laird BJ, McMillan DC. Cancer cachexia: a nutritional or a systemic inflammatory syndrome? Br J Cancer. 2022;127(3):379–382. doi: 10.1038/s41416-022-01826-2.
15. Ren B, Shen J, Qian Y, Zhou T. Sarcopenia as a determinant of the efficacy of immune checkpoint inhibitors in non-small cell lung cancer: a meta-analysis. Nutr Cancer. 2023;75(2):685–695. doi: 10.1080/01635581.2022.2153879.
16. Willemsen ACH, De Moor N, Van Dessel J, et al. The predictive and prognostic value of weight loss and body composition prior to and during immune checkpoint inhibition in recurrent or metastatic head and neck cancer patients. Cancer Med. 2023;12(7):7699–7712. doi: 10.1002/cam4.5522.
17. Muscaritoli M, Arends J, Bachmann P, et al. ESPEN practical guideline: Clinical Nutrition in cancer. Clin Nutr. 2021;40(5):2898–2913. doi: 10.1016/j.clnu.2021.02.005.
18. Aprile G, Basile D, Giaretta R, et al. The Clinical Value of Nutritional Care before and during Active Cancer Treatment. Nutrients. 2021;13(4):1196. doi: 10.3390/nu13041196.
19. Merino M, Kasamon Y, Theoret M, et al. Irreconcilable Differences: The Divorce Between Response Rates, Progression-Free Survival, and Overall Survival. J Clin Oncol. 2023;41(15):2706–2712. doi: 10.1200/JCO.23.00225.
20. Hess LM, Brnabic A, Mason O, Lee P, Barker S. Relationship between Progression-free Survival and Overall Survival in Randomized Clinical Trials of Targeted and Biologic Agents in Oncology. J Cancer. 2019;10(16):3717-3727. doi: 10.7150/jca.32205.
21. Zabor EC, Assel M. On the need for landmark analysis or time-dependent covariates. J Urol. 2023;209(6):1060-1062. doi: 10.1097/JU.0000000000003459.
22. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-1457. doi: 10.1016/S0140-6736(07)61602-X.