ISSN 2146-6505 • EISSN 2147-1894

Indexed in Web of Science

jarem

JOURNAL OF ACADEMIC RESEARCH IN MEDICINE

Original Investigations

VOLUME: 16 | ISSUE: 2 | August 2026

UNIVERSITY OF HEALTH SCIENCES TÜRKİYE GAZİOSMANPAŞA TRAINING AND RESEARCH HOSPITAL

galenos yayınevi

jarem.org

JAREM

JOURNAL OF ACADEMIC RESEARCH IN MEDICINE

• **Owned by on behalf of the University of Health Sciences Türkiye, Gaziosmanpaşa Training and Research Hospital:** Prof. Dr. Ömer N. Develioğlu
• **Editor in Chief:** Prof. Dr. Ömer N. Develioğlu
• **Publication Type:** Periodical

Jarem
JOURNAL OF ACADEMIC RESEARCH IN MEDICINE

International Editorial Board

JOURNAL OF ACADEMIC RESEARCH IN MEDICINE

International Editorial Board

First Past Editor in Chief of the journal: Prof. Dr. Barış Nuhoğlu

İstanbul Yeniyüzyıl University Gaziosmanpaşa Hospital Head of Department of Urology (The editor-in-chief duty, which started in September 2011, was transferred to Ömer N Develioğlu in March 2016)

"Please refer to the journal's webpage (https://www.jarem.org/home) for "Journal Policies", "Instructions to Authors" and "About Us".

The Journal of Academic Research in Medicine and/or its editors are members of ICMJE, COPE, WAME, CSE and EASE, and follow their recommendations. Journal of Academic Research in Medicine is indexed in Web of Science-Emerging Sources Citation Index, TÜBİTAK ULAKBİM TR Dizin, EBSCO, Gale, CINAHL, J Gate, Türk Medline, Embase and CAB International (CABI).

The journal is published online.

Owner: *University of Health Sciences Türkiye, Gaziosmanpaşa Training and Research Hospital*

Responsible Manager: *Ömer N. Develioğlu*

galenos yayınevi

Publisher Contact
Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1
34093 İstanbul, Türkiye
Phone: +90 (530) 177 30 97 / +90 (539) 307 32 03
E-mail: info@galenos.com.tr/yayin@galenos.com.tr
Web: www.galenos.com.tr
Publisher Certificate Number: 14521

Online Publishing Date: August 2026

ISSN: 2146-6505 E-ISSN: 2147-1894

International periodical journal published three times in a year.

Jarem
JOURNAL OF ACADEMIC RESEARCH IN MEDICINE

Contents

DOI: 10.4274/jarem.galenos.2026.07752
J Acad Res Med 2026;16(2):56-62

Oncologic Outcomes and HPV Persistence After LEEP and Conization for Cervical Intraepithelial Lesions: Association of Cone Depth with Positive Surgical Margins in a Single-center 10-year Retrospective Cohort

Enes Serhat Coşkun[1], Fatma Ketenci Gencer[1], Havva Betül Bacak[1], Alperen Aksan[1], Süleyman Salman[2]

[1]University of Health Sciences Türkiye, Gaziosmanpaşa Training and Research Hospital, Department of Obstetrics and Gynecology, İstanbul, Türkiye
[2]Medical Park TEM Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Türkiye

Cite this article as: Coşkun ES, Ketenci Gencer F, Bacak HB, Aksan A, Salman S. Oncologic outcomes and HPV persistence after LEEP and conization for cervical intraepithelial lesions: association of cone depth with positive surgical margins in a single-center 10-year retrospective cohort. J Acad Res Med. 2026;16(2):56-62

ABSTRACT

Objective: Positive surgical margins (PSM) after excisional treatment for cervical intraepithelial lesions affect surveillance and repeat treatment. Cone depth is modifiable, but deeper excision may increase obstetric risk. We assessed PSM prevalence, associated factors, and colposcopy-histology agreement.

Methods: We conducted a single-center retrospective cohort study of patients who underwent loop electrosurgical excision procedure (LEEP) or cervical conization between 7 January 2015 and 26 September 2025. PSM was defined as margin involvement at any surgical margin. Factors associated with PSM were evaluated using multivariable bias-reduced (Firth) logistic regression including cone depth (cm, continuous), high-grade squamous intraepithelial lesion (HSIL)-or-higher grade histology on the excision specimen (yes/no), and procedure type (LEEP vs. conization), using a complete-case approach.

Results: Among 246 patients, 92 (37.4%) underwent LEEP and 154 (62.6%) conization. Median cone depth was 1.6 cm (interquartile range 1.0-2.5) and was greater with conization than LEEP (2.0 vs. 1.5 cm; $p<0.001$). PSM occurred in 20/241 (8.3%) and did not differ between LEEP and conization (6.6% vs. 9.3%). Reconization was performed in 57/243 (23.5%) and was more frequent after conization (30.5% vs. 12.0%). Recurrence occurred in 16/246 (6.5%). In the Firth model (complete cases n=215; 18 PSM events), greater cone depth was associated with lower odds of PSM (adjusted odds ratio 0.46 per 1-cm increase; 95% confidence interval 0.24-0.88; $p=0.018$), whereas HSIL-or-higher histology and procedure type were not statistically significant. Colposcopy-histology agreement was low (Cohen's $\kappa=0.168$). Follow-up human papillomavirus (HPV) results were available for a small subset and are descriptive.

Conclusion: Greater cone depth was associated with reduced odds of PSM. Given limited event numbers and non-standardized follow-up, non-significant associations and HPV outcomes should be interpreted cautiously. Tailoring excision depth to lesion location may reduce margin positivity while balancing obstetric risks.

Keywords: Cone depth, positive surgical margins, LEEP, cervical conization, CIN, HSIL

INTRODUCTION

Cervical intraepithelial neoplasia (CIN) represents a well-characterized spectrum of precursor lesions in the pathogenesis of cervical cancer and remains a central target of screening programs that have contributed to reducing the incidence of invasive disease (1-3). The overarching aim of clinical management is to eradicate high-grade lesions, particularly high-grade squamous intraepithelial lesion (HSIL)/CIN2 to CIN3, while avoiding unnecessary overtreatment and minimizing cervical tissue loss that may compromise fertility and adversely affect obstetric outcomes (4-6). Striking this balance is especially critical in women of reproductive age and underscores the need for individualized treatment strategies.

Excisional procedures for high-grade cervical lesions have both diagnostic and therapeutic value. Loop electrosurgical excision

ORCID IDs of the authors: E.S.C. 0000-0002-6166-5458; F.K.G. 0000-0002-6076-2563; H.B.B. 0000-0001-8516-5493; A.A. 0000-0001-7623-3589; S.S. 0000-0001-7090-6105

Corresponding Author: Enes Serhat Coşkun, MD;
E-mail: enesserhatcoskun@gmail.com

Received Date: 14.02.2026 **Accepted Date:** 06.05.2026
Epub: 12.06.2026
Publication Date: 07.08.2026

J Acad Res Med 2026;16(2):56-62

procedure (LEEP) and cervical conization enable complete removal of the lesion for histopathological assessment, allowing exclusion of occult invasion, delineation of lesion margins, and provision of definitive treatment in the same setting (7,8). However, outcomes after excisional treatment are influenced by multiple factors, including extension into the endocervical canal, transformation zone type, colposcopic visibility, patient age, human papillomavirus (HPV)-related biological behavior, coexisting glandular abnormalities, and technical characteristics of the excision such as depth and volume, all of which may affect treatment success and recurrence risk (8-11).

Identification of positive surgical margins (PSM) on post-excisional pathology is among the outcomes that most strongly influences subsequent clinical management. PSM may be associated with a higher likelihood of residual disease, necessitate intensified surveillance, and, in selected cases, prompt consideration of repeat excision or escalation of treatment (11,12). Nevertheless, margin positivity does not invariably indicate residual disease, and negative margins do not completely eliminate recurrence risk; therefore, defining factors that predict PSM is clinically important (12). In this context, cone depth (i.e., the amount of tissue removed) represents a double-edged sword: insufficient depth may increase the risk of margin positivity, whereas overly deep excision has been linked to adverse obstetric outcomes such as cervical insufficiency and preterm birth (5,6,13). Accordingly, evaluating excision depth in relation to PSM using real-world clinical data may inform procedural planning, quality improvement, and patient counseling.

The diagnostic pathway leading to excisional treatment is typically guided by screening cytology, HPV testing, colposcopic assessment, and colposcopy-directed biopsy (7). In routine clinical practice, discordance may be observed between the colposcopic impression and/or biopsy findings and the final histology of the excision specimen (14,15). Such discordance can result from lesion heterogeneity, sampling limitations of biopsy, endocervical location of the transformation zone, multifocal disease, and interobserver variability in pathological interpretation (16,17). Characterizing colposcopy-to-final histology agreement under real-world conditions is therefore important for evaluating diagnostic performance and identifying opportunities for improvement.

In this single-center retrospective cohort, we aimed to determine the prevalence of PSM and to identify clinical and procedural factors associated with PSM, with particular emphasis on cone depth, among patients undergoing LEEP or cervical conization. In addition, we assessed the agreement between colposcopic impression and/or biopsy results and the final histology of the excision specimen. These findings may help inform procedural targets for excision planning and support evidence-based approaches to post-treatment surveillance and repeat-treatment decision-making.

METHODS

This single-center retrospective cohort study was conducted at University of Health Sciences Türkiye, Gaziosmanpaşa Training and Research Hospital. Clinical and pathological records of consecutive patients who underwent LEEP or cervical conization for suspected or confirmed cervical intraepithelial lesions between 07 January 2015 and 26 September 2025 were reviewed. Ethical approval was obtained from the University of Health Sciences Türkiye, Gaziosmanpaşa Training and Research Hospital Non-Interventional Ethics Committee (decision no: 134, date: 15.10.2025). Prior to analysis, all data were anonymized by removal of personal identifiers, and the study was carried out in accordance with the principles of the Declaration of Helsinki.

Demographic characteristics including age, parity, and smoking status were obtained from electronic medical records, together with preoperative screening cytology [normal/ASC-US/low-grade squamous intraepithelial lesion (LSIL)/HSIL/atypical squamous cells-cannot exclude HSIL (ASC-H)/AGC], preoperative HPV status (negative/positive), colposcopy-directed biopsy results [CIN1/CIN2/CIN3/carcinoma *in situ* (CIS)], and endocervical curettage (ECC) findings (not performed/benign/LSIL/HSIL/CIS/squamous cell carcinoma/adenocarcinoma). The type of procedure (LEEP or conization) and cone depth (cm) were recorded as operative variables. Final excision histopathology was categorized as benign, LSIL, HSIL, CIS, squamous cell carcinoma, or adenocarcinoma; for analytical purposes, a composite "CIS+" category was additionally defined to include CIS and invasive carcinomas (squamous cell carcinoma and adenocarcinoma). For the regression model, "HSIL or higher-grade histology" was defined as HSIL, CIS, or invasive carcinoma (squamous cell carcinoma or adenocarcinoma) on the excision specimen.

The primary outcome was PSM, defined as the presence of margin involvement at any surgical margin of the excision specimen. Secondary outcomes included performance of repeat excision (reconization; yes/no), recurrence (yes/no), procedure-related complications (none/bleeding/infection), and available HPV results at 12 and 24 months of follow-up (negative/positive). The relationship between colposcopic impression and final histology was examined using cross-tabulation and summarized as row percentages. Recurrence was defined as detection of histologically confirmed HSIL (CIN2+) or a higher-grade lesion during follow-up after the index excisional procedure, based on colposcopy and/or biopsy; given the retrospective, record-based design, follow-up protocols may have varied according to clinician judgment.

Excisional treatment was not routinely performed in patients with CIN1 identified on colposcopy-directed biopsy; however, in this series, all CIN1 cases that underwent excision were positive for HPV-16 and/or HPV-18 on genotyping. Accordingly, excision in the CIN1 subgroup was selected to exclude and/or treat the possibility of an underlying higher-grade lesion in the presence of high-risk HPV despite low-grade histology.

Statistical Analysis

Continuous variables were summarized as mean ± standard deviation or median [interquartile range (IQR)], as appropriate. Normality was assessed using the Shapiro-Wilk test. Between-group comparisons were performed using the Student's t-test or the Mann-Whitney U test for continuous variables and the chi-square test or Fisher's exact test, as appropriate, for categorical variables. A multivariable Firth penalized logistic regression model (bias-reduced) was used to evaluate factors associated with PSM. The prespecified model included cone depth (cm, continuous), HSIL or higher-grade histology (yes/no), and procedure type (LEEP vs. conization). Predictors were selected a priori to construct a parsimonious model given the limited number of PSM events, in line with events-per-variable considerations, and to reduce overfitting. Accordingly, the multivariable model was limited to three prespecified predictors. The selected variables were chosen to represent the key domains relevant to the study question: a modifiable excision parameter (cone depth), disease severity on the excision specimen (HSIL or higher-grade histology), and excisional technique (LEEP vs. conization). Additional clinically relevant variables (e.g., age, smoking, and preoperative HPV/cytology) were not included because of the limited number of events and variable-specific missingness, and residual confounding is therefore acknowledged as a limitation. Results are reported as adjusted odds ratios (aOR) with 95% confidence intervals (CI). Owing to missing data, analyses were conducted using a complete-case approach. The extent of missingness for each variable is summarized in Supplementary Table 1. For the PSM model, complete cases required non-missing PSM status, cone depth, final excision histopathology, and procedure type (n=215; 18 PSM events). For key between-group comparisons, absolute differences in proportions with 95% CIs were also reported to aid clinical interpretation. CIs for absolute differences were calculated using standard methods for two independent proportions. Statistical significance was defined as a two-sided $p<0.05$. Descriptive statistics and group comparisons were performed using IBM SPSS Statistics (v26.0; IBM Corp., Armonk, NY, USA), and Firth penalized regression analyses were conducted using R (R Foundation for Statistical Computing, Vienna, Austria) with the logistf package.

RESULTS

During the study period, 246 patients were evaluated; 92 (37.4%) underwent LEEP and 154 (62.6%) underwent conization. The mean age was 42.1±9.6 years, with no significant difference between procedure groups (LEEP 41.0±9.6 vs. conization 42.8±9.5; p=0.142). Median parity was 2 (IQR 1-3) in both groups (p=0.288). Among the 241 patients with available smoking data, 42.7% were current smokers, with no difference between groups (p=0.992) (Table 1).

Preoperative cytology was available for 208 patients and its distribution differed between groups (p=0.025). High-grade cytology (HSIL/ASC-H/AGC) was more frequent in the conization group than in the LEEP group (36.1% vs. 17.3%) (Table 1). The absolute difference was 18.8 percentage points (95% CI 1.0-34.1). Preoperative HPV results were recorded for 166 patients; within this subgroup, HPV positivity was 98.2% and did not differ by procedure type (p=1.000). Colposcopy-directed biopsy results were available for 237 patients and showed a marked between-group difference ($p<0.001$): CIN1 was more common in the LEEP group (43.0%), whereas CIN3 was more frequently observed in the conization group (54.3%) (Table 1). ECC results were documented in 245 patients and were comparable between groups (p=0.352).

Operative and postoperative outcomes are summarized in Table 2. Among the 218 patients with available cone depth data, the median cone depth was 1.6 cm (IQR 1.0-2.5); excisions were deeper in the conization group [2.0 cm (1.2-2.5)] than in the LEEP group [1.5 cm (0.8-2.0); $p<0.001$]. The overall complication rate was low (3/246, 1.2%), and all complications were bleeding events with no infections recorded; complication rates did not differ by procedure type (p=1.000).

Final excision histopathology was available for 243 patients and differed significantly between groups ($p<0.001$). Benign changes and LSIL were more common after LEEP (benign 34.1%, LSIL 37.4%), whereas HSIL and CIS were more frequently identified in the conization group (HSIL 51.3%, CIS 10.5%). PSM could be assessed in 241 patients and were observed in 20 (8.3%); the PSM rate did not differ between LEEP (6.6%) and conization (9.3%) (p=0.612). The absolute difference was 2.7 percentage points (95% CI -8.0 to 12.0). Reconization status was available for 243 patients, and 57 (23.5%) underwent repeat excision, which was more frequent in the conization group than in the LEEP group (30.5% vs. 12.0%; p=0.002). The absolute difference was 18.5 percentage points (95% CI 3.5-31.4). Recurrence occurred in 16 patients (6.5%); although the recurrence rate was higher after conization, the difference was not statistically significant (8.4% vs. 3.3%; p=0.184) (Table 2). The absolute difference was 5.2 percentage points (95% CI -4.2 to 12.8). Follow-up HPV results were available for a limited number of patients; HPV negativity was 18/24 (75.0%) at 12 months and 11/14 (78.6%) at 24 months, with no significant differences between procedure groups (p=0.665 and p=0.209, respectively). These follow-up comparisons should be interpreted as descriptive/exploratory because follow-up testing was not standardized and HPV results were available only for a small subset of patients.

The association between colposcopic impression and final histology is presented as row percentages in Table 3. Among cases with a colposcopic impression of CIN1, final histology showed benign changes in 37.8% and LSIL in 48.6%, whereas HSIL was identified in 13.5%. In cases with a CIN3 impression, HSIL was the most common final diagnosis (54.9%), and CIS+ was observed in 14.3%. In the small subgroup with a colposcopic impression of CIS, the proportion of CIS+ on final histology was 66.7% (Table 3). In an ordinal comparison between colposcopic impression (CIN1, CIN2, CIN3, CIS) and final histology categories (benign/LSIL/HSIL/CIS+), agreement was observed in 49.6% of cases, whereas 10.2%

J Acad Res Med 2026;16(2):56-62

Table 1. Baseline characteristics and preoperative findings by procedure (LEEP vs. conization)

Variable	Overall (n=246)	LEEP (n=92)	Conization (n=154)	p-value
Age, years	42.1±9.6	41.0±9.6	42.8±9.5	0.142
Parity	2 (1-3)	2 (1-3)	2 (1-3)	0.288
Current smoking, n (%)	103 (42.7%)	39 (43.3%)	64 (42.4%)	0.992
Preoperative cytology, (n=208)				0.025
Normal	76 (36.5%)	30 (40.0%)	46 (34.6%)	
ASC-US	38 (18.3%)	21 (28.0%)	17 (12.8%)	
LSIL	33 (15.9%)	11 (14.7%)	22 (16.5%)	
HSIL	52 (25.0%)	12 (16.0%)	40 (30.1%)	
ASC-H	8 (3.8%)	1 (1.3%)	7 (5.3%)	
AGC	1 (0.5%)	0 (0.0%)	1 (0.8%)	
Preoperative HPV status (n=166)				1.000
Negative	3 (1.8%)	1 (1.4%)	2 (2.1%)	
Positive	163 (98.2%)	69 (98.6%)	94 (97.9%)	
Colposcopy-directed biopsy (n=237)				<0.001
CIN1	37 (15.6%)	37 (43.0%)	0 (0.0%)	
CIN2	102 (43.0%)	37 (43.0%)	65 (43.0%)	
CIN3	92 (38.8%)	10 (11.6%)	82 (54.3%)	
CIS	6 (2.5%)	2 (2.3%)	4 (2.6%)	
ECC (n=245)				0.352
Not done	114 (46.5%)	44 (47.8%)	70 (45.8%)	
Benign/normal	116 (47.3%)	44 (47.8%)	72 (47.1%)	
LSIL	5 (2.0%)	3 (3.3%)	2 (1.3%)	
HSIL	9 (3.7%)	1 (1.1%)	8 (5.2%)	
CIS	0 (0.0%)	0 (0.0%)	0 (0.0%)	
SCC	1 (0.4%)	0 (0.0%)	1 (0.7%)	
Adenocarcinoma	0 (0.0%)	0 (0.0%)	0 (0.0%)	

-Values are mean ± SD, median (IQR), or n (%)
-P-values: Student's t-test or Mann-Whitney U test for continuous variables; chi-square or Fisher's exact test for categorical variables
-Preoperative HPV was available in n=166 patients
LEEP: Loop electrosurgical excision procedure, ASC-US: Atypical squamous cells of undetermined significance, LSIL: Low-grade squamous intraepithelial lesion, HSIL: High-grade squamous intraepithelial lesion, ASC-H: Atypical squamous cells-cannot exclude HSIL, AGC: Atypical glandular cells, CIS: Carcinoma *in situ*, ECC: Endocervical curettage, SD: Standard deviation, IQR: Interquartile range, HPV: Human papillomavirus

were upgraded and 40.3% were downgraded; overall agreement was low (Cohen's kappa=0.168).

In multivariable Firth penalized logistic regression evaluating factors associated with PSM (complete cases: n=215; 18 PSM events), greater cone depth was associated with lower odds of PSM (adjusted odds ratio 0.46 per 1-cm increase, 95% CI 0.24-0.88; p=0.018). HSIL or higher-grade histology and procedure type (conization vs. LEEP) were not significantly associated with PSM (Table 4).

DISCUSSION

In this single-center retrospective cohort of patients undergoing LEEP or cervical conization, the rates of PSM, recurrence, and reconization were 8.3%, 6.5%, and 23.5%, respectively. The most notable finding was that cone depth was associated with lower odds of PSM in multivariable Firth penalized logistic regression (aOR 0.46 per 1-cm increase). In contrast, HSIL or higher-grade histology and procedure type (conization vs. LEEP) were not significantly associated with PSM in the multivariable model. In addition, agreement between colposcopic impression and the final histology of the excision specimen was low; notably, among cases assessed as CIN2-3 on colposcopy, final histology could shift both toward lower-grade (downgrade) and higher-grade (upgrade) diagnoses. Taken together, these observations highlight two interrelated aspects of clinical practice: the importance of achieving an adequately deep excision during procedural planning and the inherent limitations of colposcopic assessment within the diagnostic pathway (13,18,19).

The clinical relevance of PSM stems from their association with residual disease and/or recurrent CIN2+; therefore, strategies aimed at reducing margin positivity are of practical importance. Prior studies have suggested that PSM may increase the

J Acad Res Med 2026;16(2):56-62

Table 2. Operative and postoperative outcomes by procedure (LEEP vs. conization)

Variable	Overall (n=246)	LEEP (n=92)	Conization (n=154)	p-value
Cone depth, cm	1.6 (1.0-2.5)	1.5 (0.8-2.0)	2.0 (1.2–2.5)	<0.001
Any complication, n (%)	3 (1.2%)	1 (1.1%)	2 (1.3%)	1.000
Complication type (n=246)				1.000
None	243 (98.8%)	91 (98.9%)	152 (98.7%)	
Bleeding	3 (1.2%)	1 (1.1%)	2 (1.3%)	
Infection	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Final excision pathology (n=243)				<0.001
Benign	54 (22.2%)	31 (34.1%)	23 (15.1%)	
LSIL	65 (26.7%)	34 (37.4%)	31 (20.4%)	
HSIL	101 (41.6%)	23 (25.3%)	78 (51.3%)	
CIS	19 (7.8%)	3 (3.3%)	16 (10.5%)	
SCC	3 (1.2%)	0 (0.0%)	3 (2.0%)	
Adenocarcinoma	1 (0.4%)	0 (0.0%)	1 (0.7%)	
Positive surgical margin, n (%)	20 (8.3%)	6 (6.6%)	14 (9.3%)	0.612
Re-conization performed, n (%)	57 (23.5%)	11 (12.0%)	46 (30.5%)	0.002
Recurrence, n (%)	16 (6.5%)	3 (3.3%)	13 (8.4%)	0.184
12-month HPV negative†	18/24 (75.0%)	7/10 (70.0%)	11/14 (78.6%)	0.665
24-month HPV negative†	11/14 (78.6%)	5/8 (62.5%)	6/6 (100.0%)	0.209

-Values are mean ± SD, median (IQR), or n (%)
-P-values: Student's t-test or Mann-Whitney U test for continuous variables; chi-square or Fisher's exact test for categorical variables
†: Values are presented as number of HPV-negative patients/number of patients with available HPV follow-up results (%) LEEP: Loop electrosurgical excision procedure, LSIL: Low-grade squamous intraepithelial lesion, HSIL: High-grade squamous intraepithelial lesion, CIS: Carcinoma *in situ*, SD: Standard deviation, IQR: Interquartile range, HPV: Human papillomavirus

Table 3. Colposcopic impression vs. final histology

Colposcopy impression	Benign n (%)	LSIL n (%)	HSIL n (%)	CIS+ n (%)	Total
CIN1	14 (37.8)	18 (48.6)	5 (13.5)	0 (0.0)	37
CIN2	23 (22.5)	28 (27.5)	45 (44.1)	6 (5.9)	102
CIN3	12 (13.2)	16 (17.6)	50 (54.9)	13 (14.3)	91
CIS	1 (16.7)	0 (0.0)	1 (16.7)	4 (66.7)	6

-Cells show, n (row %). Final histology grouped as Benign, LSIL, HSIL, and CIS+ (CIS or invasive carcinoma)
-Included cases with both colposcopy impression and final histology available: n=236
HSIL: High-grade squamous intraepithelial lesion, LSIL: Low-grade squamous intraepithelial lesion, CIS: Carcinoma *in situ*, CIN: Cervical intraepithelial neoplasia

Table 4. Multivariable Firth penalized logistic regression for positive surgical margins

Variable	Comparison/unit	aOR	95% CI	p-value
Cone depth (cm)	Per 1 cm increase	0.46	0.24-0.88	0.018
HSIL or higher	Yes vs. no	2.14	0.72-6.36	0.173
Procedure type	Conization vs. LEEP	1.79	0.57-5.63	0.318

- Firth penalized (bias-reduced) logistic regression
- Outcome: Positive margin=1 (negative margin=0)
- Analysis performed on complete cases: n=215; events (positive margins)=18
- Reference categories: HSIL or higher (no); procedure type (LEEP)
- HSIL or higher-grade histology was defined as HSIL, CIS, or invasive carcinoma on the excision specimen
OR: Odds ratio, CI: Confidence interval, HSIL: High-grade squamous intraepithelial lesion, CIS: Carcinoma *in situ*, LEEP: Loop electrosurgical excision procedure

likelihood of residual or recurrent lesions, and that endocervical margin involvement in particular may confer a higher risk (9,12). The relatively moderate-to-low PSM rate observed in our cohort may reflect standardized practice patterns within a single center and a selected case mix; nonetheless, the identification of cone depth as a readily modifiable parameter associated with PSM is clinically meaningful.

J Acad Res Med 2026;16(2):56-62

The inverse association between cone depth and PSM is consistent with plausible biological and technical considerations, as inadequate excision depth may lead to margin involvement in cases with endocervical extension of disease or an endocervically located transformation zone (13). However, increasing cone depth indiscriminately is not desirable, given evidence linking greater excision depth to an elevated risk of adverse obstetric outcomes, including preterm birth (5,6). Accordingly, our findings should not be interpreted as "deeper is always better," but rather as support for tailoring excision depth to lesion location while avoiding unnecessary removal of cervical tissue. This approach is particularly relevant for patient counseling and procedural planning in women of reproductive age.

From a post-treatment surveillance perspective, it is important to note that margin status alone should not be considered determinative. The 2019 risk-based management guidelines emphasize early post-treatment follow-up using HPV-based testing irrespective of margin status, with colposcopy and appropriate biopsies recommended when post-treatment HPV testing is positive (7). In our cohort, HPV results at 12 and 24 months were available only for a limited subset of patients, precluding more robust inferences regarding HPV clearance and HPV-based risk stratification. This limitation reflects a common challenge in retrospective real-world series, where follow-up testing is not uniformly standardized, and it underscores the need for prospective studies and more structured follow-up protocols.

The low level of agreement between colposcopic impression and final excision histology is also clinically relevant. Given the operator-dependent nature of colposcopy, lesion heterogeneity, and inherent sampling limitations of biopsy, perfect concordance between colposcopic assessment and histopathological outcomes is not expected (18,20). Prior studies have similarly reported low kappa values, indicating only limited concordance between colposcopic impression and histology (20). In our series, the presence of HSIL among cases with a colposcopic impression of CIN1 may be attributable to occult high-grade foci or endocervical lesion location, whereas benign or LSIL findings among cases interpreted as CIN3 may reflect overestimation at the impression or biopsy stage, or a more limited lesion burden on the excision specimen. Collectively, these observations support a risk-based approach that integrates multiple parameters, including cytology, HPV status, colposcopic findings, ECC results, and overall clinical context, rather than relying excessively on colposcopic impression alone, particularly when making management decisions for CIN2 (7).

Notably, the reconization rate exceeded the PSM rate in our cohort. This should not be interpreted as discordant, because margin status is an imperfect surrogate for residual disease and is not the sole determinant of repeat excision (7,12). In clinical practice, reconization was frequently prompted by post-treatment surveillance findings suggestive of persistence or progression (e.g., persistent HPV positivity, high-grade cytology, or suspicious colposcopy/ECC/biopsy), and some patients were referred from external centers following abnormal follow-up tests (7). Therefore, the observed reconization frequency likely reflects real-world, risk-based management rather than margin status alone.

Study Limitations

The strengths of this study include the use of a real-world, single-center cohort accrued over an extended time period, the inclusion of both LEEP and conization, and the ability to link a modifiable technical parameter such as cone depth with clinically relevant outcomes. Key limitations are inherent to the retrospective design and include missing data for certain variables (particularly follow-up HPV results), the absence of separate reporting for endocervical versus ectocervical margin involvement, and a record-based definition of recurrence. Because multivariable analyses were based on complete cases, selection bias cannot be excluded if missingness was related to patient characteristics or outcomes. Accordingly, inferences regarding HPV clearance and HPV-based risk stratification are limited, and the follow-up outcomes should be interpreted as descriptive rather than confirmatory. Importantly, the number of PSM events was limited, which constrained model complexity; therefore, we used Firth penalized logistic regression (bias-reduced) to mitigate small-sample bias and potential separation, but residual confounding cannot be excluded (21). Given the limited number of PSM events, statistical power was modest and CIs were wide for some predictors (notably HSIL or higher-grade histology and procedure type). Therefore, non-significant associations should be interpreted cautiously and should not be taken as evidence of no effect. In addition, because procedure selection (LEEP vs. conization) was associated with disease severity, selection bias and confounding by indication should be considered when interpreting differences between procedure types.

In summary, our findings suggest that cone depth represents a modifiable determinant of PSM, and that a targeted approach aimed at achieving an adequate excision depth according to lesion location, while taking obstetric risks into account, may constitute a rational strategy in clinical practice.

CONCLUSION

In this single-center retrospective cohort, the rate of PSM was 8.3%, and greater cone depth was associated with lower odds of margin positivity in multivariable Firth penalized logistic regression. The low level of agreement between colposcopic impression and final histology underscores the importance of a risk-based approach that integrates multiple preoperative data sources. Clinically, targeting an adequate excision depth according to lesion location may help reduce the risk of positive margins; however, individualized surgical planning that accounts for obstetric considerations remains essential, particularly in women of reproductive age.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the University of Health Sciences Türkiye, Gaziosmanpaşa Training and Research Hospital Non-Interventional Ethics Committee (decision no: 134, date: 15.10.2025).

J Acad Res Med 2026;16(2):56-62

Informed Consent: Retrospective study.

Footnotes

Author Contributions: Surgical and Medical Practices - E.S.C., F.K.G., H.B.B., A.A., S.S.; Concept - E.S.C., H.B.B., S.S.; Design - E.S.C., F.K.G.; Data Collection and/or Processing - E.S.C., H.B.B.; Analysis and/or Interpretation - E.S.C., A.A., S.S.; Literature Search - E.S.C., F.K.G., A.A.; Writing - E.S.C., A.A.

Conflict of Interest: The authors have no conflict of interest to declare.

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

Supplementary Table 1: https://d2v96fxpocvxx.cloudfront.net/cf9d60d6-523c-458a-a2e6-78728d3ffbb0/content-images/89b4805d-b0e7-4eb5-abfa-fc7a9d5ac2b9.pdf

REFERENCES

1. Gupta S, Nagtode N, Chandra V, Gomase K. From diagnosis to treatment: exploring the latest management trends in cervical intraepithelial neoplasia. Cureus. 2023; 15: e50291.
2. Eun TJ, Perkins RB. Screening for cervical cancer. Med Clin North Am. 2020; 104: 1063-78.
3. Pankakoski M, Sarkeala T, Anttila A, Heinävaara S. Effectiveness of cervical testing in and outside a screening program-a case-control study. Cancers (Basel). 2022; 14: 5193.
4. Perkins RB, Guido RS, Castle PE, Chelmow D, Einstein MH, Garcia F, et al.; 2019 ASCCP Risk-Based Management Consensus Guidelines Committee. 2019 ASCCP Risk-Based Management Consensus guidelines for abnormal cervical cancer screening tests and cancer precursors. J Low Genit Tract Dis. 2020; 24: 102-31. Erratum in: J Low Genit Tract Dis. 2020; 24: 427.
5. Athanasiou A, Veroniki AA, Efthimiou O, Kalliala I, Naci H, Bowden S, et al. Comparative effectiveness and risk of preterm birth of local treatments for cervical intraepithelial neoplasia and stage IA1 cervical cancer: a systematic review and network meta-analysis. Lancet Oncol. 2022; 23: 1097-108.
6. Wiik J, Kärrberg C, Nilsson S, Strander B, Jacobsson B, Sengpiel V. Associations between cervical intraepithelial neoplasia during pregnancy, previous excisional treatment, cone-length and preterm delivery: a register-based study from western Sweden. BMC Med. 2022; 20: 61.
7. Perkins RB, Guido RS, Castle PE, Chelmow D, Einstein MH, Garcia F, et al.; 2019 ASCCP Risk-Based Management Consensus Guidelines Committee. Erratum: 2019 ASCCP Risk-Based Management Consensus guidelines for abnormal cervical cancer screening tests and cancer precursors. J Low Genit Tract Dis. 2021; 25: 330-1.
8. Li Y, Gong YX, Wang Q, Gao S, Zhang H, Xie F, et al. Optimizing the detection of occult cervical cancer: a prospective multicentre study in China. Int J Womens Health. 2021; 13: 1005-15.
9. Ding T, Li L, Duan R, Chen Y, Yang B, Xi M. Risk factors analysis of recurrent disease after treatment with a loop electrosurgical excision procedure for high-grade cervical intraepithelial neoplasia. Int J Gynaecol Obstet. 2023; 160: 538-47.
10. Bogani G, DI Donato V, Sopracordevole F, Ciavattini A, Ghelardi A, Lopez S, et al. Recurrence rate after loop electrosurgical excision procedure (LEEP) and laser conization: a 5-year follow-up study. Gynecol Oncol. 2020; 159: 636-41.
11. Bruno MT, Valenti G, Cassaro N, Palermo I, Incognito GG, Cavallaro AG, et al. The coexistence of cervical intraepithelial neoplasia (CIN3) and adenocarcinoma in situ (AIS) in LEEP excisions performed for CIN3. Cancers (Basel). 2024; 16: 847.
12. Feng H, Chen H, Huang D, He S, Xue Z, Pan Z, et al. Relationship between positive margin and residual/recurrence after excision of cervical intraepithelial neoplasia: a systematic review and meta-analysis. Transl Cancer Res. 2022; 11: 1762-9.
13. Cong Q, Yu Y, Xie Y, Li Y, Sui L. Risk factors of LEEP margin positivity and optimal length of cervical conization in cervical intraepithelial neoplasia. Front Oncol. 2023; 13: 1209811.
14. Kim SI, Kim SJ, Suh DH, Kim K, No JH, Kim YB. Pathologic discrepancies between colposcopy-directed biopsy and loop electrosurgical excision procedure of the uterine cervix in women with cytologic high-grade squamous intraepithelial lesions. J Gynecol Oncol. 2020; 31: e13.
15. Booth BB, Petersen LK, Blaakaer J, Johansen T, Mertz H, Dahl K, et al. Accuracy of colposcopy-directed biopsy vs dynamic spectral imaging directed biopsy in correctly identifying the grade of cervical dysplasia in women undergoing conization: a methodological study. Acta Obstet Gynecol Scand. 2020; 99: 1064-70.
16. Jespersen MM, Booth BB, Petersen LK. Can biopsies be omitted after normal colposcopy in women referred with low-grade cervical cytology? A prospective cohort study. BMC Womens Health. 2021; 21: 394.
17. Chu MM, Cheung CS, Ngu SF, Tse KY, Ip PP, Cheung AN, et al. The value of four-quadrant cervical biopsy in women with different colposcopic impressions. Diagnostics (Basel). 2023; 13: 2384.
18. Stuebs FA, Dietl AK, Behrens A, Adler W, Geppert C, Hartmann A, et al. Concordance rate of colposcopy in detecting cervical intraepithelial lesions. Diagnostics (Basel). 2022; 12: 2436.
19. Gezer S, Kanbay Ozturk S, Balci S, Yucesoy I. The concordance between colposcopic biopsy and loop electrosurgical excision procedures in patients with known smear cytology and human papillomavirus results. North Clin Istanb. 2021; 8: 588-94.
20. Li J, Wang W, Yang P, Chen J, Dai Q, Hua P, et al. Analysis of the agreement between colposcopic impression and histopathological diagnosis of cervical biopsy in a single tertiary center of Chengdu. Arch Gynecol Obstet. 2021; 304: 1033-41.
21. Gosho M, Ohigashi T, Nagashima K, Ito Y, Maruo K. Bias in odds ratios from logistic regression methods with sparse data sets. J Epidemiol. 2023; 33: 265-75. Erratum in: J Epidemiol. 2023; 33: 332.

DOI: 10.4274/jarem.galenos.2026.71542
J Acad Res Med 2026;16(2):63-69

The Role of Systemic Inflammatory Indices in Predicting Postpartum Hemorrhage Following Cesarean Delivery

Yücel Kaya[1], Emrah Dağdeviren[2], İlteriş Yaman[3], Emrullah Akay[2], Sultan Can[4]

[1]University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital, Department of Perinatology, İstanbul, Türkiye
[2]University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital, Department of Obstetrics and Gynecology, İstanbul, Türkiye
[3]University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital, Department of Gynecologic Oncology, İstanbul, Türkiye
[4]Private Clinic, Clinic of Obstetrics and Gynecology, İstanbul, Türkiye

Cite this article as: Kaya Y, Dağdeviren E, Yaman İ, Akay E, Can S. The role of systemic inflammatory indices in predicting postpartum hemorrhage following cesarean delivery. J Acad Res Med. 2026;16(2):63-69

ABSTRACT

Objective: To determine whether preoperative inflammation-related indices calculated from routine hematological parameters are associated with objectively measured postpartum hemorrhage (PPH) following cesarean section.

Methods: A retrospective cohort design was applied at a tertiary referral center and included 226 women who delivered by cesarean section between July 2025 and January 2026. Postpartum blood loss was calculated using a standardized formula-based approach. According to the quantified blood loss, patients were classified as <1000 mL or ≥1000 mL. We conducted a comparative analysis of preoperative inflammatory markers—specifically neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, monocyte-to-lymphocyte ratio, systemic immune-inflammation index (SII), systemic inflammatory response index, and pan-immune-inflammation value—across the study cohorts. Furthermore, the prognostic utility of these indices was assessed using receiver operating characteristic (ROC) curves, and multivariable logistic regression modeling was subsequently used to identify independent associations.

Results: Baseline maternal and obstetric characteristics did not differ significantly between the study groups. Among the investigated indices, only SII values were significantly higher in women with postpartum blood loss ≥1000 mL. SII demonstrated a weak but statistically significant positive correlation with total blood loss. ROC analysis yielded an area under the curve (AUC) of 0.625 for SII. After adjustment for relevant covariates, an SII threshold of ≥1175 remained independently associated with an increased risk of clinically significant PPH.

Conclusion: Although the discriminative ability of SII as a standalone test was modest (AUC=0.625), elevated preoperative SII levels were independently associated with increased postpartum blood loss after cesarean delivery. Therefore, SII may serve as an adjunctive biomarker that complements established clinical risk factors in preoperative risk stratification.

Keywords: Postpartum hemorrhage, term birth, inflammation, biomarkers, cesarean section, risk assessment

INTRODUCTION

Postpartum hemorrhage (PPH) continues to represent a major contributor to maternal morbidity and mortality worldwide, being responsible for nearly 27% of maternal deaths (1). The World Health Organization characterizes PPH as blood loss of ≥500 mL following delivery, while losses reaching or exceeding 1000 mL are classified as severe PPH (1). Against the backdrop of increasing cesarean delivery rates worldwide, perioperative bleeding has emerged as a major management challenge in contemporary obstetric practice. Despite well-established etiological factors such as uterine atony, genital tract trauma, and coagulopathy, a considerable number of PPH cases have been observed in women without identifiable antenatal risk factors (2,3).

Labor is an inflammatory process characterized by myometrial infiltration of leukocytes, macrophages, and neutrophils that secrete proinflammatory cytokines (4). The bidirectional interplay between systemic inflammation and the coagulation system directly shapes hemostatic balance through the modulation of platelets, the endothelium, and the coagulation cascade by inflammatory mediators (4,5). Within this biological framework, it is postulated that the magnitude of the systemic inflammatory response may correspond to bleeding tendency during delivery;

ORCID IDs of the authors: Y.K. 0000-0003-4597-2922; E.D. 0000-0002-1730-3724; İ.Y. 0000-0002-5563-2887; E.A. 0000-0003-3792-7777; S.C. 0000-0003-4927-0557

Corresponding Author: Yücel Kaya, MD;
E-mail: yucelkaya0007@gmail.com
Received Date: 10.02.2026 **Accepted Date:** 07.05.2026
Epub: 11.05.2026
Publication Date: 07.08.2026

J Acad Res Med 2026;16(2):63-69

therefore, inflammatory markers could play a potential role in the assessment of postpartum bleeding risk (6,7).

In recent years, studies in the literature have demonstrated that inflammatory biomarkers based on routine hematological indices, including the neutrophil-to-lymphocyte ratio (NLR), possess prognostic relevance across a broad spectrum of obstetric pathologies (8,9). Furthermore, more contemporary composite indices—including the systemic immune-inflammation index (SII), the systemic inflammatory response index (SIRI), and the pan-immune-inflammation value (PIV)—reflect the inflammatory profile in a more integrated manner by allowing for the simultaneous evaluation of different immune cell series (10-12). Although these indices have been extensively investigated for their prognostic relevance in major obstetric conditions such as preeclampsia, data regarding their predictive performance for PPH risk among patients undergoing cesarean delivery remain limited. As a result, further studies are essential to validate the practical application of systemic inflammatory biomarkers, specifically SII, SIRI, and PIV.

The primary objective of this study is to evaluate the association between preoperatively measured complete blood count-based inflammatory indices and the objectively quantified amount of postpartum blood loss in women undergoing cesarean delivery.

METHODS

Study Design

The present investigation adopted a retrospective observational cohort design. No additional diagnostic or therapeutic procedures were implemented; the study was conducted entirely using data recorded during routine clinical practice. All study procedures were conducted in accordance with the ethical standards established by the Declaration of Helsinki. Ethical approval was granted by the Clinical Research Ethics Committee of the University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital (decision no: KAEK/28.01.2026.19, date: 03.02.2026).

Setting

This research was conducted at a single tertiary-care center in the University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital, Department of Obstetrics and Gynecology. The research cohort comprised patients who underwent cesarean deliveries at our center between July 2025 and January 2026. Obstetric management and follow-up for all cases were performed in accordance with our institution's established clinical standards.

Participants

The study population was selected from patients who delivered by cesarean section at our clinic during the study period. Inclusion criteria were defined as maternal age between 18-40 years, delivery occurring at ≥37 gestational weeks, and the presence of a singleton pregnancy. Cases of multiple pregnancies, obstetric complications (preeclampsia, gestational hypertension), chronic maternal diseases (e.g., pregestational diabetes), chorioamnionitis with premature rupture of membranes, requirement for blood transfusion or management of massive hemorrhage, placental pathologies (previa, abruption), and missing clinical data were excluded from the analysis.

Variables

The primary outcome variable was the amount of postpartum bleeding. Independent variables included preoperative hematological inflammatory markers [NLR, platelet-to-lymphocyte ratio (PLR), monocyte-to-lymphocyte ratio (MLR), SII, SIRI, and PIV], total leukocyte count, coagulation parameters [prothrombin time (PT), activated partial thromboplastin time (aPTT), and international normalized ratio (INR)], and demographic and obstetric factors [maternal age, body mass index (BMI), gravidity, parity, gestational week at delivery, and clinical characteristics related to birth].

Data Sources/Measurement

Relevant clinical data were retrieved retrospectively via the institution's digital registry and corresponding patient files. Preoperative laboratory parameters were extracted from complete blood counts and coagulation tests performed as part of the routine antepartum evaluation. The amount of postpartum blood loss was quantitatively calculated in milliliters using the Stafford formula, incorporating maternal height and weight as well as preoperative and 6-hour postoperative hematocrit values (13). The NLR, PLR, and MLR were calculated by determining the ratios of the respective cell parameters (neutrophils, lymphocytes, monocytes, and platelets). In addition, composite indices reflecting systemic inflammatory and immune status—SII, SIRI, and PIV—were calculated using the following formulas, respectively: neutrophils×platelets/lymphocytes, neutrophils×monocytes/lymphocytes, and neutrophils×platelets×monocytes/lymphocytes. Patients were stratified for analysis based on calculated postpartum blood loss into Group 1 (<1000 mL) and Group 2 (≥1000 mL).

Bias

To minimize potential selection bias in the study design, all consecutive patients who presented to the clinic within the predefined time frame and met the eligibility criteria were included. To reduce information bias, the amount of blood loss was quantified using an objective formula based on laboratory and anthropometric parameters rather than visual estimation. Moreover, data reliability was enhanced by excluding patients who received blood transfusions or developed massive hemorrhage requiring surgical intervention, because these conditions could alter hematocrit concentrations and introduce error into the mathematical calculation. All laboratory measurements and clinical data were acquired at the same center using identical analytical methods and standard protocols for both groups. Finally, to limit confounding bias, conditions that could influence baseline levels of inflammatory biomarkers—such as infection, preeclampsia, and chronic systemic diseases—were excluded.

J Acad Res Med 2026;16(2):63-69

Study Size

To determine the necessary cohort size before the study onset, we performed a statistical power estimation using the G*Power package (v3.1, developed at Heinrich Heine University, Düsseldorf) (14). Based on data reported by Akay et al. (7), the effect size (Cohen's d) was set at 0.386, with a type I error rate (α) of 0.05 and a statistical power (1-β) of 0.80, yielding a required minimum sample size of 214 participants. During the retrospective screening process, 226 patients who met the study criteria (Group 1: 185; Group 2: 41) were included in the analysis, resulting in a sample size exceeding the predetermined requirement.

Statistical Analysis

Statistical computations were performed using IBM SPSS Statistics (v26.0; IBM Corp., Armonk, NY, USA). To verify whether continuous variables conformed to a Gaussian distribution, the Kolmogorov-Smirnov test was applied. For descriptive reporting, data exhibiting a Gaussian distribution were cited as mean ± standard deviation, while non-parametric data were summarized as medians (minimum-maximum). Between-group comparisons were conducted using Student's t-test or the Mann-Whitney U test, as appropriate, based on the data distribution. Categorical data were summarized as frequencies (n) and percentages (%), with intergroup differences evaluated via the chi-square test. The relationship between inflammatory indices and PPH was examined using Spearman's rank correlation analysis or Fisher's exact test, as appropriate. The discriminative performance of SII for predicting PPH (≥1000 mL) was assessed by receiver operating characteristic (ROC) curve analysis, and the area under the curve (AUC) was reported with a 95% confidence interval (CI). The optimal cut-off value was determined using the Youden index. A multivariable binary logistic regression analysis was executed to determine independent risk factors for PPH (≥1000 mL). Parity, gestational age at delivery, and SII (≥1175) were included in the model based on their clinical relevance and potential association with PPH. The adequacy of the model was evaluated via the Hosmer-Lemeshow goodness-of-fit test. For all statistical comparisons, a two-sided $p<0.05$ was considered statistically significant.

RESULTS

The final analytical cohort encompassed 226 eligible women who underwent cesarean delivery at our institution. Stratification of the study population revealed that 185 participants (81.8%) belonged to Group 1, whereas 41 (18.1%) were assigned to Group 2. As no missing data were identified, all statistical analyses were performed on this final study cohort.

The baseline demographic and obstetric characteristics of the study groups are presented comparatively in Table 1. No statistically significant differences were detected between Group 1 and Group 2 regarding maternal age, BMI, gravidity, parity, and the gestational week at delivery (all $p>0.05$). Similarly, the groups demonstrated comparable distributions concerning birth weight, rates of primary cesarean delivery, presence of labor induction, and the proportion of cesarean deliveries performed during the active phase of labor (all $p>0.05$).

In the evaluation of preoperative laboratory parameters and inflammatory indices, no statistically significant differences were observed between the groups with respect to complete blood count components (hemoglobin, hematocrit, leukocyte, platelet, neutrophil, lymphocyte, and monocyte counts) or coagulation tests (PT, aPTT, and INR) (all $p>0.05$). Likewise, no statistically significant variations were detected between the cohorts regarding the inflammatory markers NLR, PLR, MLR, SIRI, and PIV (all $p>0.05$). In contrast, SII levels were significantly higher in Group 2 compared with Group 1 (median 1229.40 vs. 969.33, $p=0.012$) (Table 2).

The relationship between inflammatory indices and the amount of postpartum blood loss was examined using Spearman correlation analysis. A weak but statistically significant positive correlation was identified between the amount of postpartum blood loss and SII levels among the inflammatory indices ($r=0.211$, $p<0.001$). ROC curve analysis was subsequently conducted to evaluate the discriminative capacity of SII for clinically significant PPH (≥1000 mL) (Figure 1). The analysis yielded an AUC value of 0.625 for SII (95% CI: 0.519-0.731; $p=0.012$). The optimal cut-off value determined by the Youden index was 1175, yielding a sensitivity of 58.5% and a specificity of 77.8% (Table 3).

Table 1. Comparison of demographic and obstetric characteristics of the study groups

	Group 1 (n=185)	Group 2 (n=41)	p
Maternal age[b] (years)	28 (18-40)	30 (19-40)	0.869
BMI[b] (kg/m^2)	30.4 (20.2-49.3)	30.4 (23.7-44.3)	0.890
Gravidity[b]	2 (1-8)	3 (1-8)	0.502
Parity[b]	1 (0-6)	1 (0-6)	0.108
GA at delivery (weeks)[b]	38.7 (37.0-41.0)	38.9 (37.0-41.0)	0.204
Primary cesarean section[c]	102 (55.1%)	25 (61.0%)	0.495
Birth weight[a] (g)	3106.86±511.75	3133.90±584.60	0.766
Labor induction[c]	61 (33.0%)	15 (36.6%)	0.658
Active phase of labor[c]	17 (9.2%)	3 (7.3%)	1.000

[a]: Normal distribution, mean ± standard deviation, [b]: Non-normal distribution, median (minimum-maximum), [c]: Categorical data, number (percentage %), BMI: Body mass index, GA: Gestational age

J Acad Res Med 2026;16(2):63-69

Table 2. Comparison of preoperative laboratory parameters and systemic inflammatory indices between the study groups

	Group 1 (n=185)	Group 2 (n=41)	p
Hb[b] (g/dL)	11.6 (8.7-15.0)	12.0 (8.9-13.7)	0.348
HCT[b] (%)	35.1 (27.8-44.3)	36.6 (30.7-39.3)	0.164
WBC[b] (10^9/L)	10.31 (5.23-18.56)	10.26 (6.03-18.50)	0.741
PLT[b] (10^9/L)	250 (117-597)	241 (82-597)	0.827
Neutrophil[b] (10^9/L)	7.70 (2.23-18.68)	7.95 (3.06-19.78)	0.330
Lymphocyte[b] (10^9/L)	2.13 (0.65-7.90)	2.14 (0.90-20.20)	0.858
Monocyte[b] (10^9/L)	0.74 (0.26-1.66)	0.70 (0.40-6.00)	0.168
PT[b] (s)	8.4 (7.5-10.5)	8.4 (7.6-9.6)	0.818
aPTT[a] (s)	27.41±2.95	27.20±3.87	0.692
INR[b]	0.9 (0.8-1.1)	0.9 (0.8-1.1)	0.797
NLR[b]	3.60 (1.33-13.09)	3.44 (0.47-17.43)	0.882
PLR[b]	119.39 (28.99-567.14)	108.89 (7.03-412.31)	0.593
MLR[b]	0.34 (0.10-1.76)	0.29 (0.13-1.09)	0.374
SII[b]	969.33 (316.27-3050.51)	1229.40 (66.36-1790.77)	**0.012**
SIRI[b]	2.56 (0.58-12.09)	2.48 (0.81-13.83)	0.674
PIV[b]	691.90 (158.14-3542.48)	791.42 (137.08-1934.03)	0.663

Bold values indicate statistical significance (p<0.05), [a]: Normal distribution, mean ± standard deviation, [b]: Non-normal distribution, median (minimum-maximum), Hb: Hemoglobin, HCT: Hematocrit, WBC: White blood cell count, PLT: Platelet count, PT: Prothrombin time, aPTT: Activated partial thromboplastin time, INR: International normalized ratio, NLR: Neutrophil-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, MLR: Monocyte-to-lymphocyte ratio, SII: Systemic immune-inflammation index, SIRI: Systemic inflammation response index, PIV: Pan-immune-inflammation value

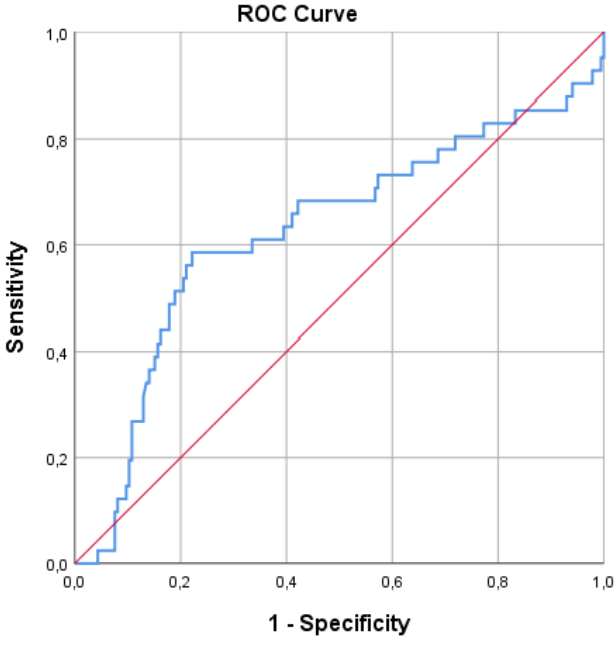

Figure 1. Receiver operating characteristic (ROC) curve illustrating the discriminative performance of the systemic immune-inflammation index for predicting postpartum hemorrhage ≥1000 mL after cesarean section. The area under the curve was 0.625 (95% confidence interval: 0.519-0.731; p=0.012)

We executed a multivariable binary logistic regression analysis to determine the independent predictors of PPH (≥1000 mL) (Table 4). The analysis demonstrated that neither parity nor gestational age at delivery was significantly associated with PPH (≥1000 mL) (p=0.153 and p=0.186, respectively). However, an SII level above the defined cut-off (≥1175) was independently associated with a statistically significant increase in the risk of PPH (≥1000 mL) (adjusted odds ratio: 4.887; 95% CI: 2.378-10.043; p<0.001). The model showed good fit, as confirmed by the Hosmer-Lemeshow goodness-of-fit test (p=0.087).

DISCUSSION

Principal Findings

This study examined the relationship between preoperative complete blood count-based inflammatory indices and objectively quantified postpartum blood loss among women undergoing cesarean delivery. The principal observation was that while routine hemogram parameters and other derived inflammatory indices did not differ between the groups, preoperative SII levels were significantly elevated in patients who developed PPH. Furthermore, our analysis unveiled a significant positive linear relationship linking preoperative SII concentrations directly to the volume of PPH (≥1000 mL). In ROC curve analysis, SII yielded an AUC of 0.625, while multivariable logistic regression revealed an independent association between elevated SII levels and PPH (≥1000 mL).

J Acad Res Med 2026;16(2):63-69

Table 3. ROC analysis of the SII for predicting postpartum hemorrhage ≥1000 mL

	Cut-off	AUC	95% CI	Sensitivity (%)	Specificity (%)	p
SII	1175	0.625	0.519-0.731	58.5	77.8	**0.012**

ROC: Receiver operating characteristic, SII: Systemic immune-inflammation index, AUC: Area under the curve, CI: Confidence interval

Table 4. Multivariate logistic regression analysis of independent predictors for postpartum hemorrhage ≥1000 mL

Variable	B	Adjusted OR [Exp(B)]	95% CI for OR	p
Parity	0.145	1.156	0.947-1.411	0.153
GA at delivery (weeks)	0.209	1.232	0.904-1.679	0.186
High SII (≥1175)	1.587	4.887	2.378-10.043	**<0.001**

Note: The multivariate model was adjusted for parity and gestational age. The cut-off value for SII was determined using the Youden index derived from the receiver operating characteristic curve analysis. The model showed good fit (Hosmer-Lemeshow goodness-of-fit test, p=0.087). Bold value indicate statistical significance (p<0.05)
B: Regression coefficient, OR: Odds ratio, CI: Confidence interval, GA: Gestational age, SII: Systemic immune-inflammation index

Interpretation

The primary observation of this study is that elevated preoperative SII values are associated with the development of clinically significant hemorrhage following cesarean section. A review of the literature examining the role of inflammatory indices in predicting PPH risk indicates that no clear consensus exists regarding the prognostic value of SII. In a cohort involving placental pathologies, SIRI and PLR were reported to have the strongest predictive value for intraoperative blood loss, whereas NLR and SII showed significant increases in cases of massive hemorrhage (15). Likewise, a recent study reported that elevated NLR, PLR, and SII levels were important prognostic markers for predicting composite adverse outcomes in severe PPH requiring fibrinogen replacement therapy (6). In contrast, Akay et al. (7) observed significantly lower SII levels in patients who developed PPH compared with the control group. This heterogeneity in the literature may stem from several factors, including variations in PPH definitions, sample selection (particularly inclusion of cases requiring massive transfusion or surgical intervention), and timing of measurements. In addition, concomitant clinical conditions that modify the inflammatory response may contribute.

The observation that SII emerged as the only significant biomarker, despite the lack of association between other inflammatory indices and PPH, may be explained by differences in the underlying cellular mechanisms upon which these indices are based. Ratios derived from three cell lineages—namely NLR, PLR, and MLR—may have failed to yield significant results due to their limited biological representativeness during the peripartum period, characterized by substantial physiological variability. SIRI, which does not incorporate platelets, and PIV, which combines multiple cellular parameters into a single formula, may inadequately capture the obstetric bleeding phenotype, thereby diluting the biological signal. Conversely, SII excludes cell lineages such as monocytes, which are typically associated with delayed responses or chronic processes, and directly reflects both inflammatory activity and platelet-mediated hemostatic capacity (10). Consequently, the distinct performance of SII compared with other indices appears to be related to its specific structure, which effectively integrates acute inflammation and hemostatic processes, suggesting that SII is a sensitive marker for assessing PPH risk.

The establishment of hemostasis in the postpartum period depends not only on mechanical uterine contractions but also on local thrombotic mechanisms involving tissue factor activation and systemic coagulation components (16,17). It has been suggested that increased inflammatory cell infiltration and cytokine release during the peripartum period may adversely affect myometrial contractility, thereby predisposing to the development of uterine atony (18,19). In other words, despite the absence of an identifiable microbial pathogen during this period, the myometrial tissue is dominated by an acute myometritis pattern characterized by massive cellular infiltration rich in neutrophils, macrophages, and mast cells (20). Additionally, histopathological examinations of cases with refractory PPH of unknown etiology following cesarean delivery have revealed that inflammatory and anaphylactoid reactions involve not only the uterine corpus but also the isthmus (21). Mandel et al. (22) further highlighted the regulatory role of platelets within the inflammation-coagulation axis, noting that platelet function may be compromised independently of platelet count under conditions of increased inflammatory activity. Taken together, data from the literature explain the pathophysiological basis by which SII, reflecting both inflammation and hemostasis, emerges as a more sensitive biomarker for predicting PPH risk than other inflammatory indices.

Clinically, a key finding of this study is that elevated SII levels are independently and significantly associated with postpartum blood loss. However, the modest but statistically significant discriminative power indicated by ROC analysis supports positioning SII not as a standalone diagnostic test but rather as an adjunctive parameter for preoperative risk stratification. As emphasized in numerous studies in the literature, the low cost and widespread availability of these hematological indices represent a major advantage, given their potential to provide complementary

value in clinical assessment (23,24). For clinical interpretation, SII values ≥1175 may help identify patients who could benefit from increased perioperative vigilance and preventive strategies for PPH. In light of the current findings, incorporating SII alongside established clinical risk factors may allow PPH risk to be addressed through a more holistic strategy and may enhance the predictive strength of preoperative evaluation.

Study Limitations

A notable strength of the present study is the quantitative assessment of postpartum blood loss, as opposed to reliance on subjective estimation. This approach reduces measurement bias, thereby enhancing the internal validity of the study and the reliability of the obtained results. Another notable strength is the exclusion of clinical conditions that could substantially influence the inflammatory profile, thereby limiting the impact of potential confounding factors. This strategy ensured a homogeneous cohort, allowing for a more robust and comprehensive assessment of the association between inflammatory indices and PPH.

Nevertheless, certain limitations of the present investigation warrant consideration. First, the retrospective study design restricts the ability to draw definitive causal inferences. Second, the single-institution design may limit the extent to which these results can be extrapolated to broader, more diverse patient populations. Another limitation is the reliance on measurements obtained at a single preoperative time point, which precludes the evaluation of dynamic fluctuations in inflammatory indices throughout pregnancy. The modest number of patients with clinically significant PPH may have constrained the statistical robustness of subgroup analyses. We did not measure all biochemical or cytokine markers that modulate the inflammatory response. Therefore, we cannot definitively exclude the presence of residual confounding factors. Therefore, future large-scale, multicenter, prospective studies are warranted to more comprehensively evaluate the clinical utility of inflammatory indices in predicting the risk of PPH.

CONCLUSION

This study demonstrates that elevated SII levels constitute an independent and significant risk factor for the development of PPH (≥1000 mL) in women undergoing cesarean delivery. However, given its modest discriminative ability (AUC=0.625), SII should be interpreted as an adjunctive biomarker rather than a standalone predictor. That SII emerged as the sole marker retaining statistical significance, in contrast to other inflammatory indices, underscores its potential clinical relevance. In women with elevated SII levels, optimizing perioperative preparation and implementing proactive preventive measures may substantially reduce maternal morbidity.

Ethics

Ethics Committee Approval: Ethical approval was granted by the Clinical Research Ethics Committee of the University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital (decision no: KAEK/28.01.2026.19, date: 03.02.2026).

Informed Consent: Due to the retrospective nature of the study, the requirement for informed consent was waived. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Footnotes

Author Contributions: Surgical and Medical Practices - Y.K., E.D.; Concept - Y.K., E.D.; Design - Y.K., E.D.; Data Collection and/or Processing - Y.K., İ.Y., E.A., S.C.; Analysis and/or Interpretation - Y.K., E.A., S.C.; Literature Search - Y.K., E.D., İ.Y., E.A., S.C.; Writing - Y.K., İ.Y.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors report that no financial support was received for this study.

REFERENCES

1. WHO Guidelines Approved by the Guidelines Review Committee. WHO recommendations on the assessment of postpartum blood loss and use of a treatment bundle for postpartum haemorrhage. Geneva: World Health Organization; 2023.
2. McLintock C. Prevention and treatment of postpartum hemorrhage: focus on hematological aspects of management. Hematology Am Soc Hematol Educ Program. 2020; 2020: 542-6.
3. Dagdeviren E, Tercan C, Yeniocak AS, Sarı N, Keskin ŞB, Yaman İ. Platelet indices as predictors of hemoglobin decline following cesarean section: a retrospective cohort analysis. Z Geburtshilfe Neonatol. 2026; 230: 122-8.
4. Jiang H, Shi H, Chen L, Yang J, Yuan P, Wang W, et al. Is there a relationship between plasma, cytokine concentrations, and the subsequent risk of postpartum hemorrhage? Am J Obstet Gynecol. 2022; 226: 835.e1-17.
5. Tercan C, Yeniocak AS, Dağdeviren E, Can S, İnce A, Azman MK, et al. Evaluating platelet indices, inflammatory markers, and the systemic immune-inflammation index as predictors of erythrocyte suspension transfusion in total abdominal hysterectomy: a matched cohort study. Kocaeli Medical Journal. 2025; 14: 91-8.
6. Tokalioglu EO, Tanacan A, Agaoglu MO, Özkavak O, Atalay A, Uzuner P, et al. The role of systemic inflammatory indices in predicting composite adverse outcomes of peripartum hemorrhage treated with fibrinogen concentrate. Indian J Hematol Blood Transfus. 2025; 41: 934-41.
7. Akay E, Şirin Donbalıoğlu G, Yeniocak AS, Dağdeviren E, Tercan C, Polat İ. Prognostic evaluation of systemic immune inflammatory index and hematological parameters in postpartum hemorrhage: a retrospective analysis. Ginekol Pol. 2025; 96: 287-93.
8. Gabbai D, Gilboa I, Lavie A, Yogev Y, Attali E. Neutrophil-to-lymphocyte ratio at admission helps to predict the need for blood transfusion after vaginal delivery. Int J Gynaecol Obstet. 2026; 173: 1489-95.
9. Özkan Y, Genç S, Kaya Y, Kaya DY, Mihmanlı V. Evaluation of neutrophil/lymphocyte and thrombocyte/lymphocyte ratios in cases of preterm premature rupture of membranes. Bagcilar Med Bull. 2024; 9: 9-14.
10. Hu B, Yang XR, Xu Y, Sun YF, Sun C, Guo W, et al. Systemic immune-inflammation index predicts prognosis of patients after curative resection for hepatocellular carcinoma. Clin Cancer Res. 2014; 20: 6212-22.
11. Qi Q, Zhuang L, Shen Y, Geng Y, Yu S, Chen H, et al. A novel systemic inflammation response index (SIRI) for predicting the survival of patients with pancreatic cancer after chemotherapy. Cancer. 2016; 122: 2158-67.
12. Lin F, Zhang LP, Xie SY, Huang HY, Chen XY, Jiang TC, et al. Pan-immune-inflammation value: a new prognostic index in operative breast cancer. Front Oncol. 2022; 12: 830138.
13. Stafford I, Dildy GA, Clark SL, Belfort MA. Visually estimated and calculated blood loss in vaginal and cesarean delivery. Am J Obstet Gynecol. 2008; 199: 519.e1-7.
14. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods. 2009; 41: 1149-60.
15. Balkaş G, Çelen Ş. Role of inflammatory and coagulation biomarkers in distinguishing placenta accreta from placenta previa and associated hemorrhage. J Clin Med. 2025; 14: 3884.
16. Warren BB, Moyer GC, Manco-Johnson MJ. Hemostasis in the pregnant woman, the placenta, the fetus, and the newborn infant. Semin Thromb Hemost. 2023; 49: 319-29.

J Acad Res Med 2026;16(2):63-69

17. Edwards HM. Aetiology and treatment of severe postpartum haemorrhage. Dan Med J. 2018; 65: B5444.
18. Shynlova O, Lee YH, Srikhajon K, Lye SJ. Physiologic uterine inflammation and labor onset: integration of endocrine and mechanical signals. Reprod Sci. 2013; 20: 154-67.
19. Zhang P, Jia Y, Song H, Fan Y, Lv Y, Geng H, et al. Novel biomarkers for prediction of atonic postpartum hemorrhage among 'low-risk' women in labor. Front Immunol. 2024; 15: 1416990.
20. Oda T, Tamura N, Ide R, Kawai K, Narumi M, Matsumoto M, et al. Postpartum acute myometritis suppresses expression of contraction-associated proteins in the gravid uterus. J Reprod Immunol. 2024; 165: 104299.
21. Jain D, Oda T, Kohmura-Kobayashi Y, Furuta-Isomura N, Yaguchi C, Uchida T, et al. Acute inflammation in the uterine isthmus coincides with postpartum acute myometritis in the uterine body involving refractory postpartum hemorrhage of unknown etiology after cesarean delivery. J Reprod Immunol. 2020; 139: 103116.
22. Mandel J, Casari M, Stepanyan M, Martyanov A, Deppermann C. Beyond hemostasis: platelet innate immune interactions and thromboinflammation. Int J Mol Sci. 2022; 23: 3868.
23. Cicek S, Coskun EI, Abdurahmanova N, Yilmaz E. Inflammation-driven pathogenesis of hyperemesis gravidarum: assessing the role of pan-immune-inflammation value and hematologic parameters. Arch Gynecol Obstet. 2025; 312: 2001-10.
24. Yang S, Shi H, Ren S, Chen Y, Jin X, Han L, et al. Association of inflammatory markers with pregnancy loss: Global Burden (1990-2021) and NHANES survey (2005-2016). BMC Pregnancy Childbirth. 2025; 25: 996.

DOI: 10.4274/jarem.galenos.2026.20053
J Acad Res Med 2026;16(2):70-78

Impact of the COVID-19 Pandemic on Psychiatric Emergencies in a Megacity

Mehmet Necmeddin Sutaşır[1], Mehmet Tatlı[2], Elif Marangoz[3], Kenan Ahmet Türkdoğan[4]

[1]University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital, Department of Emergency Medicine, İstanbul, Türkiye
[2]University of Health Sciences Türkiye, Van Training and Research Hospital, Department of Emergency Medicine, Van, Türkiye
[3]İstanbul University-Cerrahpaşa, Institute of Graduate Education, İstanbul, Türkiye
[4]İstanbul Provincial Health Directorate, Department of Emergency Health Services, İstanbul, Türkiye

Cite this article as: Sutaşır MN, Tatlı M, Marangoz E, Türkdoğan KA. Impact of the COVID-19 pandemic on psychiatric emergencies in a megacity. J Acad Res Med. 2026;16(2):70-78

ABSTRACT

Objective: To evaluate changes in the volume, demographic characteristics, diagnostic distribution, and clinical severity of prehospital psychiatric emergency presentations in İstanbul during the coronavirus disease 2019 (COVID-19) pandemic compared with a pre-pandemic reference period.

Methods: This retrospective observational time-period comparison study analyzed all psychiatric emergency encounters recorded in the emergency health events system between September 2018 and September 2021. Two equivalent 18-month intervals were compared: pre-pandemic (September 2018-March 2020) and pandemic (March 11, 2020-September 11, 2021). Descriptive statistics, chi-square tests, and relative percentage change analyses were performed.

Results: Among 91,803 eligible encounters, pandemic-period presentations increased by 18.7% compared with the pre-pandemic period. Male presentations rose by 41.6%, whereas pediatric encounters declined by 30.2%. Alcohol-related diagnoses demonstrated the most pronounced increase. Higher-acuity triage allocations became more frequent, and forensic case designations increased during the pandemic period.

Conclusion: The COVID-19 pandemic was associated with an increased number of psychiatric emergency encounters, shifts in demographic patterns, and greater clinical severity among presentations. These findings highlight the need for strengthened emergency psychiatric care capacity and improved access to early mental health interventions during large-scale public health crises.

Keywords: Psychiatric emergency, COVID-19, prehospital services

INTRODUCTION

The World Health Organization proclaimed the coronavirus disease 2019 (COVID-19) pandemic on March 11, 2020. It has become one of the most important global health crises of the twenty-first century, having a profound impact on healthcare systems all over the world (1). In addition to the direct medical effects of viral infection, the pandemic has caused unprecedented psychological distress among populations worldwide, with systematic reviews and meta-analyses indicating significant rises in the prevalence of depression, anxiety, and stress-related disorders (2,3). The enforcement of confinement measures, such as statewide lockdowns, social distance requirements, and economic limitations, has generated a distinctive array of psychosocial stresses that have significantly transformed the patterns of mental health treatment consumption (4).

Psychiatric emergency departments are essential entry sites for patients undergoing severe mental health crises, and alterations in their utilization patterns yield significant epidemiological insights into population-level mental health dynamics (5). International evidence has shown that the pandemic has had different consequences on the use of psychiatric emergency services. For example, there were fewer people using these services under stringent lockdowns, but more people started using them again after restrictions were lifted (6,7). A cross-sectional analysis examining almost 190 million emergency medical service encounters in the United States revealed that mental health-related visit rates rose during the pandemic period compared to 2019, despite an overall decline in total emergency medical service utilization (8). The pandemic has had a substantial impact on mental health, with a marked increase in symptoms of anxiety and depression. Meta-analytic estimates indicate that the prevalence of anxiety and depression was 31.9% and 31.4%, respectively, during the early phase of the pandemic (9). The COVID-19 mental disorders collaborators projected the emergence of 53.2 million new cases of major depressive disorder and 76.2 million new cases of anxiety

ORCID IDs of the authors: M.N.S. 0000-0001-6472-5092; M.T. 0000-0001-5907-9161; E.M. 0000-0001-5964-5258; K.A.T. 0000-0003-4850-5094

Corresponding Author: Mehmet Necmeddin Sutaşır, MD;
E-mail: drmehmetns@gmail.com
Received Date: 09.12.2025 **Accepted Date:** 18.05.2026
Epub: 15.06.2026
Publication Date: 07.08.2026

J Acad Res Med 2026;16(2):70-78

disorders worldwide in 2020 (10). These data indicate how the pandemic has damaged people's mental health and how crucial it is to know how emergency psychiatric services have changed.

Some groupings and diagnostic classifications have changed a lot during the pandemic. Emergency medical service encounters due to alcohol have seen substantial proportionate increases, with research indicating rates 7-24% higher in 2020 compared to pre-pandemic baseline periods (11,12). The pandemic has also been linked to more domestic violence and intimate partner violence. Systematic evaluations have shown that reported cases have gone up a lot in several countries (13,14). Also, mental health in children and teens has become a major concern. Evidence suggests a complex pattern of fewer total visits but greater proportional representation in emergency medical service utilization (15,16). Türkiye reported its first incidence of COVID-19 on March 11, 2020, and then implemented a range of public health measures to stop the virus from spreading, including as partial lockdowns, curfews, and limits on public gatherings (17). İstanbul, which has more than 15 million people and is the largest metropolitan area in the country, is a great place to look at how the use of psychiatric emergency services changed during the pandemic. Because the city has a vast healthcare system and a diversified population, it is conceivable to get results that could help other larger cities with similar issues.

Although international studies have documented heterogeneous changes in psychiatric emergency service utilization during the COVID-19 pandemic, large-scale population-based evidence from prehospital emergency medical systems in metropolitan settings remains limited. In particular, the epidemiological dynamics of psychiatric emergencies within centralized emergency medical service infrastructures in rapidly urbanizing regions have not been sufficiently characterized.

The present study aims to elucidate changes in prehospital psychiatric emergency service utilization within the centralized metropolitan emergency medical services system of İstanbul during the COVID-19 pandemic.

We hypothesized that the pandemic period would be associated with increased utilization of psychiatric emergency services, shifts in demographic distribution, and greater clinical acuity among presentations.

METHODS

This study employed a retrospective observational time-period comparison design, in which routinely collected electronic health records were used to characterize psychiatric emergency presentations across two predefined temporal intervals. This approach enabled the examination of historical patterns in service utilization, demographic distributions, and clinical characteristics within the study population, while facilitating a structured comparison of outcomes between the pre-pandemic and pandemic periods within a quasi-experimental analytical framework.

Study Setting

The study was carried out at the İstanbul Provincial Health Directorate, Department of Emergency Health Services, which coordinates prehospital emergency medical services for all 39 administrative districts of İstanbul, Türkiye's largest metropolitan area with a population exceeding 15 million. This institution was chosen for multiple strategic reasons: it has a comprehensive, centralized electronic database that records all prehospital psychiatric emergency encounters; the systematic data recording protocols guarantee consistency throughout the study period; and the institutional scope offers population-level representativeness for the İstanbul metropolitan area. The emergency health events system [ASOS (*acil sağlık olayları sistemi*)] database, established by the İstanbul Provincial Health Directorate, functioned as the main data source. This electronic platform integrates real-time documentation of emergency medical service encounters, encompassing demographic data, clinical assessments, preliminary diagnoses, triage classifications, and disposition decisions.

The analyzed dataset exclusively represents prehospital emergency medical service encounters recorded prior to hospital admission and does not include hospital emergency department clinical records.

Study Period

This analysis covered a period of 36 months, divided into two equal 18-month intervals: a pre-pandemic reference period from September 10, 2018, to March 10, 2020, and a pandemic period from March 11, 2020, to September 11, 2021. The date of March 11, 2020, was chosen to align with the World Health Organization's formal declaration of COVID-19 as a global pandemic. The use of equal-duration comparison periods aimed to reduce seasonal confounding effects and facilitate a balanced statistical comparison. This study analyzes historical data, reflecting conditions during the first 18 months of the pandemic. The temporal distance may restrict immediate clinical applicability; however, the findings offer important documentation of pandemic-era patterns that can guide future public health emergency preparedness planning.

Study Population and Eligibility Criteria

The target population comprised all individuals presenting to İstanbul Provincial Health Directorate, Emergency Health Services who received preliminary psychiatric diagnoses during the defined study period. Rather than employing probability sampling with predetermined sample size calculations, we adopted a population census methodology, systematically including all patients meeting the eligibility criteria. This comprehensive approach maximized statistical power and ensured complete representation of the population served by the emergency medical system.

Inclusion criteria: Were defined as follows: (1) presentation to İstanbul Provincial Health Directorate, Emergency Health Services during the study period (September 10, 2018 through September 11, 2021); (2) assignment of a preliminary psychiatric diagnosis by

J Acad Res Med 2026;16(2):70-78

emergency medicine physicians operating within the prehospital emergency medical service system, with consultation from other physicians when clinically indicated, and documented using provisional International Classification of Diseases, 10th revision (ICD-10) diagnostic codes, documented using ICD-10 diagnostic codes within the F00-F99 chapter (mental and behavioral disorders); and (3) availability of complete demographic data and clinical documentation necessary for analysis.

Exclusion criteria encompassed: (1) records with incomplete or missing essential data fields precluding accurate classification; (2) encounters documented with primary diagnoses outside the psychiatric spectrum; and (3) duplicate records representing the same clinical encounter.

Data Collection and Chart Review Process

Data extraction was conducted via a systematic review of electronic health records stored in the ASOS database. Before commencing data collection, we established a standardized protocol for data extraction that outlined all variables to be recorded, provided operational definitions for each data element, and detailed procedures for addressing ambiguous or conflicting information. This protocol was tested with a random sample of 100 records to evaluate feasibility and improve data collection methods. The chart review process was performed by two trained research personnel (MNS and EA), who independently extracted data from the electronic database. Both reviewers underwent standardized training on the data extraction protocol, which encompassed instruction on diagnostic coding conventions, triage classification criteria, and variable definitions. To ensure inter-rater reliability, a randomly selected subset of 5% of all records (n=4,590) was independently reviewed by both extractors. Discrepancies among reviewers were addressed through consensus discussions, with a senior investigator providing adjudication when required. The variables systematically extracted from each eligible record included: patient demographic characteristics (age at presentation, biological sex); temporal parameters (date and time of emergency service contact, day of week, season); clinical information (preliminary psychiatric diagnosis coded per ICD-10, triage category assignment using the national three-tier color-coded system, consciousness level assessment); and administrative data (forensic case designation, disposition outcome). Consciousness level was assessed by emergency medical personnel during initial patient evaluation using standard clinical judgment and categorized according to predefined operational definitions within the emergency medical service system. All extracted data were entered into a secure, password-protected electronic database, with access restricted to authorized research personnel.

These diagnostic classifications represent preliminary clinical assessments made in an emergency context and should not be interpreted as definitive psychiatric diagnoses.

Study Variables

Primary outcome variables included: (1) Frequency of psychiatric emergency presentations, quantified as the total number of encounters during each study period; and (2) distribution of psychiatric diagnoses, categorized according to ICD-10 diagnostic groupings.

Secondary outcome variables encompassed: Triage acuity classification (red code indicating life-threatening/critical status, yellow code indicating urgent status, green code indicating routine/non-urgent status); consciousness level (alert, altered, confused, unconscious, semicomatose); and forensic case designation (binary classification based on medico-legal involvement).

Independent variables included: Study period (pre-pandemic versus pandemic, treated as the primary exposure variable); patient age (analyzed both as a continuous variable and categorically in clinically meaningful groupings: <18 years, 18-30 years, 31-45 years, 46-59 years, ≥60 years); biological sex (male, female); and season of presentation (winter, spring, summer, autumn).

The absolute number of prehospital psychiatric emergency encounters increased during the pandemic period; however, the absence of denominator data on total emergency medical service utilization limits the ability to determine whether this reflects a proportional increase in psychiatric emergencies.

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics software, version 20.0 (IBM Corporation, Armonk, New York, United States). The analytic approach proceeded through several sequential phases.

Descriptive statistics: Continuous variables were summarized using measures of central tendency and dispersion (mean ± standard deviation for normally distributed data; median with interquartile range for skewed distributions). Normality of continuous variable distributions was assessed using the Kolmogorov-Smirnov test. Categorical variables were expressed as absolute frequencies and percentages.

Comparative analyses: Comparisons of categorical variables between pre-pandemic and pandemic periods were conducted using Pearson's chi-square (χ^2) test for independence. When expected cell frequencies were less than five in more than 20% of cells, Fisher's exact test was employed as an alternative. For continuous variables, independent samples t-tests were used for normally distributed data, while the Mann-Whitney U test was applied for non-normally distributed variables.

Effect size estimation: Relative percentage change between periods was calculated using the formula: [(pandemic period value - pre-pandemic period value)/pre-pandemic period value]×100. This metric provides a standardized measure of the magnitude of change observed between the two study periods. Absolute differences were also reported to provide context regarding the actual numerical changes in patient volumes.

J Acad Res Med 2026;16(2):70-78

A two-tailed p-value threshold of <0.05 was established for determining statistical significance across all analyses. Given the large sample size, we emphasize interpretation of effect sizes and clinical meaningfulness alongside statistical significance, as even small differences may achieve statistical significance with sufficiently large samples.

Multivariable logistic regression analyses were conducted to estimate adjusted odds ratios (aOR) for key outcomes. Covariates included in the models were age, sex, season of presentation, and triage category, selected based on clinical relevance and prior literature (theory-driven approach). All variables were entered simultaneously into the models using the enter method, without stepwise selection procedures.

Model fit was assessed using the Hosmer-Lemeshow goodness-of-fit test and Nagelkerke R^2 statistics. In addition, clinically plausible interaction terms, including sex×study period and age×study period, were tested; however, no statistically significant interactions were identified, and these terms were not retained in the final models.

Ethical Considerations

The study protocol obtained ethical approval from the University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital Clinical Research Ethics Committee (approval no: 2022/260, decision no: 2022-17-02, date: 05.09.2022). The retrospective investigation employed pre-existing, de-identified administrative data, leading the ethics committee to waive the requirement for individual informed consent. All data handling procedures complied with relevant data protection regulations, with personal identifying information eliminated before analysis. The research adhered to the ethical guidelines established in the Declaration of Helsinki and its subsequent revisions.

RESULTS

Study Population and Overall Findings

Over the 36-month study period (September 10, 2018-September 11, 2021), 91,803 patients received preliminary psychiatric diagnoses. The pre-pandemic period included 41,969 patients (45.7%), while the pandemic period included 49,834 patients (54.3%), representing an 18.7% relative increase (χ^2=671.34, df=1, p<0.001, Cramer's V=0.086). The odds of psychiatric emergency presentation during the pandemic period was 1.19 [95% confidence interval (CI): 1.17-1.21]. Although statistically significant, the effect size indicates a small practical effect. These findings are summarized in Table 1.

Demographic Characteristics

Gender distribution: During the pre-pandemic period, female patients predominated (n=25,889, 61.7%), with a female-to-male ratio of 1.6:1. This ratio narrowed to 1.2:1 during the pandemic period (females: n=27,060, 54.3%; males: n=22,768, 45.7%). Multivariate logistic regression analysis, adjusting for age and season, demonstrated that male gender was significantly associated with pandemic-period presentation (aOR: 1.36, 95% CI: 1.32-1.40, p<0.001, Cramer's V=0.075), reflecting a 41.6% relative increase in male presentations.

Age distribution: The pediatric population (<18 years) showed a marked decline from 11.6% (n=4,875) to 8.1% (n=4,053) of presentations (aOR: 0.67, 95% CI: 0.64-0.70, p<0.001, Cramer's V=0.060), representing a 33% reduction in odds after controlling for gender and season. Conversely, elderly patients (≥60 years) demonstrated a modest increase from 10.2% (n=4,280) to 11.0% (n=5,480) (aOR: 1.09, 95% CI: 1.04-1.14, p<0.001). Young and middle-aged adults (18-45 years) remained stable at approximately 60% of total encounters across both periods. Detailed demographic data are presented in Table 2.

Table 1. Overall psychiatric emergency encounters by study period

Study period	n	%	OR (95% CI)	p-value	Cramer's V
Pre-pandemic	41,969	45.7	Reference	-	-
Pandemic	49,834	54.3	1.19 (1.17-1.21)	<0.001	0.086
Total	**91,803**	**100.0**	-	-	-

χ^2=671.34, df=1, p<0.001, OR: Odds ratio, CI: Confidence interval

Table 2. Demographic characteristics with adjusted odds ratios

Characteristic	Pre-n	Pre-%	Pan-n	Pan-%	aOR (95% CI)	p-value
Gender						
Female	25,889	61.7	27,060	54.3	Reference	-
Male	16,075	38.3	22,768	45.7	1.36 (1.32-1.40)	<0.001*
Age group						
<18 years	4,875	11.6	4,053	8.1	0.67 (0.64-0.70)	<0.001*
18-30 years	11,884	28.3	14,655	29.4	Reference	-
31-45 years	13,214	31.5	16,156	32.4	0.99 (0.96-1.03)	0.742
≥60 years	4,280	10.2	5,480	11.0	1.09 (1.04-1.14)	<0.001*

*: P<0.001 after Bonferroni correction, aOR: Adjusted odds ratio (adjusted for age, gender, and season), CI: Confidence interval

J Acad Res Med 2026;16(2):70-78

Distribution of Psychiatric Diagnoses

Anxiety disorders remained the most common diagnosis but declined proportionally from 46.3% (n=19,436) to 40.6% (n=20,237) (aOR: 0.79, 95% CI: 0.77-0.81, p<0.001, Cramer's V=0.058). Dissociative disorders showed a more pronounced decline from 22.2% (n=9,305) to 14.5% (n=7,223) (aOR: 0.59, 95% CI: 0.57-0.61, p<0.001, Cramer's V=0.099). Schizophrenia spectrum disorders remained stable at approximately 5% (aOR: 0.93, 95% CI: 0.88-0.98, p=0.008). Complete diagnostic data are presented in Table 3. The p-values reported correspond to comparisons between the pre-pandemic and pandemic periods for each diagnostic category, obtained from multivariable logistic regression models adjusted for age, sex, season, and triage category.

Alcohol-related disorders: Alcohol-related presentations increased from 2.9% to 13.7% of all encounters. Multivariate analysis, adjusting for age, gender, season, and triage category, yielded an aOR of 5.34 (95% CI: 5.01-5.70, p<0.001), with a large effect size (Cramer's V=0.205). This finding was robust across gender-stratified analyses (males: aOR: 5.12; females: aOR: 6.01). The risk ratio was 4.74 (95% CI: 4.46-5.03), and the population attributable fraction was estimated at 81.2%. Complete diagnostic data are presented in Table 3.

Clinical Severity Indicators

Triage acuity: Ordinal logistic regression demonstrated a significant shift toward higher-acuity presentations (proportional OR: 1.68, 95% CI: 1.64-1.72, p<0.001; Brant test χ^2=2.34, p=0.126). Red code assignments increased from 18.2% to 20.1% (aOR: 1.13, 95% CI: 1.09-1.17), yellow code from 49.8% to 59.7% (aOR: 1.49, 95% CI: 1.45-1.53), while green code declined from 31.4% to 19.9% (aOR: 0.54, 95% CI: 0.52-0.56). The overall effect size was medium (Cramer's V=0.131). Sensitivity analysis excluding alcohol-related cases showed attenuated but persistent effects (proportional OR: 1.42, 95% CI: 1.38-1.46).

Forensic cases: Forensic case designations increased from 3.97% (n=1,667) to 5.83% (n=2,907) (aOR: 1.50, 95% CI: 1.41-1.60, p<0.001; risk ratio=1.47, 95% CI: 1.39-1.56; Cramer's V=0.045). Clinical severity data are detailed in Table 4.

Consciousness Level Assessment

Ordinal logistic regression revealed a significant association between pandemic period and impaired consciousness (proportional OR: 1.38, 95% CI: 1.28-1.49, p<0.001). Notably, confused consciousness increased from 0.9% (n=385) to 1.5% (n=732) (aOR: 1.60, 95% CI: 1.41-1.82, p<0.001). When alcohol-related presentations were included as a covariate, this association was attenuated but remained significant (aOR: 1.28, 95% CI: 1.12-1.47, p<0.001). Complete data are presented in Table 5.

Multiple Comparisons Correction

With 15 primary comparisons, the Bonferroni-adjusted significance threshold was α=0.0033. All reported significant associations maintained statistical significance after this correction (all p<0.001).

Table 3. Distribution of psychiatric diagnoses with adjusted odds ratios, effect sizes, and p-values

Diagnosis	Pre-n	Pre-%	Pan-n	Pan-%	aOR (95% CI)	p-value	Cramer's V
Anxiety disorders	19,436	46.3	20,237	40.6	0.79 (0.77-0.81)	<0.001*	0.058
Dissociative disorders	9,305	22.2	7,223	14.5	0.59 (0.57-0.61)	<0.001*	0.099
Schizophrenia spectrum	2,248	5.4	2,520	5.1	0.93 (0.88-0.98)	0.008	-
Bipolar disorder	1,732	4.1	1,929	3.9	0.94 (0.88-1.00)	0.006	-
Alcohol-related	**1,211**	**2.9**	**6,811**	**13.7**	**5.34 (5.01-5.70)**	<0.001*	**0.205**

Note: Percentages do not sum to 100% because only the most frequent diagnostic categories are presented. Less common diagnostic groups are not displayed to maintain clarity and readability
*: Statistically significant after Bonferroni correction (adjusted α=0.0033), aOR: Adjusted odds ratio, CI: Confidence interval

Table 4. Clinical severity indicators

Clinical feature	Pre-n	Pre-%	Pan-n	Pan-%	aOR (95% CI)	p-value
Triage category (proportional OR: 1.68, 95% CI: 1.64-1.72, Cramer's V=0.131)						
Red (critical)	7,649	18.2	10,024	20.1	1.13 (1.09-1.17)	<0.001*
Yellow (urgent)	20,904	49.8	29,757	59.7	1.49 (1.45-1.53)	<0.001*
Green (non-urgent)	13,174	31.4	9,899	19.9	0.54 (0.52-0.56)	<0.001*
Forensic status (Cramer's V=0.045)						
Forensic case	1,667	3.97	2,907	5.83	1.50 (1.41-1.60)	<0.001*
Non-forensic case	40,302	96.03	46,927	94.17	Reference	-

*: P<0.001 after Bonferroni correction, aOR: Adjusted odds ratio, CI: Confidence interval

J Acad Res Med 2026;16(2):70-78

Table 5. Consciousness level at presentation

Consciousness	Pre-n	Pre-%	Pan-n	Pan-%	aOR (95% CI)	p-value
Alert	29,005	69.1	36,321	72.9	Reference	-
Altered	388	0.9	546	1.1	1.12 (0.98-1.28)	0.089
Confused	**385**	**0.9**	**732**	**1.5**	**1.60 (1.41-1.82)**	**<0.001***
Unconscious	76	0.2	93	0.2	0.98 (0.72-1.33)	0.876
Semicomatose	12	0.03	25	0.05	1.66 (0.83-3.32)	0.151

Proportional OR: 1.38 (95% CI: 1.28-1.49); *: P<0.001, CI: Confidence interval, aOR: Adjusted odds ratio

Table 6. Summary of key findings with effect sizes

Variable	aOR (95% CI)	Cramer's V	Effect size
Overall presentations	1.19 (1.17-1.21)	0.086	Small
Male gender	1.36 (1.32-1.40)	0.075	Small
Pediatric (<18 y)	0.67 (0.64-0.70)	0.060	Small-medium
Alcohol-related disorders	**5.34 (5.01-5.70)**	**0.205**	**Large**
Triage shift	1.68 (1.64-1.72)†	0.131	Medium
Forensic cases	1.50 (1.41-1.60)	0.045	Small

†: Proportional OR; effect size: V<0.1 small, 0.1-0.3 medium, >0.3 large all p<0.001 after Bonferroni correction (adjusted α=0.0033), aOR: Adjusted odds ratio, CI: Confidence interval

DISCUSSION

This population-based observational study of 91,803 prehospital psychiatric emergency encounters provides detailed insights into changes in service utilization patterns during the COVID-19 pandemic. The study employed a retrospective observational time-period comparison design to examine changes in psychiatric emergency service utilization across two predefined temporal intervals. Therefore, the observed changes should be interpreted as reflecting shifts in prehospital help-seeking behavior and emergency medical service utilization patterns rather than direct changes in hospital-based psychiatric emergency demand.

Within this context, our analysis revealed six key associations (Table 6): (1) 18.7% increase in overall presentations (OR: 1.19, small effect); (2) disproportionate increase in male presentations (aOR: 1.36, small effect); (3) 33% reduction in pediatric presentations (aOR: 0.67, small-to-medium effect); (4) 373.7% increase in alcohol-related disorders (aOR: 5.34, large effect); (5) shift toward higher-acuity triage (proportional OR: 1.68, medium effect); and (6) increased forensic cases (aOR: 1.50, small effect). Our data reveal numerous significant correlations that require careful interpretation within the context of the study's quasi-experimental design and the wider international literature. The 18.7% increase in psychiatric emergency medical service encounters throughout the pandemic era contrasts with certain international statistics showing initial decreases in psychiatric emergency utilization (6,18). This finding should be interpreted cautiously, as denominator data on total emergency medical service utilization were not available to determine whether the proportional burden of psychiatric emergencies changed during the study period. This conclusion aligns with data from other situations that demonstrated subsequent increases following the relaxation of early lockdown measures (7,19). A study conducted in Verona, Italy, documented significant reductions in psychiatric visits in 2020 (-23.3%) and 2021 (-16.3%) compared to 2019; nonetheless, increases were observed among specific groups, particularly young adults and individuals experiencing psychosis (20). The differing results from different healthcare systems and time periods show how important contextual factors are in affecting how often people use psychiatric emergencies. These findings should be interpreted within the context of a prehospital emergency medical service system, where patient selection, diagnostic processes, and clinical priorities differ from those of hospital-based emergency departments.

The absence of a contemporaneous control group in our pre-post design limits the ability to definitively attribute observed changes to factors related to the pandemic. Secular trends in the incidence of psychiatric disorders, alterations in healthcare-seeking behavior independent of the pandemic, and adjustments to service delivery models may all account for the observed differences. The observed increase in our study may reflect factors specific to the local healthcare environment, such as the structure of the emergency medical service system and potential changes in patient flow from other facilities. The reduction of the female-to-male ratio from 1.6:1 to 1.2:1 during the pandemic, primarily due to a 41.6% rise in male presentations, indicates a significant demographic change. International literature indicates elevated rates of anxiety and depressive symptoms among females during the pandemic (21,22); however, patterns of emergency medical service utilization

J Acad Res Med 2026;16(2):70-78

may diverge from population-based prevalence estimates. The rise in male presentations may indicate varying exposure to pandemic-related stressors, such as economic instability and job loss, which have been linked to negative mental health outcomes in previous economic crises (23). Other explanations include changes in how men seek treatment or differences in how easy it is for men and women to go to different mental health facilities, which may have affected male and female patients in different ways. In addition, the observed increase in presentations among older adults (≥60 years) warrants consideration. This trend may be attributable to several factors, including increased social isolation during pandemic-related restrictions, reduced access to routine outpatient care, and heightened caregiver burden. Furthermore, neuropsychiatric complications associated with medical illness, including COVID-19 infection, may have contributed to increased psychiatric emergency presentations in this population (24).

The 30.2% drop in pediatric presentations is in line with findings from around the world that show fewer children going to the emergency room during the pandemic (15,25). A report from the Centers for Disease Control and Prevention indicated that the number of children who went to the emergency room for mental health reasons dropped by 43% between March and October 2020. However, the percentage of these trips went up a lot (26). The decline in pediatric presentations noted in our study must be interpreted with caution, as it may indicate diminished access to care, parental hesitance to pursue emergency services during the pandemic, or authentic fluctuations in the incidence of acute psychiatric crises among children and adolescents. School closures, which removed a significant avenue for recognizing mental health issues, may have led to postponed or missed treatment (27). The diagnostic distribution variations we saw in our study show complex patterns that need to be carefully thought about. Anxiety disorders continued to be the predominant diagnostic category; however, their proportionate decline (46.3% to 40.6%) coincided with absolute increases in case numbers, indicating a shift in the general makeup of psychiatric emergency presentations during the pandemic period. This conclusion is consistent with meta-analytic findings indicating a high frequency of anxiety (31.9%) and depression (31.4%) during the pandemic (9), but the correlation between population prevalence and emergency medical service utilization is complex. The diagnostic distribution observed in this study should be interpreted in light of the preliminary nature of emergency setting diagnoses, which may differ from diagnoses established through comprehensive psychiatric evaluation. This limitation is particularly relevant in prehospital emergency medical service settings, where diagnostic classification is primarily based on rapid clinical assessment and ICD-10 coding practices rather than comprehensive psychiatric evaluation. Interestingly, the proportional representation of anxiety disorder presentations decreased during the pandemic period. Several factors may explain this pattern. First, elevated baseline levels of anxiety in the general population during the pandemic may have led to a relative normalization of anxiety symptoms, reducing the likelihood of emergency service utilization for milder cases. Second, concerns about infection risk and avoidance of healthcare settings may have discouraged individuals with anxiety disorders from seeking emergency care unless symptoms were severe.

In addition, disruptions in outpatient mental health services and a shift toward alternative care modalities, such as telemedicine, may have altered help-seeking pathways and reduced reliance on emergency medical services for anxiety-related conditions. These findings suggest that changes in healthcare utilization patterns, rather than a true reduction in anxiety prevalence, may underlie the observed decrease in emergency presentations.

A notable finding of this study was the marked increase in alcohol-related presentations, rising from 2.9% to 13.7% of all encounters. This magnitude substantially exceeds that reported in most international studies and should therefore be interpreted with caution (11,28). Several factors may contribute to this observation, including a true increase in alcohol consumption and related harms during the pandemic, disruptions in access to routine addiction services, and shifts in help-seeking behavior leading to greater reliance on emergency medical services (29,30). In addition, the prehospital emergency medical service context may influence diagnostic classification, as alcohol-related presentations are often more readily identifiable during rapid clinical assessment compared to other psychiatric conditions. Furthermore, changes in service availability and referral patterns during the pandemic may have resulted in a higher proportion of alcohol-related cases being managed within emergency systems. Taken together, these findings likely reflect a combination of behavioral, healthcare system–related, and methodological factors rather than a single underlying cause. The rise in red (+31.0%) and yellow (+42.4%) triage codes and the drop in green code presentations (-24.9%) during the pandemic suggest that those who came in during that time may have had more severe clinical symptoms. This outcome corresponds with theories positing that pandemic-induced barriers to care may have led to delayed presentation with exacerbated symptoms (31). International data has also shown that psychiatric emergency presentations had more severe symptoms during the pandemic, with studies showing higher rates of suicidal thoughts, psychotic symptoms, and the need for hospitalization (32,33). Nonetheless, other possible reasons for the change in severity must be considered. Changes in triage approaches, patient self-selection (with those exhibiting milder symptoms potentially forgoing emergency care during the pandemic), and modifications to the accessibility of alternate care pathways may all impact the reported trends. In addition, changes in consciousness level observed during the pandemic period warrant further consideration. Increased rates of altered or impaired consciousness may be partly explained by substance-related conditions, including acute intoxication and withdrawal states, particularly in the context of the marked rise in alcohol-related presentations observed in this study. Delirium associated with acute medical illness, including COVID-19 infection, may also have contributed, especially among older adults.

J Acad Res Med 2026;16(2):70-78

Furthermore, delayed healthcare-seeking behavior and reduced access to routine medical care during the pandemic may have resulted in patients presenting at more advanced stages of illness, thereby increasing the likelihood of impaired consciousness at the time of emergency evaluation. Comorbid medical conditions and polypharmacy may represent additional contributing factors. The 46.9% increase in forensic cases aligns with global data showing rises in domestic violence and intimate partner violence during the pandemic (13,14,34); nevertheless, our observational methodology precludes causal conclusion. Several mechanisms may underlie this increase. Pandemic-related restrictions, including lockdown measures and social isolation, may have intensified interpersonal conflicts within households, contributing to higher rates of domestic violence. Economic stressors, unemployment, and uncertainty during the pandemic have also been associated with increased risk of aggression and violence. In addition, increased substance use particularly alcohol consumption—may have further exacerbated the likelihood of forensic incidents.

Reduced access to social support services, legal protection mechanisms, and community-based interventions during the pandemic may have limited opportunities for early intervention, potentially resulting in more severe cases requiring emergency medical attention. These findings are consistent with emerging international literature documenting increased rates of violence and forensic-related presentations during the COVID-19 period (35).

Study Limitations

This study has several strengths, including a large sample size of 91,803 patients, a three-year observation period encompassing both pre-pandemic and pandemic phases, and the use of a comprehensive population-based emergency medical service database. The application of standardized data extraction procedures and independent data review further enhances the reliability and internal validity of the findings. In addition, the metropolitan setting of İstanbul provides valuable insights that may be relevant to other large urban healthcare systems.

However, several limitations should be acknowledged. First, the quasi-experimental pre-post design without a concurrent control group limits the ability to attribute observed changes directly to the pandemic, as underlying temporal trends in psychiatric morbidity and healthcare utilization cannot be fully accounted for. Second, the absence of denominator data representing total emergency medical service utilization restricts the interpretation of proportional changes, and the findings should therefore be understood as reflecting absolute encounter numbers. Third, although season was included as a covariate in the multivariable analyses, detailed seasonal distributions were not presented, which may limit the full assessment of potential seasonal effects.

Fourth, the use of retrospective administrative data introduces the possibility of misclassification and incomplete documentation. Fifth, psychiatric diagnoses were based on preliminary assessments conducted in a prehospital emergency setting, which may limit diagnostic accuracy compared with comprehensive psychiatric evaluation. Finally, the findings are derived from a single metropolitan emergency medical service system and may not be fully generalizable to different healthcare systems, cultural contexts, or pandemic trajectories.

Future studies may benefit from integrating diagnostic categories with triage-based severity measures to provide a more comprehensive understanding of clinical presentation patterns within emergency medical service systems.

CONCLUSION

This observational study highlights significant changes in prehospital psychiatric emergency presentations within the İstanbul emergency medical service system during the COVID-19 pandemic. Key findings include increased overall utilization, demographic shifts towards male and elderly patients, a marked rise in alcohol-related disorders, and a trend towards higher clinical acuity. These findings must be understood as associations, given the constraints of a quasi-experimental design lacking a control group. The data presented do not establish causal relationships between the pandemic and the observed changes; however, they offer descriptive evidence that can inform the planning of emergency psychiatric services and identify areas that require further investigation using more rigorous methodologies.

Ethics

Ethics Committee Approval: The study protocol obtained ethical approval from the University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital Clinical Research Ethics Committee (approval no: 2022/260, decision no: 2022-17-02, date: 05.09.2022).

Informed Consent: The retrospective investigation employed pre-existing, de-identified administrative data, leading the ethics committee to waive the requirement for individual informed consent.

Footnotes

Author Contributions: Surgical and Medical Practices - M.N.S., Concept - M.T., K.A.T., Design - M.N.S., E.M., K.A.T., Data Collection and/or Processing - E.M., Analysis and/or Interpretation - M.T., Literature Search - M.N.S., Writing - M.N.S., M.T., E.M., K.A.T.

Conflict of Interest: The authors have no conflict of interest to declare.

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

REFERENCES

1. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020 [Internet]. Geneva: World Health Organization; 2020. Available from: URL: https://www.who.int/news-room/speeches/item/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020
2. Salari N, Hosseinian-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. Global Health. 2020; 16: 57.
3. Wu T, Jia X, Shi H, Niu J, Yin X, Xie J, et al. Prevalence of mental health problems during the COVID-19 pandemic: a systematic review and meta-analysis. J Affect Disord. 2021; 281: 91-8.

J Acad Res Med 2026;16(2):70-78

4. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry. 2020; 7: 547-60.
5. Goldenberg MN, Parwani V. Psychiatric emergency department volume during Covid-19 pandemic. Am J Emerg Med. 2021; 41: 233-4.
6. Simpson SA, Loh RM, Cabrera M, Cahn M, Gross A, Hadley A, et al. The impact of the COVID-19 pandemic on psychiatric emergency service volume and hospital admissions. J Acad Consult Liaison Psychiatry. 2021; 62: 588-94.
7. Enrico C, Caldiroli A, Di Brita C, Colmegna F, Nava R, Colzani LC, et al. Profile of patients attending psychiatric emergency care during the coronavirus 2019 (COVID 19) pandemic: a comparative cross-sectional study between lockdown and post-lockdown periods in Lombardy, Italy. Int J Psychiatry Clin Pract. 2022; 26: 132-8.
8. Holland KM, Jones C, Vivolo-Kantor AM, Idaikkadar N, Zwald M, Hoots B, et al. Trends in US emergency department visits for mental health, overdose, and violence outcomes before and during the COVID-19 pandemic. JAMA Psychiatry. 2021; 78: 372-9.
9. Nochaiwong S, Ruengorn C, Thavorn K, Hutton B, Awiphan R, Phosuya C, et al. Global prevalence of mental health issues among the general population during the coronavirus disease-2019 pandemic: a systematic review and meta-analysis. Sci Rep. 2021; 11: 10173.
10. COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet. 2021; 398: 1700-12.
11. Esser MB, Idaikkadar N, Kite-Powell A, Thomas C, Greenlund KJ. Trends in emergency department visits related to acute alcohol consumption before and during the COVID-19 pandemic in the United States, 2018-2020. Drug Alcohol Depend Rep. 2022; 3: 100049.
12. Murthy P, Narasimha VL. Effects of the COVID-19 pandemic and lockdown on alcohol use disorders and complications. Curr Opin Psychiatry. 2021; 34: 376-85.
13. Piquero AR, Jennings WG, Jemison E, Kaukinen C, Knaul FM. Domestic violence during the COVID-19 pandemic - evidence from a systematic review and meta-analysis. J Crim Justice. 2021; 74: 101806.
14. Kourti A, Stavridou A, Panagouli E, Psaltopoulou T, Spiliopoulou C, Tsolia M, et al. Domestic violence during the COVID-19 pandemic: a systematic review. Trauma Violence Abuse. 2023; 24: 719-45.
15. Leeb RT, Bitsko RH, Radhakrishnan L, Martinez P, Njai R, Holland KM. Mental health-related emergency department visits among children aged <18 years during the COVID-19 pandemic - United States, January 1-October 17, 2020. MMWR Morb Mortal Wkly Rep. 2020; 69: 1675-80.
16. Yard E, Radhakrishnan L, Ballesteros MF, Sheppard M, Gates A, Stein Z, et al. Emergency department visits for suspected suicide attempts among persons aged 12-25 years before and during the COVID-19 pandemic - United States, January 2019-May 2021. MMWR Morb Mortal Wkly Rep. 2021; 70: 888-94.
17. Republic of Turkey Ministry of Health. COVID-19 Information Page [Internet]. Ankara: Ministry of Health; 2020. Available from: https://covid19asi.saglik.gov.tr/?_Dil=2
18. Gómez-Ramiro M, Fico G, Anmella G, Vázquez M, Sagué-Vilavella M, Hidalgo-Mazzei D, et al. Changing trends in psychiatric emergency service admissions during the COVID-19 outbreak: report from a worldwide epicentre. J Affect Disord. 2021; 282: 26-32.
19. Dragovic M, Pascu V, Hall T, Ingram J, Waters F. Emergency department mental health presentations before and during the COVID-19 outbreak in Western Australia. Australas Psychiatry. 2020; 28: 627-31.
20. Bodini L, Bonetto C, Maccagnani A, Bonora A, Polati E, Ricci G, et al. Changes in emergency psychiatric consultations in time of COVID-19: a retrospective observational study in the Verona Academic Hospital over the two pandemic years 2020-2021. BMC Emerg Med. 2023; 23: 18.
21. Xiong J, Lipsitz O, Nasri F, Lui LMW, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. J Affect Disord. 2020; 277: 55-64.
22. Pierce M, Hope H, Ford T, Hatch S, Hotopf M, John A, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. Lancet Psychiatry. 2020; 7: 883-92.
23. Frasquilho D, Matos MG, Salonna F, Guerreiro D, Storti CC, Gaspar T, et al. Mental health outcomes in times of economic recession: a systematic literature review. BMC Public Health. 2016; 16: 115.
24. Aytaç HM, Pehlivan S. Viral pandemics as possible psycho-immunological causes of psychiatric symptoms: from past to present. Sağlık Bilimlerinde İleri Araştırmalar Dergisi. 2020; 3(Suppl 1): S92-8.
25. DeLaroche AM, Rodean J, Aronson PL, Fleegler EW, Florin TA, Goyal M, et al. Pediatric emergency department visits at US children's hospitals during the COVID-19 pandemic. Pediatrics. 2021; 147: e2020039628.
26. Villas-Boas SB, Kaplan S, White JS, Hsia RY. Adolescent total and mental health-related emergency department visits during the COVID-19 pandemic. JAMA Netw Open. 2023; 6: e2336463.
27. Golberstein E, Wen H, Miller BF. Coronavirus disease 2019 (COVID-19) and mental health for children and adolescents. JAMA Pediatr. 2020; 174: 819-20.
28. Myran D, Friesen EL, Pugliese M, Milani C, Kurdyak P, Saraswat M, et al. Changes in health service use due to alcohol during the COVID-19 pandemic among individuals with and individuals without pre-existing alcohol-related medical diagnoses. Can J Public Health. 2023; 114: 185-94.
29. Pollard MS, Tucker JS, Green HD Jr. Changes in adult alcohol use and consequences during the COVID-19 pandemic in the US. JAMA Netw Open. 2020; 3: e2022942.
30. Roberts A, Rogers J, Mason R, Siriwardena AN, Hogue T, Whitley GA, et al. Alcohol and other substance use during the COVID-19 pandemic: a systematic review. Drug Alcohol Depend. 2021; 229: 109150.
31. Rosenbaum L. The untold toll - the pandemic's effects on patients without Covid-19. N Engl J Med. 2020; 382: 2368-71.
32. Goldschmidt T, Kippe Y, Finck A, Adam M, Hamadoun H, Winkler JG, et al. Correction to: psychiatric presentations and admissions during the first wave of Covid-19 compared to 2019 in a psychiatric emergency department in Berlin, Germany: a retrospective chart review. BMC Psychiatry. 2023; 23: 164. Erratum for: BMC Psychiatry. 2023; 23: 38.
33. Bolt J, Patel F, Stone L, Pandian D, Manuel MM, Gaines N. Impact of COVID-19 on pediatric mental and behavioral health visits to the emergency department. Pediatr Emerg Care. 2022; 38: 409-15.
34. Telles LEB, Valença AM, Barros AJS, da Silva AG. Domestic violence in the COVID-19 pandemic: a forensic psychiatric perspective. Braz J Psychiatry. 2021; 43: 233-4.
35. Aytac HM. Treatment of major depression with psychotic features and Cotard's syndrome after COVID-19 ınfection in a previously healthy patient: a case report. CNS Neurol Disord Drug Targets. 2023; 22: 614-7.

DOI: 10.4274/jarem.galenos.2026.90582
J Acad Res Med 2026;16(2):79-85

Keratoacanthoma Revisited: Clinicopathological Features and Malignant Transformation in 115 Patients

Mustafa Ürün[1], Gizem Karapınar[1], Merve Bağış[2], Yıldız Gürsel Ürün[1], Nuray Can[2]

[1]Trakya University Faculty of Medicine, Department of Dermatology and Venereology, Edirne, Türkiye
[2]Trakya University Faculty of Medicine, Department of Pathology, Edirne, Türkiye

Cite this article as: Ürün M, Karapınar G, Bağış M, Gürsel Ürün Y, Can N. Keratoacanthoma revisited: clinicopathological features and malignant transformation in 115 patients. J Acad Res Med. 2026;16(2):79-85

ABSTRACT

Objective: To evaluate the demographic, clinical, and histopathological characteristics of patients diagnosed with keratoacanthoma (KA) and to identify features that may aid in distinguishing benign KA from KA with malignant transformation.

Methods: This retrospective study included 115 patients with histopathologically confirmed KA who underwent total excision. Demographic, clinical, and histopathological data were reviewed and analyzed. Statistical comparisons were performed to identify parameters differentiating benign KA from KA with malignant transformation.

Results: The mean age of the patients was 65.1±14.8 years, with a male-to-female ratio of 1.3:1. Solitary KA was the most common clinical presentation (n=114, 99.1%), with lesions predominantly located in the head and neck region (n=88, 76.5%). Histopathological evaluation revealed features of malignant transformation in 36.5% (n=42) of cases. Maximum tumor diameter and depth of invasion were significantly greater in cases exhibiting malignant transformation (p=0.012 and p<0.001, respectively).

Conclusion: Our findings suggest that histopathological parameters, rather than clinical and sociodemographic characteristics, are more strongly associated with malignant transformation in KA.

Keywords: Keratoacanthoma, histopathology, malignant transformation

INTRODUCTION

Keratoacanthoma (KA) is classified among non-melanoma skin cancer (NMSC) and most commonly arises in the head and neck region (1). It is generally regarded as a low-grade cutaneous tumor arising from the hair follicle (2). Its precise incidence remains uncertain due to frequent misdiagnosis as cutaneous squamous cell carcinoma (cSCC); however, available data suggest an estimated annual incidence of approximately 150 cases per 100,000 individuals (3).

The diagnosis of KA is based on three key components: (i) a characteristic clinical presentation of a rapidly developing crateriform lesion over weeks to months (Figure 1A and B); (ii) a distinctive three-stage clinical course in untreated cases, comprising proliferation, stabilization, and regression phases; and (iii) histopathological confirmation (4). Although histopathological findings vary depending on the stage, the stabilization phase typically demonstrates a crateriform architecture with epidermal "lipping" (overhanging epidermal edges) at the lesion margins (Figure 2) (5).

According to the current World Health Organization classification, KA cannot be reliably distinguished histologically from well-differentiated invasive cSCC and is therefore considered a variant of cSCC (6). KA and cSCC share overlapping clinical and histopathological features, which may substantially complicate the diagnostic process (1).

This study aimed to analyze the demographic, clinical, and histopathological characteristics of patients diagnosed with KA to provide further insight into its clinical behavior and management. Additionally, the study sought to identify features associated with malignant transformation in KA.

METHODS

Ethical approval for this study was obtained from the Ethics Committee of Trakya University Faculty of Medicine (protocol no:

ORCID IDs of the authors: M.Ü. 0000-0003-1272-8149; G.K. 0009-0007-2050-5902; M.B. 0000-0001-7301-1673; Y.G.Ü. 0000-0002-6137-8395; N.C. 0000-0003-2554-6432

Corresponding Author: Mustafa Ürün, Asst. Prof.;
E-mail: mustafaurun@trakya.edu.tr
Received Date: 05.03.2026 **Accepted Date:** 26.05.2026
Epub: 19.07.2026
Publication Date: 07.08.2026

J Acad Res Med 2026;16(2):79-85

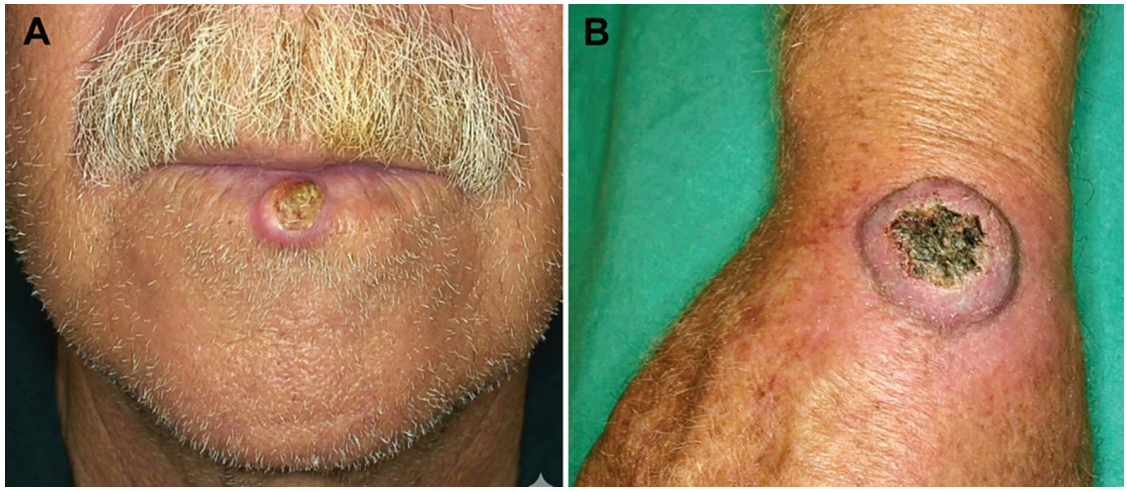

Figure 1. Clinical appearance of a crateriform lesion with central keratinous material (A). Clinical appearance of giant keratoacanthoma (B)

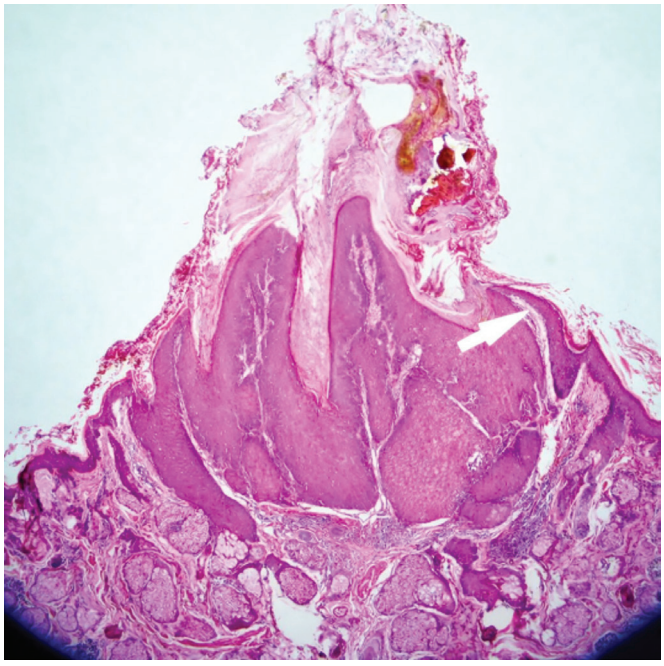

Figure 2. Keratin-filled crateriform epidermal proliferation demonstrating an inverted growth pattern, with peripheral epidermal "lipping" (overhanging epidermal edges) (white arrow) (H&E, ×100)

H&E: Hematoxylin and eosin

TÜTF-GOBAEK 2025/380, decision no: 17/03, date: 22.09.2025).

The characteristics of patients with histopathologically confirmed KA after complete excision between 2015 and 2025 were retrospectively evaluated. Sociodemographic variables (age at diagnosis, sex, family history, and season of diagnosis) and clinical variables (immunosuppression status, tumor location on sun-exposed body sites, rapid tumor growth, history of trauma, history of prior NMSC, presence of concomitant actinic keratosis or NMSC at the time of diagnosis, pruritus, and recurrence) were obtained through a comprehensive review of medical records.

KA was clinically classified into two main categories: solitary and multiple forms. Solitary KAs were further subdivided into five subgroups: classical (typical), giant KA (diameter ≥2 cm), KA centrifugum marginatum, subungual KA, and mucosal KA (4,7,8).

Histopathological parameters including maximum tumor diameter (mm), depth of invasion (mm), ulceration, perineural and/or lymphovascular invasion, and lymphohistiocytic infiltration—were recorded for each case. Missing pathological data were reassessed and completed through re-evaluation by two experienced pathologists (N.C. and M.B.). Based on the criteria reported by Vîlcea et al. (9), cases were categorized as benign KA or KA with malignant transformation; the latter was defined histopathologically as KAs containing areas of carcinoma *in situ*, microcarcinoma, well-differentiated SCC, moderately differentiated SCC, or acantholytic SCC.

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were presented as mean ± standard deviation for normally distributed continuous variables, median (interquartile range, range) for non-normally distributed continuous variables, and number (n) and percentage (%) for categorical variables. The normality of continuous variables was assessed using the Shapiro-Wilk test together with visual inspection of histograms and Q-Q plots. Comparisons between the benign KA and KA with malignant transformation groups were performed using the independent samples t-test for normally distributed continuous variables and the Mann-Whitney U test for non-normally distributed continuous variables. Categorical variables were compared using the Pearson chi-square test or Fisher's exact test, as appropriate. To identify factors associated with malignant transformation, univariate and multivariate binary logistic regression analyses were performed, and the results are presented as odds ratios (OR) with 95% confidence intervals (CI). A two-sided p-value <0.05 was considered statistically significant.

J Acad Res Med 2026;16(2):79-85

RESULTS

The mean age of the 115 patients with KA was 65.1±14.8 years, and the male-to-female ratio was 1.3:1. The highest proportion of diagnoses occurred in winter (n=38, 33.0%). A history of trauma was reported in 14.7% (n=17) of the patients, whereas rapid tumor growth was observed in 69.5% (n=80). The predominant clinical presentation was solitary KA (n=114, 99.1%), and classical KA was the most common clinical subtype (n=109, 94.8%). Because 20 patients were lost to follow-up, recurrence status was evaluated in the remaining 95 patients with available follow-up data. Accordingly, the recurrence rate was calculated as 9.5% (n=9) (Table 1).

Regarding the anatomical distribution, KA was most frequently located in the head and neck region (n=88, 76.5%). Among lesions located in the head region, the lips (n=16, 18.2%) and forehead (n=13, 14.8%) were the most commonly involved sites (Figure 3A and B).

On histopathological evaluation, 36.5% (n=42) of cases exhibited features of malignant transformation. Lymphohistiocytic infiltration was observed in 63.5% (n=73) of cases, whereas neither perineural nor lymphovascular invasion was identified in any case (Table 2).

When patients with benign KA were compared with those with KA exhibiting malignant transformation, maximum tumor diameter and depth of invasion were significantly greater in the malignant transformation group (p=0.012 and p=0.040, respectively) (Table 3).

In univariate logistic regression analysis, the largest tumor diameter and depth of tumor invasion were significantly associated with malignant transformation. The OR was 1.092 for largest tumor diameter (95% CI: 1.012-1.179, p=0.023) and 1.345 for depth of tumor invasion (95% CI: 1.006-1.799, p=0.046) (Table 4).

However, in multivariate analysis, none of the variables remained statistically significant as an independent predictor. The largest tumor diameter showed borderline significance (OR: 1.072, 95%

Table 1. Baseline demographic, clinical, and tumor characteristics of patients with keratoacanthoma (KA)

Characteristic		Total (n=115)
Age at diagnosis, years, mean ± SD*		65.1±14.8
Sex, n (%)	Female	50 (43.5)
	Male	65 (56.5)
Family history n (%)	Yes	7 (6.0)
	No	108 (94.0)
Season at diagnosis, n (%)	Spring	23 (20.0)
	Summer	28 (24.3)
	Autumn	26 (22.6)
	Winter	38 (33.0)
Immunosuppression, n (%)	Yes	21 (18.3)
	No	94 (81.7)
Sun-exposed body site, n (%)	Yes	99 (86.1)
	No	16 (13.9)
History of trauma, n (%)	Yes	17 (14.7)
	No	98 (85.3)
Rapid tumor growth, n (%)	Yes	80 (69.5)
	No	35 (30.5)
Solitary vs multiple KA, n (%)	Solitary	114 (99.1)
	Multiple	1 (0.9)
History of non-melanoma skin cancer, n (%)	Present	19 (16.5)
	Absent	96 (83.5)
Concurrent non-melanoma skin cancer at diagnosis, n (%)	Present	20 (17.4)
	Absent	95 (82.6)
Actinic keratosis at diagnosis, n (%)	Present	25 (21.7)
	Absent	90 (78.3)

Table 1. Continued

Characteristic		Total (n=115)
Clinical variants, n (%)	Classic	109 (94.8)
	Giant KA	6 (5.2)
	KA centrifugum marginatum	0 (0)
	Subungual KA	0 (0)
	Mucosal KA	0 (0)
Recurrence, n (%)	Yes	9 (9.5)
	No	86 (90.5)
	Unknown	20 (17.4)
Time to recurrence, months, median (IQR)**		31 (20-63)
Site of recurrence, (n=9), n (%)	Same site	4 (44.4)
	Different site	5 (55.6)

*: Normally distributed continuous variables are presented as mean ± standard deviation, **: Non-normally distributed continuous variables are presented as median (IQR), SD: Standard deviation, n: Number of patients, IQR: Interquartile range

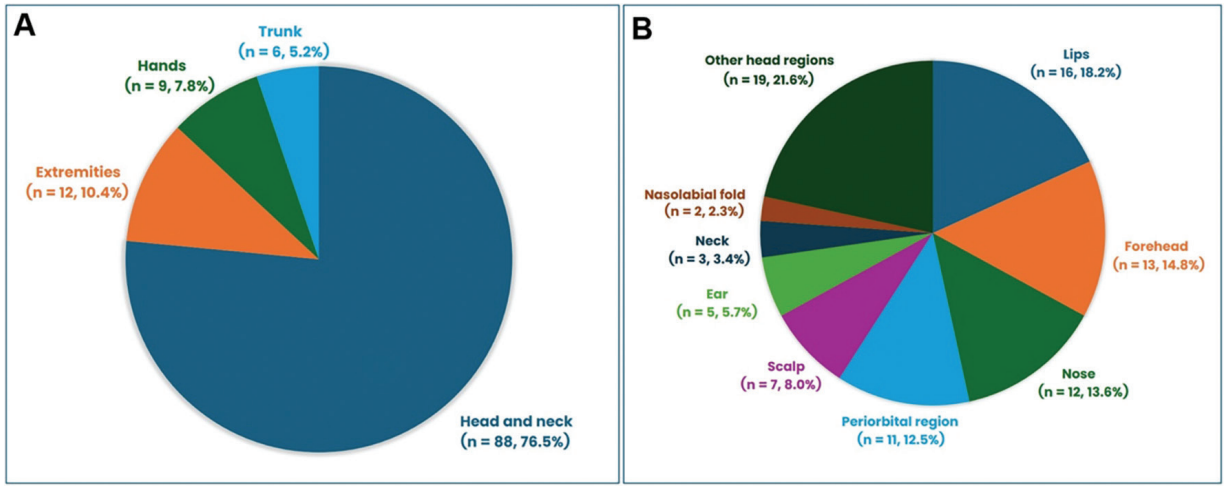

Figure 3. Anatomical distribution of lesion locations in patients with keratoacanthoma (n=115) (A). Distribution of keratoacanthoma lesions within the head and neck region (n=88) (B)

Table 2. Histopathological characteristics of keratoacanthoma (KA) lesions

Characteristic	Total (n=115)
KA with malignant transformation, n (%)	42 (36.5)
Benign KA, n (%)	73 (63.5)
Maximum tumor diameter (mm), median (IQR)*	6 (5)
Depth of invasion (mm), median (IQR)*	2 (2)
Ulceration, n (%)	28 (24.3)
Perineural invasion, n (%)	0 (0)
Lymphovascular invasion, n (%)	0 (0)
Lymphohistiocytic infiltrate, n (%)	73 (63.5)

*: Non-normally distributed continuous variables are presented as median (IQR), n: Number of patients, IQR: interquartile range

J Acad Res Med 2026;16(2):79-85

Table 3. Comparison of clinical and histopathologic characteristics between benign keratoacanthoma (KA) and KA with malignant transformation

Characteristic		Benign KA (n=73)	KA with malignant transformation (n=42)	p-value*
Age at diagnosis, years, mean ± SD		63.12±14.94	68.52±14.20	0.060
Sex, n (%)	Female	34 (46.6)	16 (38.1)	0.377
	Male	39 (53.4)	26 (61.9)	
Family history, n (%)		5 (6.8)	2 (4.8)	0.175
Immunosuppression, n (%)		11 (15.1)	10 (23.8)	0.243
Tumor location, n (%)	Trunk	4 (5.5)	2 (4.8)	0.870
	Extremities	8 (11.0)	4 (9.5)	0.790
	Head and neck	58 (79.5)	30 (71.4)	0.328
	Hands	3 (4.1)	6 (14.3)	0.061
History of non-melanoma skin cancer, n (%)		12 (16.4)	7 (16.7)	0.802
Concurrent non-melanoma skin cancer at diagnosis, n (%)		15 (20.5)	5 (11.9)	0.239
Actinic keratosis at diagnosis, n (%)		19 (26.0)	6 (14.2)	0.231
Maximum tumor diameter (mm), median (IQR)**		6 (5.5)	8 (6)	**0.012**
Depth of invasion, (mm), median (IQR)**		2 (1.5)	2 (1.5)	**0.040**
Ulceration, n (%)		15 (20.5)	13 (31.0)	0.211
Lymphohistiocytic infiltrate, n (%)		42 (57.5)	31 (73.8)	0.081

*: Continuous variables were compared using the independent samples t-test or Mann-Whitney U test, as appropriate. Categorical variables were compared using Pearson's chi-square test or Fisher's exact test, **: Non-normally distributed continuous variables are presented as median (IQR), n: Number of patients, IQR: Interquartile range, SD: Standard deviation

Table 4. Factors associated with malignant transformation of keratoacanthoma: univariate and multivariate logistic regression analyses*

Characteristic	Univariate		Multivariate	
	OR (95% CI)	p-value	OR (95% CI)	p-value
Largest diameter of the tumor (mm)	1.092 (1.012-1.179)	0.023	1.072 (0.992-1.158)	0.078
Depth of tumor invasion (mm)	1.345 (1.006-1.799)	0.046	1.208 (0.888-1.643)	0.229

*: Univariate and multivariate binary logistic regression analyses were performed, and the results are presented as odds ratios (ORs) with 95% confidence intervals (CIs)

CI: 0.992-1.158, p=0.078), while depth of tumor invasion was no longer significant (OR: 1.208, 95% CI: 0.888-1.643, p=0.229) (Table 4).

DISCUSSION

The incidence of KA peaks between 50 and 69 years of age, with a higher prevalence in males (4,10). Consistent with the literature, our study demonstrated a predominance of middle-to- older age and male sex among patients with KA.

Dufresne et al. (11) reported a higher incidence of KA during the spring and summer months, attributing this pattern to ultraviolet (UV) exposure. In contrast, our study identified the highest proportion of diagnoses in winter. This discrepancy may be explained by either (i) true geographic or environmental variation in seasonal patterns or (ii) a temporal lag between lesion onset and clinical diagnosis. Although lesions may develop during periods of increased UV exposure, the rapid growth of KA over weeks to months—potentially combined with delayed patient presentation—may shift the timing of diagnosis to winter (10). Future studies comparing the season of diagnosis with the estimated onset of lesions may further clarify this association.

The etiology of KA remains incompletely understood. It is believed to originate from the hair follicle and most commonly arises on hair-bearing, chronically sun-exposed skin in older adults. Proposed etiological factors include UV radiation, chemical carcinogens (e.g., tar), immunosuppression, and foreign materials such as tattoos or collagen fillers. KA has also been associated with genetic syndromes (including Muir-Torre syndrome and xeroderma pigmentosum), trauma (e.g., surgery, laser therapy, or cryotherapy), and certain medications, including BRAF inhibitors (vemurafenib, dabrafenib) and immunosuppressive disease-

modifying antirheumatic drugs such as leflunomide (1,4). These risk factors substantially overlap with those reported for cSCC (12). This shared risk profile, together with the frequent involvement of sun-exposed body sites of the head and neck, may complicate clinical differentiation between KA and cSCC. In our study, KA lesions were most frequently located on the head, particularly the lips, consistent with the findings of Vîlcea et al. (9). Several studies have identified the lower lip or vermilion border as a common site for cSCC (13). This shared anatomic predilection further complicates clinical differentiation between KA and cSCC.

Solitary KA is the most common clinical variant and typically presents as a well-circumscribed, firm, pink nodule measuring approximately 1-2 cm in diameter, with a characteristic central keratotic plug (14). Giant KA is a rare variant that may exceed 2 cm in diameter and, in some cases, fail to undergo spontaneous regression. Histopathologically, giant KA shares similar features with classical solitary KA (15). Due to their size and anatomical location, giant KAs may be associated with greater functional and cosmetic morbidity (16). Consistent with previous reports, solitary KA was the predominant variant in our study. It is possible that some giant KA cases represent initially solitary lesions that enlarged due to delayed presentation or diagnosis. These findings highlight the importance of improving awareness among clinicians and patients to promote early presentation, timely diagnosis, and appropriate management.

KA is associated with the presence or subsequent development of other NMSCs and premalignant lesions, such as actinic keratosis, and is considered a clinically relevant marker of increased cutaneous cancer risk (9). In a large population-based study, 89% of individuals with KA had a prior history of treatment for at least one skin cancer or actinic keratosis (17). In our study, a prior history of NMSC was documented in 16% of cases, while concomitant NMSC at diagnosis was observed in 21.7%. The lower rates observed in our cohort may partly reflect incomplete documentation inherent to retrospective study designs. Nonetheless, careful screening for coexisting NMSCs and premalignant lesions in patients with KA remains clinically warranted.

Distinguishing benign KA from KA with malignant transformation may be challenging because of substantial clinical and histopathological overlap. Features suggestive of malignancy include metastatic potential, aggressive clinical behavior (such as local recurrence or destructive growth), marked cytological atypia, vascular and/or perineural invasion, an infiltrative growth pattern, and an increased proliferative index (14). Vîlcea et al. (9) reported a malignant transformation rate of 73.71% and concluded that clinical features alone were insufficient predictors of transformation. The lower rate of malignant transformation (36.5%) observed in our series may be attributed to variations in sample characteristics, case selection criteria, or differences in diagnostic thresholds. Collectively, these findings support the notion that clinical characteristics alone may be insufficient for definitive differentiation; rather, tumor size and depth of invasion appear to be more discriminative parameters.

The literature indicates that KA cases with malignant transformation are more likely to occur in older individuals (9,18). In line with these reports, malignant transformation was also observed more frequently in older patients in our study. One possible explanation is that age-related immunosenescence and cumulative UV exposure may contribute to a cutaneous microenvironment that facilitates malignant transformation (19).

The association between malignant transformation in KA and tumor location remains uncertain. One study reported that most KAs with malignant transformation were located on sun-exposed body sites (18). In line with the study by Vîlcea et al. (9), tumor location was not significantly associated with malignant transformation in the present study. Further studies are needed to clarify this relationship.

The natural course of KA involves sequential proliferation, stabilization and regression, with spontaneous involution attributed to increased apoptosis, Wnt-retinoic acid pathway crosstalk, and heightened antitumour immune activity; when these regulatory pathways are inadequate, lesions may fail to regress and instead resemble the biology of cSCC (3). Previous studies have reported lower rates of spontaneous regression in KA cases with malignant transformation (20). This may be associated with the larger tumor diameter and greater depth of tumor invasion observed in KAs with malignant transformation, as also demonstrated in our study. However, the available evidence on this issue remains limited. Our findings may therefore provide useful guidance for clinicians in this regard.

A review examining the histopathological features of KA and cSCC noted that ulceration and perineural invasion occurred more frequently in cSCC, suggesting that these findings may reflect more aggressive biological behavior (4). Another study reported lymphohistiocytic infiltration in all KA cases, whereas no instances of vascular or perineural invasion were observed. (9). The absence of lymphovascular and perineural invasion may be considered a supportive histopathological finding favoring KA in the appropriate clinicopathological context; however, this finding alone is not sufficient to reliably distinguish KA from cSCC, particularly in the absence of a direct cSCC control group.

Complete surgical excision remains the first-line treatment and the gold standard for KA, as it allows definitive histopathological assessment and is associated with favorable outcomes. Recurrence has been reported in approximately 8% of cases following surgical treatment (1,21). Our recurrence rate of 9.5% was somewhat higher than previously reported rates, which may be partly attributable to differences in follow-up duration, patient characteristics, or treatment approaches.

In our study, univariate analysis identified the largest tumor diameter and depth of tumor invasion as predictive factors for malignant transformation in KA. However, these associations were not confirmed in the multivariate analysis. This may largely be explained by the limited sample size and by the biological and statistical overlap between clinicopathological parameters,

J Acad Res Med 2026;16(2):79-85

particularly tumor diameter and depth of invasion. Therefore, the independent prognostic significance of these variables in malignant transformation should be validated in larger prospective studies.

Study Limitations

This study has several limitations. First, its retrospective design may introduce inherent bias. Second, patients' sociodemographic characteristics could not be fully analyzed due to missing or incomplete medical records. Third, the assessment of recurrence may have been limited by the loss of 20 patients to follow-up and the relatively short mean follow-up duration of 31 months, which may have precluded a comprehensive evaluation of long-term recurrence. Fourth, in the comparison between benign KA and KA with malignant transformation, only the largest tumor diameter and depth of tumor invasion were statistically significant. Therefore, only these two variables were included in the multivariable logistic regression analysis. The limited number of events may have affected the reliability of the multivariable analysis.

CONCLUSION

In conclusion, our findings suggest that KA follows a clinical and demographic pattern similar to cSCC, predominantly affecting elderly males on sun-exposed body sites. Histopathologically, the absence of lymphovascular and perineural invasion may support the diagnosis of KA. Moreover, in cases exhibiting malignant transformation, tumor diameter and depth of invasion appear to be more discriminative parameters than clinical or sociodemographic variables. Histopathological examination remains the most reliable and essential method for diagnosing malignant transformation in KA.

Ethics

Ethics Committee Approval: Ethical approval for this study was obtained from the ethics committee of Trakya University Faculty of Medicine (protocol no: TÜTF-GOBAEK 2025/380, decision no: 17/03, date: 22.09.2025).

Informed Consent: Retrospective study.

Footnotes

Author Contributions: Concept - M.Ü., Y.G.Ü.; Design - M.Ü., G.K., Y.G.Ü.; Data Collection and/or Processing - G.K., M.B., N.C.; Analysis and/or Interpretation - M.Ü., Y.G.Ü., N.C.; Literature Search - M.Ü., G.K., M.B., Y.G.Ü., N.C.; Writing - M.Ü., G.K., M.B., Y.G.Ü., N.C.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors report that no financial support was received for this study

REFERENCES

1. Hyeraci M, Didona D, Abeni D, Magri F, Ricci F, Bertagnin C, et al. New insights into pathogenesis and management of keratoacanthoma: a narrative review. Int J Mol Sci. 2025; 26: 10040.
2. Takai T. Advances in histopathological diagnosis of keratoacanthoma. J Dermatol. 2017; 44: 304-14.
3. Tisack A, Fotouhi A, Fidai C, Friedman BJ, Ozog D, Veenstra J. A clinical and biological review of keratoacanthoma. Br J Dermatol. 2021; 185: 487-98.
4. Kwiek B, Schwartz RA. Keratoacanthoma (KA): an update and review. J Am Acad Dermatol. 2016; 74: 1220-33.
5. Elder DE, Massi D, Scolyer RA, Willemze R. WHO classification of skin tumours. 4th ed. Lyon: IARC Press; 2018.
6. Goldman-Lévy G, Barnhill R, Bastian BC, Kempf W, Elder D, Gerami P, et al. WHO classification of skin tumours: key updates in the fifth edition. Histopathology. 2026; 88: 555-68.
7. Dominiak N, Hayes B, Swick J, Leach B, Maize J Sr, Ralston J. Keratoacanthoma centrifugum marginatum: a diagnostic and therapeutic challenge. JAAD Case Rep. 2016; 2: 206-8.
8. Alghamdi NA, Taj MA, Alamri SA, Taj JA, Alshowaikhat SJ, Jaber T, et al. Solitary keratoacanthoma of the scalp: a case report. Cureus. 2023; 15: e51176.
9. Vîlcea AM, Stoica LE, Georgescu CV, Popescu FC, Ciurea RN, Vîlcea ID, et al. Clinical, histopathological and immunohistochemical study of keratoacanthoma. Rom J Morphol Embryol. 2021; 62: 445-56.
10. Bahmad HF, Stoyanov K, Mendez T, Trinh S, Terp K, Qian L, et al. Keratoacanthoma versus squamous-cell carcinoma: histopathological features and molecular markers. Dermatopathology (Basel). 2024; 11: 272-85.
11. Dufresne RG, Marrero GM, Robinson-Bostom L. Seasonal presentation of keratoacanthomas in Rhode Island. Br J Dermatol. 1997; 136: 227-9.
12. Artosi F, Costanza G, Di Prete M, Garofalo V, Lozzi F, Dika E, et al. Epidemiological and clinical analysis of exposure-related factors in non-melanoma skin cancer: a retrospective cohort study. Environ Res. 2024; 247: 118117.
13. Solak SS, Yöndem H, Urun YG, Cezik M, Can N. High prevalence of high-risk cutaneous squamous cell carcinoma in the thrace region of Turkey. Turk J Dermatol. 2020; 14: 83-9.
14. Ko CJ. Keratoacanthoma: facts and controversies. Clin Dermatol. 2010; 28: 254-61.
15. Bogner PN, Cheney RT, Zeitouni NC. Giant keratoacanthoma: case report and review of the English literature. Am J Dermatopathol. 2014; 36: 252-7.
16. Garcia-Zuazaga J, Ke M, Lee P. Giant keratoacanthoma of the upper extremity treated with mohs micrographic surgery: a case report and review of current treatment modalities. J Clin Aesthet Dermatol. 2009; 2: 22-5.
17. Kolmodin A, Pandeya NP, Olsen CM, Dusingize JC, Whiteman DC, Claeson M. Patient and tumour characteristics of keratoacanthoma in a large, community-based cohort study from Queensland, Australia. Acta Derm Venereol. 2021; 101: adv00469.
18. Sánchez Yus E, Simón P, Requena L, Ambrojo P, de Eusebio E. Solitary keratoacanthoma: a self-healing proliferation that frequently becomes malignant. Am J Dermatopathol. 2000; 22: 305-10.
19. D'Arino A, Caputo S, Eibenschutz L, Piemonte P, Buccini P, Frascione P, et al. Skin cancer microenvironment: what we can learn from skin aging? Int J Mol Sci. 2023; 24: 14043.
20. Saavedra-Portales S, Corradini E, Goldman Y, Pantoja Á. Self-involution of a solitary keratoacanthoma: a case report with photographic documentation. J Cutan Aesthet Surg. 2021; 14: 433-5.
21. Tran DC, Li S, Henry S, Wood DJ, Chang ALS. An 18-year retrospective study on the outcomes of keratoacanthomas with different treatment modalities at a single academic centre. Br J Dermatol. 2017; 177: 1749-51.

DOI: 10.4274/jarem.galenos.2026.36002
J Acad Res Med 2026;16(2):86-90

Evaluation of Surgical Success Rate and Influencing Factors in Patients with Primary Rhegmatogenous Retinal Detachment Treated with Pneumatic Retinopexy

Kezban Bulut, Dilber Çelik Yaprak, Yusuf Can Aydın, Güzide Akçay

University of Health Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital, Department of Ophthalmology, İstanbul, Türkiye

Cite this article as: Bulut K, Çelik Yaprak D, Aydın YC, Akçay G. Evaluation of surgical success rate and influencing factors in patients with primary rhegmatogenous retinal detachment treated with pneumatic retinopexy. J Acad Res Med. 2026;16(2):86-90

ABSTRACT

Objective: This study aims to evaluate the anatomical and visual success rates of primary pneumatic retinopexy (PR) in patients with rhegmatogenous retinal detachment (RRD) and to investigate the factors influencing surgical outcomes.

Methods: A retrospective review was conducted on patients with RRD treated with primary PR at University of Health Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital between May 2011 and May 2024. Demographic data, best-corrected visual acuity (BCVA), lens status, number and location of retinal breaks, macular involvement, gas tamponade type, intraocular pressure, and anatomical success were analyzed. Patients with failed PR underwent pars plana vitrectomy (PPV).

Results: The study included 32 eyes of 32 patients (19 males, 13 females) with a mean age of 53 years. PR achieved anatomical success in 17 patients (53.1%). Success was significantly higher in macula-on cases (85.7%) compared to macula-off cases (27.8%) (p=0.001). Preoperative BCVA improved significantly postoperatively (1.61 to 1.25 logMAR; p=0.003). PPV performed in 15 patients yielded significant visual improvement (2.14 to 0.93 logMAR; p=0.001). Macular involvement was a negative prognostic factor, while tamponade type and lens status were not statistically significant.

Conclusion: PR is a minimally invasive and cost-effective alternative to vitrectomy in selected RRD cases. PR should be considered for patients with superior retinal breaks and no macular involvement.

Keywords: Rhegmatogenous retinal detachment, pneumatic retinopexy, retinal surgery

INTRODUCTION

Retinal detachment (RD) is one of the most significant causes of vision loss among the working-age population. Rhegmatogenous RD (RRD) occurs as a result of a full-thickness retinal break, allowing liquefied vitreous to enter the subretinal space and separate the neurosensory retina from the underlying retinal pigment epithelium (1). Its annual incidence is approximately 1 in 10,000, and nearly 50% of cases develop spontaneously (2). Contemporary management of RRD primarily relies on three established surgical approaches: scleral buckling (SB), pars plana vitrectomy (PPV), and pneumatic retinopexy (PR) (3).

PR was first described in the mid-1980s by Dominguez (4), Hilton and Grizzard (5) as a treatment modality for RRD. PR is based on the principle of intravitreal expansile gas injection combined with postoperative positioning to achieve temporary retinal break tamponade, followed by definitive retinopexy. The most important advantages of PR include its minimally invasive nature, the absence of a requirement for sedation or general anesthesia, its applicability under outpatient clinic conditions in a short time, and a lower risk of surgical complications. In addition, compared to SB and PPV, PR is preferred due to its faster visual rehabilitation, the ability to be performed without extensive preparation, and its cost-effectiveness (6). However, despite all these advantages, uncertainties regarding the efficacy of PR have caused the technique to remain controversial.

In this study, we aimed to evaluate the functional (visual) and anatomical success rates of patients who underwent primary PR for the treatment of RRD, as well as to investigate the factors affecting surgical success.

METHODS

A retrospective review was conducted of patients who were diagnosed with RRD at the Eye Clinic of University of Health Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital between May 2011 and May 2024, had superior retinal breaks, and underwent primary PR.

ORCID IDs of the authors: K.B. 0000-0002-8315-4923; D.Ç.Y. 0000-0002-2144-3340; Y.C.A. 0009-0006-1518-2320; G.A. 0000-0002-6111-9772

Corresponding Author: Kezban Bulut, MD;
E-mail: yepelekk@hotmail.com

Received Date: 15.12.2025 **Accepted Date:** 16.06.2026
Epub: 07.07.2026
Publication Date: 07.08.2026

J Acad Res Med 2026;16(2):86-90

Patients who were followed for at least 6 months were included in the study. Preoperative and postoperative best-corrected visual acuity (BCVA), intraocular pressure (IOP), and detailed anterior and posterior segment examinations were evaluated in all included patients. For statistical purposes, all BCVA values were standardized and expressed in logMAR units. Standard logMAR conversion values were used for low visual acuity measurements (e.g., counting fingers, hand movements, light perception), such as 2.0 logMAR for counting fingers and 2.3 logMAR for hand movements. In addition, lens status, number and localization of retinal breaks, presence of macular involvement, type of gas tamponade used [sulfur hexafluoride (SF_6) or perfluoropropane (C_3F_8)], and anatomical success rates were analyzed.

In cases in which primary PR failed, PPV combined with phacoemulsification and intraocular lens implantation was performed. In these patients, preoperative and postoperative visual acuity as well as the development of proliferative vitreoretinopathy (PVR) were also evaluated.

The inclusion criteria were RRD cases with retinal breaks located between the 8 and 4 o'clock positions. Exclusion criteria included inferior quadrant retinal breaks, retinal breaks larger than one clock hour, cases in which retinal breaks could not be detected by fundus examination or could not be visualized due to media opacities (such as cataract or vitreous hemorrhage), presence of PVR, history of previous retinal surgery, and inability to comply with the recommended postoperative positioning.

Ethical Approval

This study was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained from the Local Institutional Review Board of University of Health Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital (approval number: 2025/010.99/17/28, date: 25.06.2025). Due to the retrospective nature of the study, the requirement for informed consent was waived.

Surgical Technique and Postoperative Management

All procedures were performed according to a standardized protocol by one of two experienced vitreoretinal surgeons (M.N.B. or G.A.). After all retinal breaks were identified by fundus examination, patients were informed in detail about the procedure, alternative treatment options, and possible complications, and written informed consent was obtained from all participants. Written informed consent for the surgical intervention was obtained from all patients prior to surgery.

All surgeries were performed under topical anesthesia in an outpatient setting. Intravitreal gas injection was administered through the pars plana at a distance of 3.5 mm in pseudophakic eyes and 4 mm in phakic eyes using a 26-gauge needle, delivering either 0.5-0.6 mL of pure SF_6 or 0.3 mL of pure C_3F_8 to create a single gas bubble. Following the injection, anterior chamber paracentesis was performed to prevent IOP elevation, and light perception was checked immediately after the procedure.

All patients were instructed to maintain appropriate postoperative head positioning to ensure that the gas bubble effectively tamponaded the retinal break. Within the first three days after gas injection, 2-3 rows of laser photocoagulation were applied around the retinal tear. Postoperative topical antibiotic and corticosteroid eye drops were prescribed routinely.

Patients were followed up at postoperative 1 week, and at 1, 3, and 6 months, and subsequently at 6-month intervals. At each visit, BCVA, IOP, and detailed anterior and posterior segment examinations were performed.

Complications and Secondary Surgical Protocol

Postoperative complications such as elevated IOP, cataract progression, vitreous hemorrhage, re-detachment, and the development of PVR were carefully recorded throughout the follow-up period.

In cases in which primary PR failed to achieve anatomical retinal reattachment, secondary surgical intervention with PPV combined with phacoemulsification and intraocular lens implantation was performed when indicated. During secondary surgery, standard three-port PPV was applied, subretinal fluid was drained when necessary, endolaser photocoagulation was applied around the retinal breaks, and long-acting gas tamponade was used. The development of postoperative PVR and final anatomical and functional outcomes were evaluated in these patients.

Statistical Analysis

Statistical analysis was performed using SPSS software. Descriptive statistics were presented as mean ± standard deviation or median (interquartile range) for continuous variables, and as frequencies (n) and percentages (%) for categorical variables. Visual acuity measurements were converted to logMAR values for statistical comparisons. For patients with very low vision, standard logMAR conversion values were assigned: 2.0 logMAR for counting fingers and 2.3 logMAR for hand movements.

The normality of the data distribution was tested using the Shapiro-Wilk test. Preoperative and postoperative functional (BCVA) and physiological (IOP) parameters were compared using the Wilcoxon signed-rank test (for non-normally distributed data). Categorical variables were compared using the chi-square test or Fisher's exact test. A p-value of <0.05 was considered statistically significant.

RESULTS

Thirty-two eyes of 32 patients (19 male, 13 female) were included in the study. The mean age of the patients was 51.41±14.27. While macular involvement was not observed in eighteen patients, RD with macular involvement was present in 14 patients. The mean number of tears was determined to be 1.3. (Table 1).

SF_6 was used as an gas tamponade in 14 patients, and C_3F_8 was used in 18 patients. While anatomical success was achieved in 17 patients (53.1%) after PR, retinal reattachment could not be achieved in 15 patients (46.9%), and 100% final anatomical success was achieved in all patients with secondary surgery.

J Acad Res Med 2026;16(2):86-90

While the preoperative median BCVA was 2.3 logMAR, the postoperative BCVA was 1.00 logMAR. This change was statistically significant (p=0.003). In patients where anatomical success could not be achieved with PR and who subsequently underwent PPV, the preoperative BCVA was measured as 2.3 logMAR, and the postoperative BCVA as 0.8 logMAR; this difference was statistically significant (p=0.001). Overall, a statistically significant improvement in BCVA was observed across the entire cohort (p=0.003), with a more pronounced gain in patients requiring secondary PPV (p<0.01) (Table 2).

Among the 18 patients with macular involvement, surgical failure occurred in 13, whereas failure was observed in only 2 of the 14 patients without macular involvement (fovea-on). The anatomical success rate of primary PR was 85.7% in macula-on RRD compared with 27.8% in macula-off RRD. Macular involvement was identified as a statistically significant risk factor for surgical failure (p=0.001).

With respect to the type of tamponade, surgical failure occurred in 6 of the 18 patients treated with C_3F_8 and in 9 of the 14 patients treated with SF_6; however, the difference between the two groups was not statistically significant (p=0.08). Similarly, according to lens status, surgical failure developed in 9 of the 18 phakic patients and in 6 of the 14 pseudophakic patients. This difference was also not found to be statistically significant (p=0.68) (Table 3).

Table 1. Demographic and clinical characteristics of the patients

Parameter	Value
Age (years), mean ± SD	51.41±14.27
Laterality (right/left), n (%)	17 (53.1)/15 (46.9)
Sex (male/female), n (%)	19 (59.4)/13 (40.6)
Macular status (on/off), n (%)	14 (43.8)/18 (56.3)
Number of retinal breaks, mean (min-max)	1.3 (1-3)
Gas type (SF_6/C_3F_8), n (%)	14 (43.8)/18 (56.3)
Lens status (phakic/pseudophakic), n (%)	18 (56.3)/14 (43.8)
Retinal status after PR (attached/detached), n (%)	17 (53.1)/15 (46.9)

PR: Pneumatic retinopexy, SF_6: Sulfur hexafluoride, C_3F_8: Perfluoropropane, SD: Standard deviation

Table 2. Comparison of preoperative and postoperative clinical parameters

Parameter	Median (IQR)	p-value
Preoperative BCVA (logMAR)	2.3 (0.1-2.7)	0.003
Postoperative PR BCVA (logMAR)	1.0 (0.1-2.7)	
Preoperative IOP (mmHg)	14 (11-18)	0.165
Postoperative IOP (mmHg)	15 (12-17)	
Preoperative PPV BCVA (logMAR)	2.3 (0.3-2.7)	0.001
Postoperative PPV BCVA (logMAR)	0.8 (0.2-1.3)	

Wilcoxon signed rank-test
BCVA: Best-corrected visual acuity, IOP: Intraocular pressure, PR: Pneumatic retinopexy, PPV: Pars plana vitrectomy, IQR: Interquartile range

The median preoperative IOP was 14 mmHg, while the median postoperative IOP was 15 mmHg. No clinically or statistically significant IOP elevation was observed following PR, and no cases of secondary glaucoma developed during follow-up.

DISCUSSION

PR represents a minimally invasive and cost-effective alternative to vitrectomy in the management of selected cases of RRD. In the present retrospective study, primary PR achieved an anatomical success rate of 53.1%. Although this rate appears moderate, it falls within the broad range of outcomes reported in the literature, reported single-procedure success rates for PR show considerable variability across published series (7). Importantly, all cases with primary PR failure achieved complete retinal reattachment following secondary PPV combined with phacoemulsification and intraocular lens implantation, yielding a final anatomical success rate of 100%. These findings support the concept that PR can be safely positioned as an initial, less invasive step within a stepwise surgical strategy for appropriately selected patients.

In our study, patient selection was based on the PIVOT eligibility criteria, which are also applied in our routine clinical practice. One potential explanation for the relatively lower primary success rate observed is that PR was performed only once in all patients, and repeat PR was not attempted in cases of primary failure. Instead, these patients were directly referred for PPV. This pragmatic approach, while reducing cumulative procedural burden, may have contributed to the lower single-procedure success rate compared with series allowing repeated PR attempts.

Chandelier-assisted PR, as described by Habib et al. (8), has been reported to significantly enhance single-procedure success by facilitating improved visualization of peripheral retinal breaks, achieving success rates of 91.7% after one procedure and 100% after a second intervention. The authors emphasized the ability of this technique to reduce the likelihood of missed small peripheral breaks and to minimize intraoperative complications such as iatrogenic lens injury. In our cohort, missed small peripheral retinal tears likely contributed to primary surgical failure, particularly in eyes with extensive RD. Consistently, we observed that primary

Table 3. Surgical success rates among patient subgroups

Variable	Retina detached, n	Retina attached, n	p-value
Macular status			0.001[a]
Macula-on	2	12	
Macula-off	13	5	
Gas type			0.082[b]
C_3F_8	6	12	
SF_6	9	5	
Lens status			0.688[b]
Phakic	9	9	
Pseudophakic	6	8	

[a]: Chi-square test, [b]: Fisher's exact test, C_3F_8: Perfluoropropane, SF_6: Sulfur hexafluoride

J Acad Res Med 2026;16(2):86-90

PR failure occurred more frequently in cases with widespread detachment and foveal involvement.

Lens status did not significantly influence anatomical success in our cohort. While earlier studies reported higher PR success rates in phakic eyes compared with pseudophakic or aphakic eyes (9), our results are consistent with the observations of Fabian et al. (10), who also reported comparable outcomes between these groups. The prevailing hypothesis underlying reduced success in pseudophakic and aphakic eyes is the higher probability of multiple missed retinal breaks. Accordingly, meticulous peripheral retinal examination remains essential when considering PR in pseudophakic RRD.

Macular status emerged as the most significant predictor of surgical outcome in our series. While an anatomical success rate of 85.7% was achieved in macula-on RRD, this rate decreased dramatically to 27.8% in macula-off cases (p=0.001). In our cohort, macular status emerged as the strongest predictor of anatomical outcome following PR. These findings are in strong agreement with previous reports (11). Grizzard et al. (12) previously demonstrated that increasing extent of detachment is associated with poorer PR outcomes. In extensive RRD, visualization of small and peripheral breaks becomes increasingly challenging due to media disturbance caused by subretinal fluid accumulation, thereby predisposing these eyes to residual undetected breaks and subsequent surgical failure. Our findings further support this mechanistic explanation.

Similarly, the type of gas tamponade (C_3F_8 vs. SF_6) was not significantly associated with surgical success in our analysis (p=0.08), in agreement with the findings reported by İpekli et al. (13). These authors suggested that injected gas volume, rather than gas type itself, may be the principal determinant of tamponade efficacy. Unfortunately, due to the retrospective nature of our data and limited sample size, quantitative analysis of gas volume as a prognostic variable could not be performed.

From a functional standpoint, PR resulted in a statistically significant improvement in BCVA (p=0.003), underscoring the functional benefit of this minimally invasive technique and its ability to facilitate rapid visual rehabilitation. Moreover, patients who required secondary PPV after PR failure also demonstrated a marked improvement in BCVA (p=0.001), confirming the effectiveness of PPV as a highly reliable second-line intervention. These findings are consistent with prior reports and with the PIVOT study, which similarly demonstrated favorable visual outcomes following PPV after primary PR failure (11,14).

Study Limitations

The principal limitation of this study lies in its retrospective design. Additional limitations include the relatively small sample size and the short follow-up duration of six months. Nevertheless, data from the PIVOT trial have demonstrated that eyes achieving stable retinal reattachment within the first three months exhibit a negligible risk of late redetachment. Future prospective, multicenter studies with longer follow-up periods are warranted to further refine real-world estimates of primary and final reattachment rates, patient-reported visual outcomes, and late recurrence patterns under standardized PIVOT-guided surgical protocols. Additionally, the retrospective design precluded standardized assessment of patient-reported outcomes and precise quantification of injected gas volume.

CONCLUSION

In conclusion, optimal patient selection, strict postoperative positioning, close follow-up, and high patient compliance are critical determinants of PR success. Based on our findings, PR should primarily be considered in patients with macula-on RRD and superiorly located retinal breaks, in whom it offers an effective, minimally invasive, and vision-preserving first-line treatment option. When applied in carefully selected macula-on RRD cases with superior breaks, PR represents an effective first-line strategy that preserves the option of highly successful secondary surgery without compromising final anatomical or functional outcomes.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the local Institutional Review Board of University of Health Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital (approval number: 2025/010.99/17/28, date: 25.06.2025).

Informed Consent: Due to the retrospective nature of the study, the requirement for informed consent was waived.

Footnotes

Author Contributions: Surgical and Medical Practices - K.B., G.A.; Concept - K.B., D.Ç.Y., Y.C.A., G.A.; Design - K.B., D.Ç.Y., Y.C.A., G.A.; Data Collection and/or Processing - K.B., D.Ç.Y., Y.C.A., G.A.; Analysis and/or Interpretation - K.B., D.Ç.Y., Y.C.A., G.A.; Literature Search - K.B., D.Ç.Y., Y.C.A., G.A.; Writing - K.B., D.Ç.Y., Y.C.A., G.A.

Conflict of Interest: The authors have no conflict of interest to declare.

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

REFERENCES

1. Tareen S, Tahir MA, Cheema AM. Surgical audit of outcome of rhegmatogenous retinal detachment repair at Vitreoretinal unit JPMC in year 2014. Pak J Med Sci. 2016; 32: 101-5.
2. Mikhail MA, Mangioris G, Casalino G, McGimpsey S, Sharkey J, Best R, et al. Outcome of primary rhegmatogenous retinal detachment surgery in a tertiary referral centre in Northern Ireland - a regional study. Ulster Med J. 2017; 86: 15-9. Erratum in: Ulster Med J. 2017; 86: 126.
3. Znaor L, Medic A, Binder S, Vucinovic A, Marin Lovric J, Puljak L. Pars plana vitrectomy versus scleral buckling for repairing simple rhegmatogenous retinal detachments. Cochrane Database Syst Rev. 2019; 3: CD009562.
4. Dominguez DA. Cirugia precoz y ambulatoria del desprendimiento de retina. Arch Soc Esp Oftalmol. 1985; 48: 47-54.
5. Hilton GF, Grizzard WS. Pneumatic retinopexy. A two-step outpatient operation without conjunctival incision. Ophthalmology. 1986; 93: 626-41.
6. Figueiredo N, Warder DC, Muni RH, Lee WW, Yong SO, Kertes PJ. Pneumatic retinopexy as a treatment for rhegmatogenous retinal detachment in pediatric patients meeting PIVOT criteria. Can J Ophthalmol. 2022; 57: 359-63.
7. Roshanshad A, Shirzadi S, Binder S, Arevalo JF. Pneumatic retinopexy versus pars plana vitrectomy for the management of retinal detachment: a systematic review and meta-analysis. Ophthalmol Ther. 2023; 12: 705-19.

J Acad Res Med 2026;16(2):86-90

8. Habib AE, Abdel-Kader AA, Elnahry AG. Chandelier-assisted pneumatic retinopexy for rhegmatogenous retinal detachment repair in young adults. Indian J Ophthalmol. 2021; 69: 979-81.
9. Chan CK, Lin SG, Nuthi AS, Salib DM. Pneumatic retinopexy for the repair of retinal detachments: a comprehensive review (1986-2007). Surv Ophthalmol. 2008; 53: 443-78.
10. Fabian ID, Kinori M, Efrati M, Alhalel A, Desatnik H, Hai OV, et al. Pneumatic retinopexy for the repair of primary rhegmatogenous retinal detachment: a 10-year retrospective analysis. JAMA Ophthalmol. 2013; 131: 166-71.
11. Iannetta D, Valsecchi N, Finzi A, Mastropasqua R, Muni RH, Fontana L. Pneumatic retinopexy for primary rhegmatogenous retinal detachment: from a clinical trial to the real-life experience. BMC Ophthalmol. 2024; 24: 287.
12. Grizzard WS, Hilton GF, Hammer ME, Taren D, Brinton DA. Pneumatic retinopexy failures. Cause, prevention, timing, and management. Ophthalmology. 1995; 102: 929-36.
13. İpekli Z, Pehlivanoğlu S, Artunay Ö. Efficacy of pneumatic retinopexy in young adults with rhegmatogenous retinal detachment. Ther Adv Ophthalmol. 2023; 15: 25158414231208279.
14. Hillier RJ, Felfeli T, Berger AR, Wong DT, Altomare F, Dai D, et al. The pneumatic retinopexy versus vitrectomy for the management of primary rhegmatogenous retinal detachment outcomes randomized trial (PIVOT). Ophthalmology. 2019; 126: 531-9.

DOI: 10.4274/jarem.galenos.2026.94546
J Acad Res Med 2026;16(2):91-97

Validity and Reliability of the Turkish Version of the Screening for Anxiety and Depression in Tinnitus-Hyperacusis-Misophonia

Nizamettin Burak Avcı, İlknur Taşdemir Gürşen

Trakya University Faculty of Health Sciences, Department of Audiology, Edirne, Türkiye

Cite this article as: Avcı NB, Taşdemir Gürşen İ. Validity and reliability of the Turkish version of the screening for anxiety and depression in tinnitus-hyperacusis-misophonia. J Acad Res Med. 2026;16(2):91-97

ABSTRACT

Objective: The purpose of this study was to assess the validity and reliability of the Turkish version of the screening for anxiety and depression in tinnitus-hyperacusis-misophonia (SAD-T) in individuals with tinnitus.

Methods: The SAD-T has four items and was translated into Turkish using the "back-translation" method. The study included 105 individuals (33 M, 72 F) with subjective tinnitus whose mean age was 31.87±15.16 (min: 20, max: 65). Cronbach's α was used to measure internal consistency, and the intraclass correlation coefficient (ICC) was used to measure test-retest reliability in a subgroup of thirty participants, based on repeated assessments over two weeks. Confirmatory factor analysis was used to assess construct validity. The depression anxiety stress scale-21 (DASS-21) was used to evaluate the concurrent validity of the Turkish version of SAD-T (SAD-T-TR). Additionally, the DASS-21 subscale scores were used to categorize participants into anxiety-positive and anxiety-negative, and depression-positive and depression-negative groups to assess discriminant validity.

Results: Every item had a corrected item-total correlation coefficient higher than 0.30. The scale demonstrated excellent test-retest reliability (ICC=0.945) and good internal consistency (Cronbach's α=0.879). The model demonstrated an acceptable fit [χ^2 (1)=3.046, p=0.081; comparative fit index=0.991; Tucker-Lewis index=0.945; root mean square error of approximation (RMSEA)=0.140], supporting a unidimensional structure despite an inflated RMSEA due to low degrees of freedom (df=1). SAD-T-TR scores were moderately positively correlated with DASS-21 scores (r=0.680 for total, r=0.670 for depression, r=0.586 for anxiety, r=0.639 for stress, p<0.001). SAD-T-TR scores were significantly higher in participants with anxiety and depression than in those without these conditions (p<0.001).

Conclusion: SAD-T-TR is a valid and reliable tool for screening depression and anxiety in individuals with tinnitus. The scale may be a helpful tool for clinical practice and research in tinnitus populations because it is brief and simple to administer.

Keywords: Tinnitus, anxiety, depression, surveys and questionnaires, self-report, psychometrics

INTRODUCTION

Tinnitus is a common auditory symptom characterized by the perception of sound without an external sound source (1). Tinnitus is considered a complex disorder that affects individuals' cognitive, emotional, and psychosocial functioning, and is not solely an auditory disorder. In individuals with tinnitus, the persistence and perceived severity of the symptom are associated with limitations in daily living activities, sleep problems, and decreased quality of life, which can increase the risk of developing comorbid psychological diagnoses (2).

The literature shows that the prevalence of anxiety and depression is significantly higher among individuals with tinnitus than among those without tinnitus. A systematic review reported an average depression prevalence of 33% in patients with tinnitus, varying between 6% and 84% across studies (3). Similarly, Bhatt et al. (4) found that approximately 26% of adults with tinnitus experienced anxiety and 25.6% experienced depression; these rates remained at approximately 9% in individuals without tinnitus, and the likelihood of these psychological problems increased with the severity of tinnitus. These findings demonstrate that tinnitus significantly affects not only individuals' auditory perception but also their psychological well-being.

Therefore, it is crucial for clinical practice and research to evaluate comorbid psychiatric problems, including anxiety and depression, in people with tinnitus. Self-report measures significantly contribute to our understanding of the psychosocial effects of tinnitus by directly reflecting the emotional, cognitive, and

ORCID IDs of the authors: N.B.A. 0000-0002-5474-457X; İ.T.G. 0000-0002-8304-6674

Corresponding Author: Nizamettin Burak Avcı, MD;
E-mail: nizamettinburakavci@gmail.com

Received Date: 12.03.2026 **Accepted Date:** 16.07.2026
Epub: 23.07.2026
Publication Date: 07.08.2026

J Acad Res Med 2026;16(2):91-97

behavioral challenges people face daily. Clinicians can use these simple-to-use, time-saving, and cost-effective scales to evaluate the presence, severity, and functional impact of symptoms. They also facilitate early referral and treatment planning.

Due to the limited availability of objective measurement methods for assessing tinnitus, self-report scales are predominantly used in clinical practice and research (5,6). In this context, the most commonly used scales are the tinnitus handicap inventory (THI) (7,8), the tinnitus functional index (9), and the tinnitus impact questionnaire (10), and these scales are used in studies examining the relationships between tinnitus and anxiety and depression.

Researchers often use general mental health scales instead of tinnitus-specific scales in the evaluation of anxiety and depression in people with tinnitus (6,11). The patient health questionnaire-9, Beck depression inventory, hospital anxiety and depression scale - depression subscale, and depression, anxiety and stress scales (DASS-21) - depression subscale are among the most commonly used scales for depression screening. These scales have been shown to exhibit significant correlations with tinnitus severity and perceived level of discomfort (3,12,13). Short and reliable screening scales like the generalized anxiety disorder 7-item and 2-item, the hospital anxiety and DASS-21 - anxiety subscale are frequently used in anxiety assessment (3,13,14).

A novel method of screening for psychiatric disorders in this population is provided by the screening for anxiety and depression in tinnitus-hyperacusis-misophonia (SAD-T) scale, which was created especially for tinnitus, misophonia, and hyperacusis (15). The SAD-T was developed as a condition-specific adaptation of the patient health questionnaire-4 (PHQ-4) to facilitate screening for anxiety and depression symptoms in individuals with tinnitus, hyperacusis, and misophonia. While retaining the brevity of the PHQ-4, it incorporates symptom attribution within the context of these auditory conditions, which may improve its clinical relevance for affected populations. However, studies of validity and reliability are necessary to ensure the valid and reliable use of such scales across diverse languages and cultures. Although the SAD-T was originally designed for tinnitus, hyperacusis, and misophonia populations, the present study focused exclusively on individuals with tinnitus. Therefore, the findings should not be generalized to populations with hyperacusis or misophonia without further validation studies. The current literature indicates a scarcity of studies examining the Turkish psychometric properties of scales utilized for screening anxiety and depression in individuals with tinnitus. Thus, this study aims to assess the Turkish validity and reliability of the SAD-T scale among individuals with tinnitus.

METHODS

This study was designed to be both descriptive and cross-sectional. Between March and November 2025, the study was conducted in the Department of Audiology at Trakya University Faculty of Health Sciences. The Trakya University Non-Interventional Ethics Committee approved it with decision number 05/19 and protocol code 2025/104 (date: 06.01.2025). Written informed consent was obtained from all participants.

Participants

The study included 105 individuals with tinnitus, aged between 20 and 65 years. The inclusion criteria for the participants were: subjective tinnitus for at least 6 months; age older than 18 years; absence of any known neurological disorder; and absence of hearing loss.

All participants underwent hearing evaluations prior to inclusion in the study. Hearing thresholds were measured using an Interacoustics AC40 audiometer. Air- and bone-conduction thresholds were obtained for both ears at 500, 1000, 2000, and 4000 Hz. Participants with a four-frequency pure tone average of 25 dB HL or lower were included in the study.

Participants were classified into subgroups based on their scores on the DASS-21. For discriminant validity analyses, participants were classified as anxiety-positive or anxiety-negative based on the DASS-21 anxiety subscale cut-off scores, and as depression-positive or depression-negative based on the DASS-21 depression subscale cut-off scores.

Materials

The participants underwent hearing evaluations, and those who met the requirements were asked to complete the THI, SAD-T, and DASS-21 questionnaires in person. To evaluate test–retest reliability, 30 participants were randomly selected from the study sample and asked to complete the SAD-T again two weeks after the initial assessment.

The THI was used to assess the impact of tinnitus on an individual's daily life. Turkish validity and reliability were established by Aksoy et al. (8), with a Cronbach's α value of 0.88. The total score ranges from 0 to 100, with higher scores indicating a greater negative impact of tinnitus on the individual.

SAD-T is used to assess symptoms of anxiety and depression that may require referral for psychiatric evaluation. The SAD-T includes four items that match the PHQ-4 (16). The total SAD-T score is obtained by summing all item scores, resulting in a range of 0-12, and the scale demonstrated good internal consistency (Cronbach's α=0.91) (17).

DASS-21 is a 21-item self-report measure designed to assess the severity of general psychological distress and symptoms associated with depression and anxiety. It consists of 3 subscales, each with 7 items. These subscales are depression, anxiety, and stress. Turkish reliability validity was done by Sarıçam (18) and Cronbach's α values for the subscales were found to be between 0.87 and 0.81. Subscale scores represent the total score of the relevant items. All subscales are scored between 0 and 21 points. When no symptoms are present, the cut-off score for the depression subscale is 4, the anxiety subscale is 3, and the stress subscale is 7 (19).

Translation Process

The original English form of the SAD-T scale was adapted into Turkish using the back-and-forth translation method in accordance with the proposed standard adaptation steps (20).

J Acad Res Med 2026;16(2):91-97

In the initial stage, two subject-matter experts who were native Turkish speakers and fluent in English independently translated the scale into Turkish. After comparing the draft texts, the authors reached a single Turkish version by consensus. Two independent translators, who were unfamiliar with the original version of the scale, then retranslated the form into English. The back-translations were compared with the original scale to evaluate conceptual and linguistic equivalence, and the preliminary Turkish version of SAD-T (SAD-T-TR) was created after applying the necessary corrections. The preliminary form was administered to 5 adults with tinnitus, and the comprehensibility and cultural appropriateness of the items were evaluated. The final SAD-T-TR was developed and utilized in validity and reliability assessments following minor language modifications identified in the pilot application.

Statistical Analysis

IBM SPSS Statistics version 27 and AMOS version 23 were used for all statistical analyses. The normality of the data distribution was assessed using the Kolmogorov-Smirnov test and visual inspection of histogram graphs. The results indicated that the data were not normally distributed. Categorical data were reported as percentages and frequencies, whereas continuous variables were reported as median (interquartile range).

The distribution of the SAD-T-TR scores was examined for floor and ceiling effects. It is accepted that the floor and ceiling effects occur when the number of participants with the lowest or highest scores exceeds 15% of the total number of participants. Item analysis was performed on all the scale items, and the item was considered reliable when the item-total correlation coefficient was higher than 0.3 (21).

Analyses of construct, concurrent, and discriminant validity were used to assess the SAD-T-TR. The factor structure of the SAD-T-TR was tested for construct validity using confirmatory factor analysis (CFA). The chi-square statistic (χ^2), chi-square/degrees of freedom ratio (χ^2/df), comparative fit index (CFI), Tucker-Lewis index (TLI), incremental fit index (IFI), normed fit index (NFI), goodness-of-fit index (GFI), root mean square error of approximation (RMSEA), and root mean square residual (RMR) were among the goodness-of-fit indices used to assess model fit. Model fit was assessed using the following criteria: χ^2/df <5, CFI, TLI, IFI, NFI, and GFI ≥0.90 indicating acceptable fit, and RMSEA ≤0.08 indicating acceptable model fit (22,23). The concurrent validity of the SAD-T-TR was evaluated by analyzing Spearman correlation coefficients between SAD-T-TR and DASS-21 scores. The correlation coefficient of concurrent validity was considered a weak correlation under 0.3, a moderate correlation between 0.3 and 0.7 and a high correlation with 0.7 and above (24). Differences in SAD-T-TR scores between participants with and without anxiety and depression symptoms, based on the DASS-21 anxiety and depression subscales, were evaluated using the Mann-Whitney U test to examine discriminant validity. Statistical significance was defined as a p-value of less than 0.05.

Internal consistency and test-retest reliability were used to evaluate the scale. Cronbach's α coefficients were rated excellent when r>0.80 (25). Test-retest reliability was assessed using the intraclass correlation coefficient (ICC). ICC values of 0.8 and above were accepted to be perfectly correlated (26).

RESULTS

A total of 105 individuals with tinnitus (33 M, 72 F) were included in the study. The mean age of the participants was 31.87±15.16 (min: 20, max: 65). According to DASS-A, 47 individuals had anxiety, whereas DASS-D identified depression in 29 individuals. The mean tinnitus duration was 42.99±45.92 months (range: 6-160 months), and the mean THI score was 15.99±15.46 (range: 1-50). Based on the THI severity classification, 66, 20, and 19 participants had slight, mild, and moderate tinnitus, respectively. None of the participants reported ear pain. Tinnitus was bilateral in 86 participants (81.9%) and unilateral in 19 participants (18.1%), of whom 11 had right-sided tinnitus and 8 had left-sided tinnitus. Participant characteristics are shown in Table 1.

The median score for the SAD-T-TR across all participants was 1.00 (3.00). No statistically significant difference was found in SAD-T-TR scores between male and female participants (p=0.384). A moderate positive correlation was observed between SAD-T-TR and THI scores (r=0.350, p<0.001). No significant correlations were found between SAD-T-TR and tinnitus duration (r=-0.155, p=0.114) or between SAD-T-TR and age (r=0.136, p=0.167). Two participants (1.9%) scored the highest, and 28 participants (26.7%) scored the lowest. The SAD-T-TR scale exhibited a floor effect, and no data were lost. Item characteristics, means and standard deviations, and corrected item-total correlations are shown in Table 2. The item-total correlation coefficients of all items were higher than 0.3.

Validity Analysis

Concurrent validity was examined by comparing the DASS-21 total and subscale scores with the SAD-T-TR scores (Table 3). A moderately positive correlation was observed between SAD-T-TR and each DASS-21 subscale, as well as with the total score (p<0.001).

Discriminant validity was evaluated by classifying participants based on DASS-21 anxiety and depression subscale scores. SAD-T-TR scores were compared between tinnitus participants with and without anxiety and between those with and without depression (Table 4). Participants with anxiety and depression showed significantly higher SAD-T-TR scores than those without anxiety or depression.

The construct validity of the SAD-T-TR was tested using CFA. The original study suggested a one-factor structure for the SAD-T (15); therefore, CFA was performed to test this structure (Figure 1). All standardized factor loadings were ≥0.68, indicating strong item-factor relationships. A correlated error term was specified between items 1 and 3 (r=0.35) based on the modification indices. In PHQ-4 validation studies, the anxiety and depression subfactors have been shown to be highly intercorrelated, and under a unidimensional model this structural overlap may manifest as residual covariance between items tapping the two constructs

J Acad Res Med 2026;16(2):91-97

Table 1. Participant characteristics

Characteristic		n (%)
Gender	Male	33 (31.4%)
	Female	72 (68.6%)
Education	Primary education	3 (2.9%)
	Middle school	2 (1.9%)
	High school	14 (13.3%)
	Graduate	80 (76.2%)
	Postgraduate	6 (5.7%)
Anxiety	No (DASS-A score ≤3)	58 (55.2%)
	Yes (DASS-A score >3)	47 (44.8%)
Depression	No (DASS-D score ≤4)	76 (72.4%)
	Yes (DASS-D score >4)	29 (27.6%)
THI classification	Slight	66 (62.9%)
	Mild	20 (19.0%)
	Modarate	19 (18.1%)

DASS-A: Depression anxiety stress scale-21-anxiety subscale, DASS-D: Depression anxiety stress scale-21-depression subscale, THI: Tinnitus handicap inventory

Table 2. Item characteristics of SAD-T-TR

Item	Question	Mean	SD	Corrected item-total correlation (r)
Item 1	Feeling nervous, anxious or on edge	0.50	0.76	0.707
Item 2	Not being able to stop or control worrying	0.61	0.83	0.749
Item 3	Little interest or pleasure in doing things	0.59	0.84	0.728
Item 4	Feeling down, depressed or hopeless	0.66	0.80	0.770

SAD-T-TR: Turkish version of screening for anxiety and depression in tinnitus-hyperacusis-misophonia, SD: Standard deviation

Table 3. Concurrent validity findings

	DASS-total	DASS-A	DASS-D	DASS-S	SAD-T-TR
DASS-total	1.000				
DASS-A	0.931*	1.000			
DASS-D	0.936*	0.867*	1.000		
DASS-S	0.899*	0.732*	0.759*	1.000	
SAD-T-TR	0.680*	0.670*	0.586*	0.639*	1.000

*: Spearman correlation, p-value is significant at 0.001 level, DASS-A: Depression anxiety stress scale-21-anxiety subscale, DASS-D: Depression anxiety stress scale-21-depression subscale, DASS-S: Depression anxiety stress scale-21-stress subscale, SAD-T-TR: Turkish version of the screening for anxiety and depression in tinnitus-hyperacusis-misophonia

(27). The model yielded an acceptable fit to the data [χ^2 (1)=3.046, p=0.081, χ^2/df=3.046]. Incremental fit indices indicated an excellent model fit (CFI=0.991, TLI=0.945, IFI=0.991, NFI=0.987). Absolute fit indices were also satisfactory (GFI=0.986, RMR=0.010). Although the RMSEA value was 0.140 (90% confidence interval: 0.000-0.332), RMSEA is known to be unstable and systematically inflated in models with very few degrees of freedom; simulation studies have demonstrated that RMSEA should not be interpreted in isolation when df=1 (28). Given the non-significant chi-square, high incremental fit indices, and the inherent psychometric constraints of four-item scales, these results should be interpreted with appropriate caution. Nevertheless, the overall pattern of findings supports the unidimensional structure of the SAD-T-TR.

Reliability Analysis

The internal consistency and test–retest reliability of the SAD-T-TR are presented in Table 5. The ICC value for test-retest reliability was 0.945, and the α value was 0.879. The ICC (test-retest reliability) and Cronbach's α (internal consistency) indicated that SAD-T-TR has excellent reliability.

DISCUSSION

The present study aimed to evaluate the validity and reliability of the SAD-T-TR in individuals with tinnitus. Overall, the results show that the SAD-T-TR is a valid and reliable instrument with satisfactory psychometric properties. In line with previous

J Acad Res Med 2026;16(2):91-97

Table 4. Discriminant validity findings

	SAD-T-TR median (IQR)	U	Z	p-value*	Effect size (r)
Anxiety (+) (n=47)	4.00 (4.00)	345.000	-6.697	<0.001	0.65
Anxiety (-) (n=58)	1.00 (1.00)				
Depression (+) (n=29)	4.00 (4.00)	434.000	-4.887	<0.001	0.48
Depression (-) (n=76)	1.00 (2.00)				

Note: Anxiety (+) defined as DASS-A score >3, depression (+) defined as DASS-D score >4
*: Mann-Whitney U test, SAD-T-TR: Turkish version of the screening for anxiety and depression in tinnitus-hyperacusis-misophonia, IQR: Interquartile range, DASS-A: Depression anxiety stress scale-21-anxiety subscale, DASS-D: Depression anxiety stress scale-21-depression subscale

Table 5. Internal consistency and test-retest reliability of the SAD-T-TR

	Test-retest reliability		Internal consistency
	Intraclass correlation coefficient	95% CI	Cronbach's α
SAD-T-TR	0.945	0.876-0.975	0.879

SAD-T-TR: Turkish version of the screening for anxiety and depression in tinnitus-hyperacusis-misophonia, CI: Confidence interval

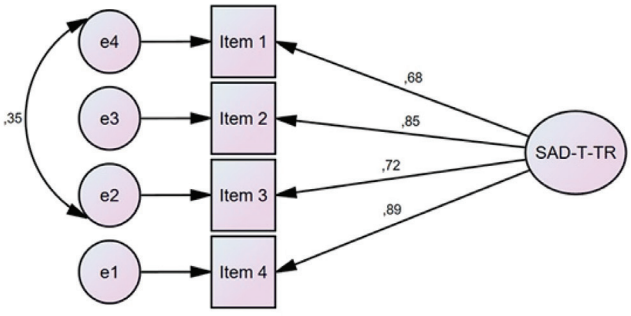

Figure 1. Confirmatory factor analysis of SAD-T-TR
SAD-T-TR: Turkish version of the screening for anxiety and depression in tinnitus-hyperacusis-misophonia

research (15), CFA validated the scale's one-factor structure. The scale showed sufficient construct, concurrent, and discriminant validity as well as good internal consistency and test-retest reliability. People with tinnitus often report psychological distress, especially anxiety and depression, which is thought to be one of the main causes of tinnitus-related disability and a lower quality of life (12,29). Thus, quick screening methods that can quickly detect psychological symptoms could be very helpful for tinnitus treatment in audiology clinic. SAD-T was developed as an ultra-brief screening instrument with the express purpose of identifying symptoms of anxiety and depression in people seeking treatment for tinnitus.

The SAD-T-TR had a high internal consistency (Cronbach's α=0.879), which means that the items measure a common underlying construct. Item-total correlations for SAD-T-TR items ranged between 0.707 and 0.770. These correlation coefficients demonstrate strong consistency and distinguish among different levels of the measured trait. Aazh et al. (17) also found a Cronbach's α value of 0.91 for the SAD-T in a clinical tinnitus population. The item-total correlation coefficients ranged from 0.76 to 0.84, indicating that each item made a strong contribution to the total score. The scale also has high test-retest reliability (ICC=0.945), which means that it gives consistent results over time. These results collectively indicate that the SAD-T is reliable across diverse populations and languages.

The current study found that SAD-T-TR scores demonstrated a floor effect (26.7%), exceeding the commonly accepted threshold of 15%. Similar floor effects have been reported in ultra-brief screening instruments, including the PHQ-4, particularly in samples with low levels of psychological distress (30,31). A systematic review have also indicated that such effects may occur in very brief screening tools designed primarily to identify clinically relevant symptoms rather than to capture the full range of symptom severity (32). This finding may partly reflect the distribution of psychological symptoms within the present tinnitus sample, which included individuals with minimal or no clinically significant distress. The primary aim of the SAD-T is rapid screening rather than detailed symptom profiling; this emphasis may further contribute to this pattern. Nevertheless, the presence of a floor effect should be considered a potential limitation, as it may reduce the scale's sensitivity in distinguishing subtle differences among individuals with low levels of psychological distress.

The CFA supported a one-factor structure for the SAD-T-TR, consistent with the original validation study (15). This unidimensional structure is supported by satisfactory model fit indices and strong standardized factor loadings (≥0.68), indicating that all items adequately represent a single latent construct. Conversely, the PHQ-4 are generally regarded as two-factor models representing distinct constructs of anxiety and depression (27,33). The difference between these structures may be explained by the conceptual purpose of the instruments and the characteristics of the sample. The PHQ-4 is used to assess two areas of mental health, anxiety and depression, while the SAD-T is a brief instrument that assesses general emotional distress specifically related to tinnitus, hyperacusis, and misophonia.

J Acad Res Med 2026;16(2):91-97

In individuals with tinnitus, anxiety and depressive symptoms frequently co-occur and may reflect a shared underlying distress dimension rather than clearly separable constructs. These findings support the notion that the unidimensional structure identified for the SAD-T may be population-specific. Unlike studies using the PHQ-4 that typically distinguish anxiety and depression as separate constructs, individuals with tinnitus may experience these symptoms as part of a broader, more integrated emotional-distress response.

The results of the concurrent validity tests add to the evidence that the scale is validly constructed. A moderate correlation was observed between SAD-T-TR scores and DASS-21 total and subscale scores. The current study also supports discriminant validity. People diagnosed with anxiety or depression based on their DASS-21 scores had SAD-T-TR scores much higher than those who did not have these symptoms. This finding indicates that the scale can effectively distinguish between individuals exhibiting clinically significant psychological symptoms and those who do not. These findings confirm that the SAD-T-TR scale can identify psychological symptoms more quickly and consistently.

There were no significant differences in SAD-T-TR scores between men and women. This result could suggest that the psychological distress associated with tinnitus is influenced less by gender and more by how the person perceives tinnitus and how it affects their daily life. Other studies on tinnitus have also suggested that emotional distress related to tinnitus may depend more on tinnitus severity and how well people deal with it than on things like gender (17). Moreover, associations between sound sensitivity symptoms and mental health disorders have been documented in previous research (34).

Psychological symptoms may also play a role in how individuals with tinnitus perceive and cope with their condition, as well as in their response to treatment. Previous research indicates that higher SAD-T scores correlate with increased tinnitus handicap and diminished confidence in managing tinnitus (11,35). In our study, a moderate positive correlation was found between SAD-T-TR and THI scores (r=0.35), indicating that higher levels of anxiety and depressive symptoms are associated with a greater impact of tinnitus. Although this association was statistically significant, its magnitude indicates that psychological distress represents only one of several factors contributing to the perceived impact of tinnitus. Nevertheless, these findings support the potential value of including a brief psychological screening tool as part of a comprehensive tinnitus assessment.

Study Limitations

Several limitations should be considered when interpreting the findings. First, the sample's relatively large age variability may further restrict the generalizability of the results to specific age groups. Although no significant sex-related differences were observed in SAD-T scores, the predominance of women and individuals with higher educational attainment in the sample may limit the generalizability of the findings. Future studies involving more demographically diverse tinnitus populations are warranted. Second, the study relied on self-report measures that may be affected by response bias, potentially leading to inaccuracies in the collected data and limiting the reliability of the findings. Third, participants were included only if their hearing thresholds were below 25 dB HL, consistent with the standard audiological definition of normal hearing and with the inclusion criteria used in previous tinnitus research (36,37). Although this criterion was applied to minimize the confounding influence of peripheral hearing loss on psychological distress scores, it excludes a substantial proportion of the tinnitus population with hearing loss, thereby limiting the generalizability of the findings. Fourth, since only individuals with tinnitus were included, the findings should not be generalized to individuals with hyperacusis or misophonia without further validation studies. Finally, the CFA was based on a four-item scale with df=1, which constrains the reliability of fit indices such as RMSEA and warrants cautious interpretation of the model fit results. Future studies involving more demographically diverse samples and larger populations with tinnitus, hyperacusis, and misophonia are warranted to further examine the psychometric properties and clinical utility of the SAD-T-TR.

CONCLUSION

The SAD-T-TR exhibited satisfactory psychometric properties in individuals with tinnitus. The scale demonstrated strong internal consistency, excellent test-retest reliability and adequate construct validity, concurrent validity, and discriminant validity. The CFA validated the one-factor structure of the scale. The SAD-T-TR may be a useful screening tool for detecting symptoms of anxiety and depression in people with tinnitus because it is brief, easy to use, and psychometrically sound. Using concise instruments in clinical audiology may enhance prompt detection of psychological distress and promote a more holistic and interdisciplinary approach to tinnitus management. Subsequent research should examine the scale's sensitivity to treatment-related variations and assess its applicability in larger tinnitus populations.

Ethics

Ethics Committee Approval: Trakya University Non-Interventional Ethics Committee approved it with decision number 05/19 and protocol code 2025/104 (date: 06.01.2025).

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

Footnotes

Author Contributions: Concept - N.B.A.; Design - N.B.A., İ.T.G.; Data Collection and/or Processing - N.B.A., İ.T.G.; Analysis and/or Interpretation - N.B.A.; Literature Search - İ.T.G.; Writing - N.B.A.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors report that no financial support was received for this study.

REFERENCES

1. Møller AR. Epidemiology of tinnitus in adults. Textbook of tinnitus: Springer; 2011. p. 29-37.
2. Jacques D, Nozeret Y, Zdanowicz N, Reynaert C, Garin P, Gilain C. Tinnitus and psychiatric comorbidities in liaison psychiatry analysis of three years in an audiophonology centre. Psychiatr Danub. 2013; 25(Suppl 2): S102-4.

J Acad Res Med 2026;16(2):91-97

3. Salazar JW, Meisel K, Smith ER, Quiggle A, McCoy DB, Amans MR. Depression in patients with tinnitus: a systematic review. Otolaryngol Head Neck Surg. 2019; 161: 28-35.
4. Bhatt JM, Bhattacharyya N, Lin HW. Relationships between tinnitus and the prevalence of anxiety and depression. Laryngoscope. 2017; 127: 466-9.
5. Henry JA. "Measurement" of tinnitus. Otol Neurotol. 2016; 37: e276-85.
6. Hall DA, Haider H, Szczepek AJ, Lau P, Rabau S, Jones-Diette J, et al. Systematic review of outcome domains and instruments used in clinical trials of tinnitus treatments in adults. Trials. 2016; 17: 270.
7. Newman CW, Jacobson GP, Spitzer JB. Development of the tinnitus handicap inventory. Arch Otolaryngol Head Neck Surg. 1996; 122: 143-8.
8. Aksoy S, Firat Y, Alpar R. The tinnitus handicap inventory: a study of validity and reliability. Int Tinnitus J. 2007; 13: 94-8.
9. Meikle MB, Henry JA, Griest SE, Stewart BJ, Abrams HB, McArdle R, et al. The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. Ear Hear. 2012; 33: 153-76. Erratum in: Ear Hear. 2012; 33: 443.
10. Aazh H, Hayes C, Moore BCJ, Vitoratou S. Psychometric evaluation of the tinnitus impact questionnaire using patients seeking help for tinnitus or tinnitus with hyperacusis. Int J Audiol. 2023; 62: 835-44.
11. Aazh H, Moore BCJ, Erfanian M. Confirmatory factor analysis of the tinnitus impact questionnaire using data from patients seeking help for tinnitus alone or tinnitus combined with hyperacusis. PLoS One. 2024; 19: e0302837.
12. Hackenberg B, Döge J, O'Brien K, Bohnert A, Lackner KJ, Beutel ME, et al. Tinnitus and its relation to depression, anxiety, and stress-a population-based cohort study. J Clin Med. 2023; 12: 1169.
13. Husain WSW, Othman A, Othman NAN, Mohamad WNW, Zakaria MN. Determining the internal and external reliability of depression, anxiety and stress scales (DASS-21) in assessing psychological symptoms among patients with tinnitus. NeuroQuantology. 2018; 16: 97-102.
14. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. Gen Hosp Psychiatry. 2016; 39: 24-31.
15. Aazh H, Hayes C, Erfanian M, Moore BCJ, Vitoratou S. Confirmatory factor analysis of the hyperacusis impact questionnaire, sound sensitivity symptoms questionnaire, and screening for anxiety and depression in tinnitus, including preliminary analyses of the parent versions for use with children. J Am Acad Audiol. 2024; 35: 81-92.
16. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics. 2009; 50: 613-21.
17. Aazh H, Hayes C, Moore BCJ, Danesh AA, Vitoratou S. Psychometric evaluation of the hyperacusis impact questionnaire (HIQ) and sound sensitivity symptoms questionnaire (SSSQ) using a clinical population of adult patients with tinnitus alone or combined with hyperacusis. J Am Acad Audiol. 2022; 33: 248-58.
18. Sarıçam H. The psychometric properties of Turkish version of depression anxiety stress scale-21 (DASS-21) in health control and clinical samples. JCBPR. 2018; 7: 19-30.
19. Lovibond SH. Manual for the depression anxiety stress scales. Sydney Psychology Foundation. 1995.
20. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. Qual Life Res. 1998; 7: 323-35.
21. Field A. Discovering statistics using SPSS. 3rd ed. London: SAGE Publications Ltd; 2009.
22. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling. 1999; 6: 1-55.
23. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. MPR-online. 2003; 8: 23-74.
24. Abma IL, Rovers M, van der Wees PJ. Appraising convergent validity of patient-reported outcome measures in systematic reviews: constructing hypotheses and interpreting outcomes. BMC Res Notes. 2016; 9: 226.
25. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press. 2024.
26. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005; 19: 231-40.
27. Adzrago D, Walker TJ, Williams F. Reliability and validity of the patient health questionnaire-4 scale and its subscales of depression and anxiety among US adults based on nativity. BMC Psychiatry. 2024; 24: 213.
28. Kenny DA, Kaniskan B, McCoach DB. The performance of RMSEA in models with small degrees of freedom. Sociological Methods & Research. 2015; 44: 486-507.
29. Haider HF, Ribeiro SF, Hoare DJ, Fialho G, Hall DA, Antunes M, et al. Quality of life and psychological distress in Portuguese older individuals with tinnitus. Brain Sci. 2021; 11: 953.
30. Materu J, Kuringe E, Nyato D, Galishi A, Mwanamsangu A, Katebalila M, et al. The psychometric properties of PHQ-4 anxiety and depression screening scale among out of school adolescent girls and young women in Tanzania: a cross-sectional study. BMC Psychiatry. 2020; 20: 321.
31. Meidl V, Dallmann P, Leonhart R, Bretthauer B, Busch A, Kubosch EJ, et al. Validation of the patient health questionnaire-4 for longitudinal mental health evaluation in elite Para athletes. PM R. 2024; 16: 141-9.
32. Caro-Fuentes S, Sanabria-Mazo JP. A systematic review of the psychometric properties of the patient health questionnaire-4 in clinical and nonclinical populations. J Acad Consult Liaison Psychiatry. 2024; 65: 178-94.
33. Christodoulaki A, Baralou V, Konstantakopoulos G, Touloumi G. Validation of the patient health questionnaire-4 (PHQ-4) to screen for depression and anxiety in the Greek general population. J Psychosom Res. 2022; 160: 110970.
34. Siepsiak M, Rosenthal MZ, Raj-Koziak D, Dragan W. Psychiatric and audiologic features of misophonia: use of a clinical control group with auditory over-responsivity. J Psychosom Res. 2022; 156: 110777.
35. Aazh H, Kartsonaki C, Moore BCJ. Psychometric evaluation of the 4C tinnitus management questionnaire for patients with tinnitus alone or tinnitus combined with hyperacusis. Int J Audiol. 2024; 63: 21-9.
36. Goodman A. Reference zero levels for pure-tone audiometer. ASHA. 1965; 7: 1.
37. Spankovich C, Gonzalez VB, Su D, Bishop CE. Self reported hearing difficulty, tinnitus, and normal audiometric thresholds, the National Health and Nutrition Examination Survey 1999-2002. Hear Res. 2018; 358: 30-6.

DOI: 10.4274/jarem.galenos.2026.36693
J Acad Res Med 2026;16(2):98-105

MRI-based Morphometric Assessment of the Pediatric Cervical Spinal Cord and Spinal Canal

Muhammed Samed Cansız[1], Mehmet Öztürk[1], Zuhal Bayramoğlu[2], Nurullah Dağ[3]

[1]Selçuk University Faculty of Medicine, Department of Radiology, Konya, Türkiye
[2]İstanbul University-İstanbul Faculty of Medicine, Department of Pediatric Radiology, İstanbul, Türkiye
[3]İnönü University Faculty of Medicine, Department of Radiology, Malatya, Türkiye

Cite this article as: Cansız MS, Öztürk M, Bayramoğlu Z, Dağ N. MRI-based morphometric assessment of the pediatric cervical spinal cord and spinal canal. J Acad Res Med. 2026;16(2):98-105

ABSTRACT

Objective: To establish magnetic resonance imaging (MRI)-based reference morphometrics of the pediatric cervical spinal cord and spinal canal in children and adolescents with normal cervical MRI findings, and to assess age-, sex-, and level-related variation.

Methods: Cervical spine MRI examinations from 300 children and adolescents, including 150 girls and 150 boys aged 3-17 years, were retrospectively analyzed. All subjects had normal cervical spinal cord and spinal canal morphology on imaging. Participants were equally distributed by sex within three age groups: 3-6 years, Group 1; 7-12 years, Group 2; and 13-17 years, Group 3. T2-weighted sagittal and axial images were obtained using a 1.5-T MRI system. At the mid-vertebral levels of C2, C4, and C6, anteroposterior diameters and cross-sectional areas of the spinal cord and spinal canal were measured using standardized protocols. Appropriate parametric or non-parametric tests were used; $p<0.05$ was considered significant.

Results: Across C2, C4, and C6, boys had significantly larger anteroposterior diameters and cross-sectional areas of the spinal cord and canal than girls (all $p<0.05$). The spinal cord cross-sectional area showed the most consistent age-related increase. At C4, mean cord area increased from 52.3±4.1 mm^2 in Group 1 to 58.7±4.6 mm^2 in Group 2 and 63.4±5.2 mm^2 in Group 3. At C2, cord anteroposterior diameter increased progressively with age, 6.8±0.5, 7.3±0.6, and 7.9±0.7 mm for Groups 1-3, respectively; $p<0.001$. At C6, canal anteroposterior diameter and canal area increased with age, whereas the cord anteroposterior diameter did not change significantly.

Conclusion: Pediatric cervical cord and canal dimensions vary by sex, age, and vertebral level in children and adolescents with normal cervical MRI findings. Cross-sectional area measurements may complement anteroposterior diameters. These MRI-based reference values may support pediatric cervical MRI interpretation and developmentally informed surgical planning.

Keywords: Pediatric cervical spine, morphometry, magnetic resonance imaging

INTRODUCTION

The cervical spine and its neural contents represent an integrated biomechanical and neuroanatomical unit whose structural integrity is fundamental to both mechanical stability and neurological function (1). While the vertebral column provides mobile support for the head and protects the spinal cord and nerve roots, it also plays a critical role in coordinating sensory inputs, including vestibular, visual, and auditory information, which together maintain posture and balance.

Throughout the pediatric age range, continuous growth and maturation of the cervical spine shape its morphometric characteristics through a dynamic interplay of developmental biology, biomechanical factors, and clinical influences. Variations in spinal canal and cord dimensions, as well as pathological changes that may arise, carry important implications for vulnerability to injury, accurate symptom interpretation, and informed treatment planning.

Establishing accurate reference morphometric data for the pediatric cervical spinal cord and canal serves purposes far beyond descriptive documentation; such reference values are clinically essential (2). Reliable age- and sex-specific reference values facilitate early identification of developmental deviations, enhance differential diagnostic reasoning, and enable more confident radiological interpretation when pediatric patients present with cervical complaints or incidental imaging findings. Furthermore, management of pediatric cervical pathology frequently depends on size-dependent clinical considerations, particularly in surgical planning for instrumentation and fusion procedures. Therefore,

ORCID IDs of the authors: M.S.C. 0000-0003-4608-8194; M.Ö. 0000-0002-4316-5038; Z.B. 0000-0002-2080-2647; N.D. 0000-0002-9342-0244

Corresponding Author: Mehmet Öztürk, Prof. Dr.;
E-mail: drmehmet2121@gmail.com

Received Date: 31.01.2026 **Accepted Date:** 21.07.2026
Epub: 27.07.2026
Publication Date: 07.08.2026

J Acad Res Med 2026;16(2):98-105

age- and sex-stratified morphometric reference standards are invaluable for guiding procedural decisions and informing risk assessment (3).

Previous investigations of pediatric cervical morphometry have predominantly relied on conventional radiography and computed tomography to characterize the bony anatomy. However, these imaging modalities suffer from limited soft-tissue contrast, which constrains the availability and precision of reference data for spinal cord measurements, especially anteroposterior diameters and cross-sectional areas (4,5). In contrast, magnetic resonance imaging (MRI) offers superior soft-tissue visualization and multiplanar imaging capabilities, and is currently recognized as the reference standard for comprehensive evaluation of the cervical spine and spinal cord. MRI thus represents the optimal imaging platform for precise, level-specific measurement of both the spinal canal and neural structures, permitting detailed characterization of developmental morphometry.

This retrospective study aimed to establish MRI-based reference morphometric data for the cervical spinal cord and spinal canal in children and adolescents with normal cervical spine MRI examinations. Our specific objectives were to quantify anteroposterior diameters and cross-sectional areas at three anatomically distinct cervical levels (C2, C4, and C6) and to systematically evaluate morphometric variation by age, sex, and vertebral level in a well-stratified pediatric cohort.

METHODS

Study Design, Setting, and Ethics

This retrospective, single-center morphometric study was conducted at our institution, using cervical spine MRI examinations acquired from May 2018 through October 2023. The research protocol was approved by the Local Institutional Ethics Committee of Selçuk University (decision no: 2024/448, date: 18.09.2024). Given the retrospective study design, informed consent procedures were waived, and the study was conducted in accordance with institutional policies and ethics committee guidelines.

Study Population

The study cohort consisted of 300 pediatric subjects, including 150 girls and 150 boys, aged 3-17 years. The cohort was derived from pediatric patients who underwent cervical spine MRI for clinical indications and who demonstrated normal cervical spinal canal and spinal cord morphology on imaging.

To achieve balanced demographic representation across both age and sex dimensions, a stratified sampling approach was employed. For each single-year age interval from 3 through 17 years, 10 girls and 10 boys were selected consecutively, yielding 20 subjects per age category across 15 age groups, for a total sample size of 300.

Eligible MRI examinations were identified in a subset of pediatric patients younger than 17 years who underwent cervical spine imaging for various clinical indications, including neck pain, shoulder pain, and suspected disc herniation. All included examinations demonstrated normal spinal morphology with no imaging evidence of pathology affecting the cervical spinal canal or spinal cord.

Participants were stratified into three developmentally defined age groups: Group 1, 3-6 years, early childhood; Group 2, 7-12 years, middle childhood; and Group 3, 13-17 years, adolescence.

Inclusion Criteria

Participants met the following criteria for study inclusion:

1. Age between 3 and 17 years;
2. Availability of diagnostic-quality cervical MRI including both T2-weighted sagittal and axial image sequences;
3. Absence of imaging evidence for pathology affecting cervical spinal canal or spinal cord morphometry.

Exclusion Criteria

Examinations were excluded if imaging or clinical records indicated conditions likely to substantially alter the dimensions of the spinal canal or cord or to introduce systematic morphometric bias. These conditions included spinal stenosis, compressive myelopathy, spinal cord neoplasm, syringomyelia, inflammatory spinal pathology, spinal cord ischemic changes, spinal cord injury, Chiari malformation, tonsillar herniation, and prior head and neck surgery.

Additionally, studies demonstrating substantial motion artifact, inadequate anatomical coverage of the cervical spine, or incomplete image sequences that precluded standardized measurements at the designated vertebral levels were excluded.

MRI Protocol and Morphometric Measurements

All imaging examinations were performed on a 1.5-T Siemens Magnetom MRI system (Siemens Medical Solutions, Erlangen, Germany). To minimize variability attributable to patient positioning, all participants were scanned in a standardized neutral supine position.

Morphometric analysis was conducted using T2-weighted sequences:

1. Sagittal turbo spin-echo sequence with the following parameters: repetition time=3000 ms, echo time=108 ms, and slice thickness=3.4 mm;
2. Axial T2-weighted sequence with repetition time=4280 ms, echo time=105 ms, and slice thickness=3.0 mm.

Field of view, acquisition matrix dimensions, and in-plane resolution were systematically retrieved from DICOM image headers and documented to ensure consistency across all reported measurements.

All quantitative measurements were performed on a calibrated picture archiving and communication system workstation with integrated electronic calipers and region-of-interest analysis tools, using syngo.via software (Siemens Healthineers, Erlangen, Germany).

J Acad Res Med 2026;16(2):98-105

Morphometric measurements were obtained from precisely defined mid-vertebral-body at three cervical levels (C2, C4, and C6) using a standardized anatomical approach. On mid-sagittal T2-weighted images, the anteroposterior diameters of both the spinal canal and the spinal cord were measured at each level, perpendicular to the long axis of the canal, thereby minimizing potential overestimation due to oblique image planes.

On corresponding axial T2-weighted images, the cross-sectional areas of the spinal canal and spinal cord were quantified by manual delineation of regions of interest. The canal region of interest was traced along the inner boundary of the osseoligamentous canal, while the cord region of interest was outlined along the outer contour of the spinal cord.

To minimize partial-volume averaging effects and ensure accurate level identification, the single axial slice that best represented the mid-vertebral body level and provided optimal definition of anatomical margins was selected for measurement at each level, as shown in Figures 1 and 2.

Measurement Reliability and Quality Assurance

All morphometric measurements were performed by two independent readers with substantial experience in pediatric MRI, using a consensus methodology. Reader 1 had 15 years of experience, and reader 2 had 5 years of experience. Both observers were blinded to all clinical information and to the original imaging indication throughout the measurement process.

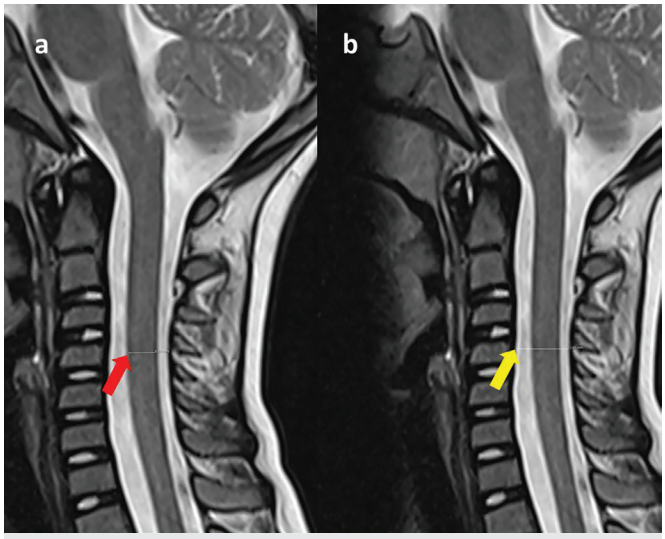

Figure 1. Sagittal T2-weighted cervical magnetic resonance imaging demonstrating anteroposterior diameter measurements at the C4 level. Panel (a) illustrates the measurement of the spinal cord anteroposterior diameter (indicated by the red arrow), while panel (b) demonstrates the corresponding measurement of the spinal canal anteroposterior diameter (indicated by the yellow arrow). Both measurements were obtained at the mid-vertebral body level, which represents the standardized location for diameter assessment in this study. The sagittal plane provides visualization of the entire cervical column, enabling precise identification of the C4 vertebral level and accurate linear measurements in the anteroposterior dimension

For reliability assessment, both readers independently measured the selected subset of 50 examinations before consensus review. Inter-rater reliability was calculated using these independent preliminary measurements. After reliability assessment, discrepant measurements were resolved by joint review, and a single consensus value was used for final statistical analysis.

Intra-observer reliability was assessed by repeated measurements performed by the primary reader after a two-week interval to minimize recall bias. Intra-class correlation coefficients were calculated using a two-way mixed-effects model with absolute agreement. All intraclass correlation coefficients (ICC) values for both inter-rater and intra-observer reliability exceeded 0.80, indicating excellent reproducibility across all evaluated parameters.

Representative ICC values were as follows: C2 cord diameter=0.87; C2 cord area=0.89; C2 canal diameter=0.85; C2 canal area=0.88. Comparable values were obtained for the C4 and C6 levels.

Statistical Analysis

All statistical computations were performed using IBM SPSS Statistics, version 23.0 (IBM Corporation, Armonk, NY, USA). A two-tailed significance level of $p<0.05$ was applied as the threshold for statistical significance throughout all analyses unless explicitly stated otherwise.

Primary Outcome Variables

Primary outcome variables measured at each of the three vertebral levels, C2, C4, and C6, comprised:

1. Anteroposterior spinal canal diameter and anteroposterior spinal cord diameter, measured in the sagittal plane;
2. Spinal canal cross-sectional area and spinal cord cross-sectional area, measured in the axial plane.

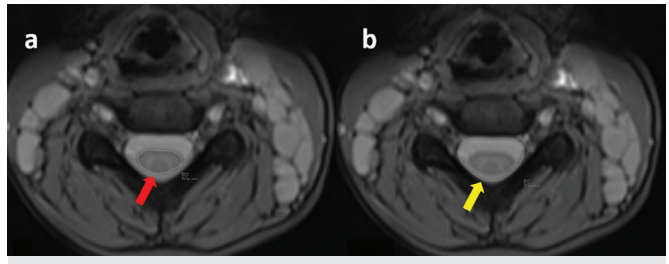

Figure 2. Axial T2-weighted cervical magnetic resonance imaging demonstrating cross-sectional area measurements at the C4 level. Panel (a) shows the measurement of the spinal cord cross-sectional area (indicated by the red arrow), while panel (b) displays the corresponding measurement of the spinal canal cross-sectional area (indicated by the yellow arrow). Both measurements were performed by manually tracing the outer contours of the spinal cord and spinal canal on the axial image acquired at the mid-vertebral body level of C4. Cross-sectional area measurements provide a more comprehensive assessment of the available space than linear anteroposterior diameter measurements alone, particularly when assessing canal morphology that deviates from a simple circular configuration

J Acad Res Med 2026;16(2):98-105

Descriptive Statistics

Continuous variables were assessed for normality using the Shapiro-Wilk test together with visual inspection of histograms and quantile-quantile plots. Variance homogeneity was evaluated using Levene's test. Because most morphometric variables were not normally distributed, data were summarized as medians with interquartile ranges (IQR). The IQR was calculated as the difference between the 75th and 25th percentiles. Categorical variables were summarized as frequencies and percentages.

All statistical analyses were performed separately for each vertebral level and stratified by age group and biological sex.

Comparative Statistical Tests

Comparisons between girls and boys were conducted using the independent-samples Student's t-test when assumptions for parametric tests were satisfied; otherwise, Welch's t-test was used when variances were unequal. For non-normally distributed continuous variables, the non-parametric Mann-Whitney U test was applied.

Comparisons among the three age-defined groups were performed using one-way analysis of variance when parametric assumptions were satisfied; for non-normally distributed data, the non-parametric Kruskal-Wallis H test was used.

Post-hoc pairwise comparisons between groups following significant Kruskal-Wallis test results were conducted using Dunn's multiple-comparison test with Bonferroni adjustment. Using three pairwise comparisons per parameter, the Bonferroni-adjusted significance threshold was established at $p<0.0167$. Accordingly, pairwise comparisons with p-values below this threshold were considered statistically significant.

Correlation Analysis

Spearman correlation coefficients were calculated to assess the association between age and each morphometric parameter at each vertebral level.

RESULTS

Study Population Characteristics

The analysis included 300 pediatric subjects, comprising 150 boys and 150 girls. The median age of the cohort was 10 years (IQR, 6-14 years; IQR=8 years). The age distribution was identical between boys and girls, with no statistically significant sex-related difference observed (p=1.000). This perfect age matching reflected the successful implementation of the stratified sampling design.

Sex-stratified descriptive statistics for all measured morphometric parameters are presented in Table 1.

Sex-related Morphometric Differences

At all evaluated cervical vertebral levels (C2, C4, and C6), boys demonstrated consistently larger spinal cord and spinal canal dimensions than girls. Regarding spinal cord anteroposterior diameter, boys exhibited significantly greater measurements than girls at all three levels: C2, p=0.012; C4, p=0.001; and C6, p=0.018.

Similarly, spinal cord cross-sectional area was significantly larger in boys than in girls at C2 (p=0.003), C4 (p=0.009), and C6 (p=0.003).

The spinal canal demonstrated a parallel pattern of sexual dimorphism. Canal anteroposterior diameter measurements were significantly higher in boys than in girls at all examined levels: C2, p=0.009; C4, p=0.001; and C6, p=0.001.

Table 1. Descriptive statistics of age and cervical spinal cord/spinal canal measurements by sex

Parameter	Male (n=150) Median (IQR)	Female (n=150) Median (IQR)	Total sample (n=300) Median (IQR)	p-value
Age (years)	10 (6-14, 8)	10 (6-14, 8)	10 (6-14, 8)	1.000
C2 spinal cord AP diameter (mm)	6.3 (5.87-6.6)	6.2 (5.6-6.6)	6.3 (5.7-6.6)	0.012
C2 spinal canal AP diameter (mm)	13.8 (12.87-14.7)	13.5 (12.6-14.2)	13.6 (12.8-14.4)	0.009
C2 spinal cord area (mm^2)	67 (59-76)	63.35 (55-72)	65.2 (56.87-73.8)	0.003
C2 spinal canal area (mm^2)	240 (214-273)	231 (205-260)	235.25 (209.7-266)	0.027
C4 spinal cord AP diameter (mm)	6.3 (5.9-6.6)	6.0 (5.6-6.4)	6.1 (5.7-6.6)	0.001
C4 spinal canal AP diameter (mm)	11.9 (11.2-12.6)	11.35 (10.6-12.2)	11.7 (10.8-12.4)	0.001
C4 spinal cord area (mm^2)	69 (59.1-76.95)	63.7 (55-73.9)	65.75 (57.22-75.6)	0.009
C4 spinal canal area (mm^2)	188.3 (165-207.4)	180.75 (166.3-199.37)	186.3 (165.5-202.72)	0.010
C6 spinal cord AP diameter (mm)	6.0 (5.6-6.3)	5.9 (5.4-6.2)	6.0 (5.6-6.3)	0.018
C6 spinal canal AP diameter (mm)	12.2 (11.57-12.8)	11.8 (10.97-12.4)	12.0 (11.3-12.6)	0.001
C6 spinal cord area (mm^2)	66.75 (60.15-73.8)	63.45 (55-71.2)	65.8 (56.8-72.1)	0.003
C6 spinal canal area (mm^2)	193.95 (173.4-211.1)	187.75 (169.3-199.4)	190 (170-207)	0.010

Note: Data are presented as median [interquartile range (IQR), IQR difference]. The IQR difference was calculated as the difference between the 75th percentile and the 25th percentile. Anteroposterior (AP); C2, C4, and C6 indicate cervical vertebral levels. Sex-based comparisons were performed using the Mann-Whitney U test. A p-value of <0.05 was considered statistically significant

J Acad Res Med 2026;16(2):98-105

Additionally, spinal canal cross-sectional area was significantly greater in boys at C2, p=0.027; C4, p=0.010; and C6, p=0.010. These consistent differences across measurement parameters and vertebral levels underscore significant sex-related variation in cervical morphometry throughout the pediatric age range.

Age-related Morphometric Changes

Age-stratified analyses revealed a progressive increase in cervical morphometric measurements with advancing age. The spinal cord cross-sectional area demonstrated the most consistent age-related increase across all examined vertebral levels. Specifically, spinal cord cross-sectional area increased significantly from Group 1 to Group 2 and from Group 2 to Group 3 at the C2, C4, and C6 levels. In addition, the C2 spinal cord anteroposterior diameter showed a significant stepwise enlargement across all age groups.

At the C4 level, both the spinal cord anteroposterior diameter and the spinal canal cross-sectional area were significantly greater in Groups 2 and 3 than in Group 1. At the C6 level, the spinal canal anteroposterior diameter and cross-sectional area were significantly greater in Groups 2 and 3 than in Group 1, whereas the C6 spinal cord anteroposterior diameter did not differ significantly among age groups. These findings are summarized in Table 2.

Further comparison between the youngest and oldest age groups revealed that, at the C2 level, both the canal anteroposterior diameter and the canal cross-sectional area were significantly larger in Group 3 (13-17 years) than in Group 1 (3-6 years). A similar pattern was observed at the C4 level for the anteroposterior diameter of the canal when comparing Group 3 with Group 1.

Correlation Analysis

Correlation analyses between age and morphometric parameters are presented in Table 3. Age showed a moderate positive correlation with C2 spinal cord anteroposterior diameter (Spearman r=0.48, p=0.001), as illustrated in Figure 3, and a strong

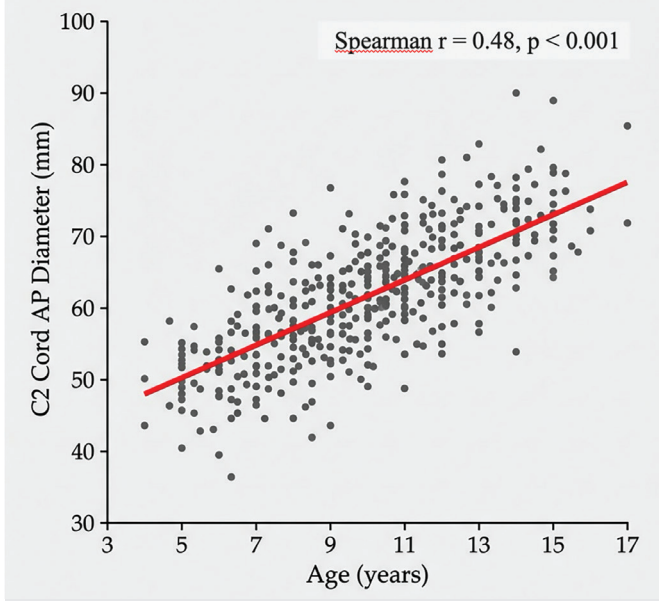

Figure 3. Scatter plot showing the relationship between age and C2 spinal cord anteroposterior (AP) diameter in pediatric subjects (n=300). The red line represents the linear regression fit. Spearman rank correlation coefficient (r=0.48, p<0.001) indicates a moderate monotonic relationship between age and C2 cord diameter

Table 2. Descriptive data of spinal cord and canal measurements according to age groups

Parameter	Group 1 (3-6 years) Median (IQR)	Group 2 (7-12 years) Median (IQR)	Group 3 (13-17 years) Median (IQR)	p-value	Pairwise comparisons (p*)
C2 cord diameter (mm)	5.6 (5.4-6.3)	6.3 (5.9-6.6)	6.6 (6.1-6.9)	0.001	1 vs. 2: 0.001; 1 vs. 3: 0.001; 2 vs. 3: 0.001
C2 canal diameter (mm)	13.4 (12.4-14.3)	13.6 (12.8-14.3)	13.9 (12.72-14.7)	0.037	1 vs. 3: 0.016
C2 cord area (mm^2)	51.35 (45.7-59.7)	64.2 (58.8-72.1)	73.7 (67.62-90.37)	0.001	1 vs. 2: 0.001; 1 vs. 3: 0.001; 2 vs. 3: 0.001
C2 canal area (mm^2)	227 (202-260)	238.4 (208-262.7)	240 (219-280)	0.012	1 vs. 3: 0.003
C4 cord diameter (mm)	5.7 (5.52-6.17)	6.2 (5.9-6.6)	6.3 (6.0-6.6)	0.001	1 vs. 2: 0.001; 1 vs. 3: 0.001
C4 canal diameter (mm)	11.2 (10.5-11.9)	11.9 (11.0-12.4)	11.9 (11.2-12.67)	0.001	1 vs. 2: 0.001; 1 vs. 3: 0.001
C4 cord area (mm^2)	53.3 (46.9-58.6)	66.6 (60.7-74.4)	74.8 (67.7-81.0)	0.001	1 vs. 2: 0.001; 1 vs. 3: 0.001; 2 vs. 3: 0.001
C4 canal area (mm^2)	169 (150-190)	189 (171-207)	191.3 (172-208)	0.001	1 vs. 2: 0.001; 1 vs. 3: 0.001
C6 cord diameter (mm)	5.9 (5.6-6.25)	6.0 (5.52-6.3)	6.0 (5.6-6.3)	0.56	-
C6 canal diameter (mm)	11.7 (10.7-12.2)	12.1 (11.3-12.7)	12.2 (11.4-12.8)	0.001	1 vs. 2: 0.001; 1 vs. 3: 0.001
C6 cord area (mm^2)	57.3 (49.3-65.6)	65.9 (56.8-71.7)	69.3 (64.7-74.9)	0.001	1 vs. 2: 0.001; 1 vs. 3: 0.001; 2 vs. 3: 0.001
C6 canal area (mm^2)	184 (163-201)	192.4 (172.7-209)	190 (174-208)	0.014	1 vs. 2: 0.006; 1 vs. 3: 0.016

Note: Data are presented as median [interquartile range (IQR), IQR difference]. The IQR difference was calculated as the difference between the 75th percentile and the 25th percentile. Anteroposterior; C2, C4, and C6 indicate cervical vertebral levels. Age groups were defined as follows: Group 1, 3-6 years; Group 2, 7-12 years; Group 3, 13-17 years. Overall group comparisons were performed using the Kruskal-Wallis test, *: P-values indicate pairwise comparison results obtained using Dunn's post-hoc test with Bonferroni correction. A Bonferroni-adjusted p-value of <0.0167 was considered statistically significant.

J Acad Res Med 2026;16(2):98-105

Table 3. Results of correlation analysis of parameters with age

Parameter	Spearman r	p-value
C2 spinal cord diameter (mm)	0.48	0.001
C2 spinal canal diameter (mm)	0.09	0.089
C2 spinal cord area (mm^2)	0.68	0.001
C2 spinal canal area (mm^2)	0.16	0.004
C4 spinal cord diameter (mm)	0.37	0.001
C4 spinal canal diameter (mm)	0.22	0.001
C4 spinal cord area (mm^2)	0.68	0.001
C4 spinal canal area (mm^2)	0.26	0.001
C6 spinal cord diameter (mm)	0.04	0.34
C6 spinal canal diameter (mm)	0.19	0.001
C6 spinal cord area (mm^2)	0.44	0.001
C6 spinal canal area (mm^2)	0.12	0.022

Note: Spearman rank correlation analysis was used to evaluate the association between age and each morphometric parameter. Anteroposterior; C2, C4, and C6 indicate cervical vertebral levels. A p-value of <0.05 was considered statistically significant

positive correlation with spinal cord cross-sectional area at both C2 and C4 levels (Spearman r=0.68 for both parameters, p=0.001).

C6 spinal cord diameter showed no significant correlation with age (Spearman r=0.04, p=0.34). In contrast, the C6 spinal cord area showed a significant positive correlation with age. These findings further support the observation that age-related spinal cord development is more consistently reflected in the cross-sectional area than in the linear anteroposterior diameter alone.

Summary of Key Findings

Collectively, these analyses demonstrate that dimensions of the pediatric cervical spinal canal and spinal cord progressively enlarge with advancing age during childhood and adolescence. Furthermore, boys consistently exhibit larger morphometric measurements than girls across all examined cervical vertebral levels, confirming marked sexual dimorphism in the developing cervical spine.

DISCUSSION

Pediatric cervical spine and spinal cord morphometry carries substantial clinical significance for both diagnostic interpretation and surgical planning. These structures undergo considerable developmental change throughout childhood and adolescence, necessitating developmentally appropriate reference values.

This study provides contemporary, level-specific, MRI-based reference data for the pediatric cervical spinal canal and spinal cord at the C2, C4, and C6 levels, systematically stratified by age group and sex. Our dual-parameter approach, combining sagittal anteroposterior diameters and axial cross-sectional areas, was designed to characterize morphometry in an anatomically meaningful way and to be directly applicable to routine clinical MRI assessment.

This approach is clinically relevant because sagittal diameters remain standard in routine diagnostic reporting, while axial areas more accurately represent the space available for the spinal cord, particularly when canal or cord morphology deviates from simple circular or elliptical geometry (1,6).

A principal finding was the presence of consistent sexual dimorphism across all evaluated levels. Male subjects had larger spinal cord and spinal canal measurements than female subjects, and these differences were evident in both anteroposterior diameters and cross-sectional areas.

This pattern is consistent with previous MRI studies documenting sex-related differences in cervical canal dimensions in asymptomatic populations, suggesting that sex-related morphometric variation is readily detectable in the absence of overt pathology. Such dimorphism likely reflects broader sex-related differences in somatic growth, vertebral development, and overall body habitus during childhood and adolescence (4,5).

Importantly, our cohort was stratified into single-year age intervals with equal sex distribution, resulting in balanced age distributions for males and females. This design feature substantially strengthens the interpretability of sex-based comparisons by minimizing the confounding effect of age.

From a clinical perspective, these findings support the adoption of sex-specific reference expectations in pediatric reporting. This distinction is particularly important when interpreting borderline measurements or evaluating suspected stenosis, where modest absolute differences may substantially influence clinical assessment and patient management (1,6).

Clinically, this observation is also relevant to pediatric cervical instrumentation and fusion decision-making, which must account for ongoing growth and segment-specific anatomy, particularly at the upper cervical region and craniovertebral junction, where stabilization strategies are carefully tailored to unique biomechanical properties and developmental considerations (3,7).

Age-related morphometric changes showed distinct level-dependent patterns. Spinal cord cross-sectional area increased progressively with age at C2, C4, and C6, indicating continued development of spinal cord dimensions throughout childhood and adolescence. However, linear anteroposterior diameter measurements did not increase uniformly at all levels. For example, the C2 spinal cord anteroposterior diameter showed a significant age-related increase, whereas the C6 spinal cord anteroposterior diameter remained relatively stable across age groups.

This finding suggests that pediatric cervical cord development may involve shape-related or multidirectional remodeling that is not fully captured by a single linear diameter measurement. Therefore, cross-sectional area measurements may provide complementary information to sagittal anteroposterior diameters in pediatric cervical MRI interpretation. Beyond developmental assessment, comprehensive imaging evaluation is also clinically relevant in spinal trauma, where correlations

J Acad Res Med 2026;16(2):98-105

between clinical findings and radiological abnormalities have been reported (8).

Cervical morphometry demonstrates intrinsic level specificity. Measurements at C2 differ substantially from those at C4 and C6, consistent with known anatomical and biomechanical variation across the cervical spine. Morphometric analyses of osseous cervical structures and large-scale imaging studies of asymptomatic subjects similarly emphasize that canal and vertebral morphology vary by level and change systematically with age (9,10).

At C2, an age-related increase in spinal cord anteroposterior diameter suggests that upper cervical morphometry exhibits more prominent developmental scaling in the sagittal dimension. In contrast, lower cervical levels appear to demonstrate greater growth in other parameters, notably canal dimensions or cord cross-sectional area, rather than in cord anteroposterior diameter in isolation.

This divergence carries clinical implications: cord diameter and cord area do not necessarily change proportionally, and reliance on a single linear measurement may obscure relevant developmental remodeling when growth demonstrates anisotropic or shape-driven characteristics (11-13).

This principle is clinically important because pediatric cases are sometimes assessed using simplified single-threshold rules or reference standards derived from adult populations. However, stenosis risk assessment thresholds, such as measures of canal caliber or canal-to-body ratios, are sensitive to baseline anatomy and do not translate directly to pediatric populations without reference values anchored to age, sex, and level (1,6).

Our findings therefore support a more nuanced, developmentally informed interpretive approach. Pediatric cervical spine assessment benefits substantially from developmentally anchored reference values, and the diagnostic significance of any measurement should be contextualized within the child's age, sex, and the specific vertebral level being evaluated.

Providing both anteroposterior diameters and cross-sectional areas addresses an important clinical need. Linear anteroposterior diameter remains straightforward to obtain and is widely utilized in studies examining cervical canal dimensions and their association with degenerative or pathological changes (1,2). However, cross-sectional area may more directly reflect true available space, particularly when evaluating canal compromise or cord crowding, where asymmetric narrowing and shape changes can substantially reduce the available space without necessarily affecting anteroposterior diameter (11-13).

Moreover, MRI-based cord assessment is increasingly valued not solely for morphometry but also for tissue characterization and biomarker identification, underscoring the value of standardized, level-specific MRI measurements in advancing pediatric neuroimaging (12,13).

The reference morphometric data presented in this study may serve as practical benchmarks for pediatric cervical spine MRI interpretation, facilitate earlier recognition of clinically meaningful canal narrowing or cord compromise, and support developmentally informed surgical planning. Cross-sectional areas of the cord and canal may be particularly informative when cord crowding is suspected because area-based metrics capture morphological variation that a single anteroposterior diameter may not fully represent.

Nevertheless, strict numeric cut-offs warrant cautious application given inherent methodological variability across centers and imaging systems, including scanner differences, sequence parameters, and region-of-interest methodology, as well as the pronounced level- and age-dependence demonstrated in this analysis (11-13). The most appropriate use of these reference values is therefore as comparative ranges and benchmarks for clinical interpretation, rather than as absolute decision thresholds applied in isolation.

Study Limitations

Several limitations warrant explicit discussion. First, the retrospective single-center design may limit generalizability to other populations and imaging systems. Reference morphometry can vary substantially with imaging techniques and population characteristics; external validation across multiple centers and systems would strengthen applicability (5,10).

Second, although all included examinations demonstrated normal cervical spinal canal and spinal cord morphology, the cohort comprised children who underwent MRI for clinical indications rather than asymptomatic volunteers. Therefore, the term "normative" should be interpreted in the context of a clinically imaged pediatric population with normal MRI findings. Prospective studies involving asymptomatic community-based cohorts may provide further validation.

Third, anthropometric variables, including height, weight, and body mass index, were not incorporated into this analysis. These factors likely contribute significantly to inter-individual variability and may partially mediate the observed sex differences, as suggested by adult morphometric studies (14).

Fourth, measurements were performed manually. Although standardized manual delineation of regions of interest can yield robust results, this approach remains sensitive to boundary selection and choice of slice position. This limitation is particularly relevant when cord margins lack optimal conspicuity or when small variations in level selection influence area estimates.

Emerging automated and semi-automated processing pipelines, such as those provided by open-source spinal cord analysis frameworks, offer promising approaches to enhance reproducibility, accelerate throughput, and facilitate scalability for larger prospective reference datasets (15).

Multiple parameters were evaluated across several vertebral levels. No global multiplicity correction was applied across all statistical endpoints. Accordingly, findings, particularly those from exploratory comparisons, should be interpreted with appropriate caution regarding multiple comparisons.

J Acad Res Med 2026;16(2):98-105

CONCLUSION

This study provides comprehensive, level-specific, MRI-based morphometric reference data for the cervical spinal cord and spinal canal in children and adolescents with normal cervical MRI findings. Dimensions of the cervical cord and canal varied systematically with age, sex, and vertebral level. Boys had significantly larger cord and canal measurements than girls, and most parameters increased with age, although growth patterns differed among the C2, C4, and C6 levels. These reference values may assist radiologists and clinicians in interpreting pediatric cervical spine MRI examinations, in recognizing clinically relevant morphometric abnormalities, and in supporting developmentally informed surgical planning in pediatric cervical pathology.

Ethics

Ethics Committee Approval: The research protocol was approved by the Local Institutional Ethics Committee of Selçuk University (decision no: 2024/448, date: 18.09.2024).

Informed Consent: Given the retrospective study design, informed consent procedures were waived, and the study was conducted in accordance with institutional policies and ethics committee guidelines.

Acknowledgments

A generative AI language model, ChatGPT-5.2, was used to assist with grammar checking, to improve clarity and style, and to prepare a translation draft for this manuscript. The AI tool was used solely for editorial purposes; the study design, data analysis, and interpretation of results were conducted by the author(s).

Footnotes

Author Contributions: Surgical and Medical Practices - M.S.C., Z.B.; Concept - M.Ö., Z.B.; Design - M.S.C., M.Ö., Z.B.; Data Collection and/ or Processing - M.S.C., M.Ö., N.D.; Analysis and/ or Interpretation - M.Ö., N.D.; Literature Search - M.S.C., N.D.; Writing - M.Ö., N.D.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors report that no financial support was received for this study.

REFERENCES

1. Morishita Y, Naito M, Hymanson H, Miyazaki M, Wu G, Wang JC. The relationship between the cervical spinal canal diameter and the pathological changes in the cervical spine. Eur Spine J. 2009; 18: 877-83.
2. Kim KH, Park JY, Kuh SU, Chin DK, Kim KS, Cho YE. Changes in spinal canal diameter and vertebral body height with age. Yonsei Med J. 2013; 54: 1498-504.
3. Hwang SW, Gressot LV, Rangel-Castilla L, Whitehead WE, Curry DJ, Bollo RJ, et al. Outcomes of instrumented fusion in the pediatric cervical spine. J Neurosurg Spine. 2012; 17: 397-409.
4. Tatarek NE. Variation in the human cervical neural canal. Spine J. 2005; 5: 623-31.
5. Ulbrich EJ, Schraner C, Boesch C, Hodler J, Busato A, Anderson SE, et al. Normative MR cervical spinal canal dimensions. Radiology. 2014; 271: 172-82.
6. Tierney RT, Maldjian C, Mattacola CG, Straub SJ, Sitler MR. Cervical spine stenosis measures in normal subjects. J Athl Train. 2002; 37: 190-3.
7. Anderson RC, Ragel BT, Mocco J, Bohman LE, Brockmeyer DL. Selection of a rigid internal fixation construct for stabilization at the craniovertebral junction in pediatric patients. J Neurosurg. 2007; 107(1 Suppl): 36-42.
8. Badzhi LJ, Hardcastle TC, Naidoo P. Correlation of clinical and radiological findings in patients with spinal trauma at Inkosi Albert Luthuli Central Hospital. SA J Radiol. 2025; 29: 3248.
9. Abdullah KG, Steinmetz MP, Mroz TE. Morphometric and volumetric analysis of the lateral masses of the lower cervical spine. Spine (Phila Pa 1976). 2009; 34: 1476-9.
10. Kato F, Yukawa Y, Suda K, Yamagata M, Ueta T. Normal morphology, age-related changes and abnormal findings of the cervical spine. Part II: magnetic resonance imaging of over 1,200 asymptomatic subjects. Eur Spine J. 2012; 21: 1499-507.
11. Prasad SS, O'Malley M, Caplan M, Shackleford IM, Pydisetty RK. MRI measurements of the cervical spine and their correlation to Pavlov's ratio. Spine (Phila Pa 1976). 2003; 28: 1263-8.
12. Cadotte DW, Cadotte A, Cohen-Adad J, Fleet D, Livne M, Wilson JR, et al. Characterizing the location of spinal and vertebral levels in the human cervical spinal cord. AJNR Am J Neuroradiol. 2015; 36: 803-10.
13. Martin AR, De Leener B, Cohen-Adad J, Cadotte DW, Kalsi-Ryan S, Lange SF, et al. A novel MRI biomarker of spinal cord white matter injury: T2*-weighted white matter to gray matter signal intensity ratio. AJNR Am J Neuroradiol. 2017; 38: 1266-73.
14. Dağ N, Erdoğan Kaydu N, Cansız MS, Öztürk M. Morphometric analysis of cervical spinal cord and spinal canal with magnetic resonance imaging in Turkish adults. J Contemp Med. 2021; 11: 811-5.
15. De Leener B, Lévy S, Dupont SM, Fonov VS, Stikov N, Louis Collins D, et al. SCT: spinal cord toolbox, an open-source software for processing spinal cord MRI data. Neuroimage. 2017; 145: 24-43.

DOI: 10.4274/jarem.galenos.2026.81994
J Acad Res Med 2026;16(2):106-112

Therapeutic Plasma Exchange Procedures and Outcomes in a Pediatric Intensive Care Unit: A Single-center Experience

Esin Aydın Aksoy[1], Selman Kesici[2], Şule Ünal Cangül[3], Salih Aksu[4], Benan Bayrakçı[2]

[1]University of Health Sciences Türkiye, Şişli Hamidiye Etfal Training and Research Hospital, Department of Pediatrics, İstanbul, Türkiye
[2]Hacettepe University, Life Support Practice and Research Center, Department of Pediatric Critical Care Medicine, Ankara, Türkiye
[3]Hacettepe University Faculty of Medicine, Department of Pediatric Hematology, Ankara, Türkiye
[4]Hacettepe University Hospital, Blood Bank and Apheresis Unit, Department of Hematology, Ankara, Türkiye

Cite this article as: Aydın Aksoy E, Kesici S, Ünal Cangül Ş, Aksu S, Bayrakçı B. Therapeutic plasma exchange procedures and outcomes in a pediatric intensive care unit: a single-center experience. J Acad Res Med. 2026;16(2):106-112

ABSTRACT

Objective: Therapeutic plasma exchange (TPE) is an apheresis procedure in which the patient's plasma is removed and replaced with replacement fluid. We report an 11-year single-center experience with pediatric TPE in a pediatric intensive care unit (PICU).

Methods: This retrospective, single-center, observational study included 51 pediatric patients who underwent 248 TPE procedures in the PICU over an 11-year period. The indications for TPE were grouped as follows: 1. microangiopathic hemolytic anemia, including sepsis, disseminated intravascular coagulation (DIC), thrombocytopenia-associated multiorgan failure (TAMOF), hemolytic uremic syndrome (HUS), and hepatitis; 2. hemophagocytic lymphohistiocytosis (HLH) or macrophage activation syndrome (MAS); 3. intoxications; 4. neurological diseases. The clinical and laboratory data of these patients, including organ failure and scores on the pediatric logistic organ dysfunction (PELOD) and Glasgow coma scale, were recorded.

Results: The time for TPE after admission to PICU was 2.2±5.1 days. Twenty-eight (54.9%) patients were discharged, and 23 (45.1%) died. Minor complications that can be attributed to TPE were observed in 14.9% of the patients. Earlier initiation of TPE appeared to be associated with more favorable outcomes. A significant increase in platelet count was observed in Group 1, and favorable responses were also noted in selected intoxication cases. In contrast, no clear additional benefit was demonstrated in the HLH/MAS group.

Conclusion: TPE may be considered an option in pediatric patients with sepsis, DIC, TAMOF, HUS, and hepatitis (Group 1), as well as in selected intoxication cases. In Group 1, improvements in platelet counts and PELOD scores were the principal responses to treatment.

Keywords: Therapeutic plasma exchange, pediatric, critically ill children, intensive care unit

INTRODUCTION

"Apheresis" means "forced removal" in Greek. In many diseases, the problem is not only the accumulation of harmful substances but also the depletion of other substances. In such patients, removal and replacement of the patient's plasma may be beneficial. This procedure, called therapeutic plasma exchange (TPE), is non-selective and removes both normal and pathological plasma components, including antibodies, immune complexes, and cytokines (1). Although the removal of the pathologic substances is the major mechanism of action of TPE, various immunomodulatory effects of TPE, including the shift of the TH1/TH2 balance towards TH2 and the inhibition of interleukin-2 and interferon-γ production were reported (2).

The American Society for Apheresis (ASFA) published a categorized list of indications for therapeutic apheresis using an evidence-based approach (3). ASFA guidelines are the main reference for decision-making regarding indications in pediatric and adult patients. When taking indications in consideration, it can be seen that number of diseases which can be treated with TPE has increased significantly over the years (4).

Limited data are available on the safety and efficacy of TPE in pediatric patients. We aimed to present a single-center, 11-year

ORCID IDs of the authors: E.A.A. 0000-0002-2155-2897; S.K. 0000-0003-4746-6986; Ş.Ü.C. 0000-0002-3842-8788; S.A. 0000-0003-0634-3808; B.B. 0000-0003-3307-0948

Corresponding Author: Esin Aydın Aksoy, MD;
E-mail: dresinaaksoy@gmail.com
Received Date: 25.07.2025 **Accepted Date:** 22.07.2026
Epub: 30.07.2026
Publication Date: 07.08.2026

J Acad Res Med 2026;16(2):106-112

TPE experience in a pediatric intensive care unit (PICU), including indications, complications, outcomes of TPE use, and prognostic factors affecting TPE success.

METHODS

This study was a retrospective, single-center, observational study. The medical records of pediatric patients who underwent TPE in the PICU over an 11-year period were retrospectively reviewed. A total of 248 TPE procedures, performed in 51 patients, were included in the analysis. All pediatric patients who underwent at least one TPE procedure in the PICU during the study period were eligible for inclusion. Patients with insufficient medical records were excluded from the relevant analyses.

All TPE procedures were performed after written informed consent had been obtained from the parents or legal guardians.

The patients were categorized into four groups (Table 1):

Group 1: Microangiopathic hemolytic anemia group including patients with sepsis, disseminated intravascular coagulation (DIC), thrombocytopenia-associated multiorgan failure (TAMOF), HUS, and hepatitis (n=23).

Group 2: Patients with hemophagocytic lymphohistiocytosis (HLH)or macrophage activation syndrome (MAS), (n=11).

Group 3: Patients with intoxications at fatal doses with amitriptyline, carbamazepine, colchicine, mushrooms, or digoxin (n=12).

Group 4: Patients with neurological diseases, including Guillain-Barre syndrome, encephalitis, and Rasmussen encephalitis (n=5).

Complete blood count, serum liver and kidney function tests, and electrolytes were ordered before and after each TPE session for all patients.

The serum BUN and creatinine results from patients who underwent hemodialysis were excluded from the analyses. Serum ferritin, lactate dehydrogenase (LDH), triglyceride levels, aPTT, and PT were studied in Groups 1 and 2; thrombin time, plasma anti-thrombin III, and fibrinogen were studied in Groups 1-3. The clinical and laboratory data of these patients, including Pediatric Logistic Organ Dysfunction (PELOD) scores (5), organ failure scores [as recommended by Goldstein et al. (6)], and Glasgow coma scale (GCS) scores were noted (7).

Before the initiation of TPE, patients with hemoglobin levels <9 g/dL or platelet counts <50 x 10^9/L were transfused. Fresenius As. Tec 204™, COM TEC, or OPTIA devices were used for the TPE procedure. Veno-venous access was used, and either albumin or fresh-frozen plasma (FFP) was preferred as the replacement fluid. In each procedure, TPE was performed with an exchange of 1 to 1.5 plasma volumes.

The termination of TPE was decided by numerous clinical and laboratory criteria: in patients with TAMOF, TPE was ceased when organ failures remitted and platelet count >100 x10^9/L, in those with HLH, when serum LDH levels normalized and platelet count >100 x10^9/L, in patients with intoxication, when serum levels of the responsible agent decreased and intoxication signs and symptoms receded, and in those with Guillain-Barre syndrome, after the completion of five cycles.

Statistical Analyses

Statistical analyses were performed using SPSS for Windows, version 15.0 (SPSS Inc., Chicago, IL, USA). The normality of continuous variables was assessed using the Kolmogorov-Smirnov test. Continuous variables were summarized as mean ± standard deviation together with minimum and maximum values, which were considered more informative given the wide range of laboratory parameters observed in critically ill pediatric patients. Categorical variables were expressed as numbers and percentages.

Due to the non-normal distribution of several variables, paired comparisons of laboratory measurements before and after TPE [platelet count, albumin, aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), LDH, prothrombin time, international normalized ratio (INR), and serum electrolyte levels] were performed using the Wilcoxon signed-rank test. Changes in PELOD and organ failure scores before and after TPE were also analyzed using the Wilcoxon signed-rank test. Comparisons between categorical variables were performed using the chi-square test or Fisher's exact test, as appropriate.

Standardized mortality rates (SMRs) were calculated using the exact Poisson method, and the total SMR was 0.73 (95% confidence interval: 0.44-1.06). A two-sided p-value of <0.05 was considered statistically significant.

The study protocol was approved by the Institutional Ethics Committee of Hacettepe University (approval number: LUT 12/92-46, date: 11.09.2012). Written informed consent for the TPE procedure was obtained from the parents or legal guardians as part of routine clinical practice.

RESULTS

The mean age of the patients was 8±5.2 (0.1-16) years and 28 (54.9%) were female. A total of 248 procedures were performed on 51 patients. The underlying indications for TPE were summarized in Table 1. The mean number of TPE cycles performed on one patient was 4.9±4.6 (1-19). The mean time of one TPE cycle was 87.5±29 minutes (24-180). The overall mean time elapsed from admission to the PICU until the initiation of TPE was 2.2±5.1 days (1-32). For groups 1-4, the mean times to TPE initiation in Groups 1-4 were 2.3±6.5 days, 2.6±4.9 days, 0.2±0.4 days, and 5±3.8 days, respectively.

The devices used for TPE were COMTEC (53.8%), OPTIA (33.3%), and Fresenius AS 204 (12.9%). The central venous sites for TPE were the jugular vein (89.7%) and the femoral vein (10.3%). The diameter of the catheters ranged from 4 to 12 Fr.

The replacement fluid was FFP in 209 (90.5%) and 5% albumin in the remainder. FFP was used in all patients in Group 1, in 94.9% of those in Group 2, and in 93.3% of those in Group 3. All patients in Group 4 underwent TPE with 5% albumin. Citrate was used as an anticoagulant throughout the procedures, and

J Acad Res Med 2026;16(2):106-112

prophylactic calcium gluconate was administered to all patients during TPE.

All patients' GCSs increased significantly after TPE compared with before TPE (p=0.038). The post-TPE values in Group 3 patients were significantly higher than those in the other groups (p=0.005). Among the survivors, the difference in pre- and post-TPE GCS values of Group 1 was significantly higher than in the other groups (p=0.043).

Of the 51 patients, 23 (45.1%) died and 28 (54.9%) were discharged. The mortality rates in Groups 1, 2, 3, and 4 were 60%, 63%, 8%, and 20%, respectively (Table 1). The overall mean time until death after admission to the PICU was 27.5±40.3 days (1-186). Invasive mechanical ventilation was used in 35 patients, of whom 22 (62.8%) died before being weaned from the ventilator. The remaining 13 patients (37.2%) were extubated. One of the extubated patients died on the 10th post-extubation day. Seventy-five percent of patients in Group 1 were successfully extubated within 48 hours (early), while all patients in Group 3 were off mechanical ventilation within 24 hours (very early). In Groups 2 and 4, early extubation was not possible.

In Group 1, the survival rates of patients who received the first TPE session within 24 hours (n=10), within 25-48 hours (n=5), and within 49-72 hours (n=6) of admission were 40%, 40%, and 50%, respectively. Those in Group 1 who had their first TPE after the 3rd day of admission (n=2) died.

The SMRs were calculated according to the PRISM scores and summarized in Table 2.

Table 1. Clinical characteristics of the patients

	Indication for TPE	Number of cycles/ patient mean ± SD	Deceased patients n (%)	Pre-TPE PELOD score mean ± SD median (25th-75th percentile)	Post-TPE PELOD score mean ± SD median (25th-75th percentile)	P-value for the difference of pre-TPE, post-TPE PELOD scores*	Pre-TPE organ failure score mean ± SD median (25th-75th percentile)	Post-TPE organ failure score mean ± SD median (25th-75th percentile)	P-value for the difference of pre-TPE, post-TPE organ failure scores*
Group 1 (n=23)	-Sepsis and TAMOF (n=10) -HUS (n=7) -DIC (n=4) -Hepatitis (n=2)	6.4 (1-19)	14 (60)	All groups 14.1±11.2 12 (10-21)	All groups 12.6±13.6 11 (1-23)	All groups p=0.234	All groups 3.7 (±1.5) 4 (2-5)	All groups 3.0 (±2.0) 3 (1-5)	All groups p=0.035
				Among survivors 9.1±7.0 10 (1.5-12)	Among survivors 2.9±4.4 1 (0-6)	Among survivors p=0.007	Among survivors 2.8±1.6 2 (2-4.5)	Among survivors 1.0±1.4 0 (0-2)	Among survivors p=0.020
Group 2 (n=11)	-MAS (n=7) [secondary to JRA (n=5) or SLE (n=2) -HLH (n=4)] [secondary to infections (n=2) or propionic acidemia (n=2)]	5.4 (1-13)	7 (63)	All groups 8.6±8.9 10 (2-12)	All groups 11.7±8.4 11 (1-23)	All groups p=0.239	All groups 2.7±1.2 2 (2-3)	All groups 3.3±1.2 3 (2-4)	All groups p=0.096
				Among survivors 4.0±6.0 1.5 (0.3-10.3)	Among survivors 2.5±0.6 2.5 (2-3)	Among survivors p=0.713	Among survivors 2.2±0.9 2.5 (1.3-3)	Among survivors 2.5±0.6 2.5 (2-3)	Among survivors p=0.317
Group 3 (n=12)	-Amitriptyline (n=5) -Carbamazepine (n=2) -Colchicine (n=2) -Mushroom (n=2) -Digoxin (n=1)	1.8 (1-5)	1 (8)	All groups 9.6±9.9 10 (1-17.8)	All groups 0.6±0.9 0 (0-1)	All groups p=0.017	All groups 1.5±1.0 2 (0.3-2)	All groups 0.4±0.6 0 (0-1)	All groups p=0.016
Group 4 (n=5)	-Guillain-Barre syndrome (n=3) -Encephalitis (n=1) -Rasmussen encephalitis (n=1)	3.6 (1-5)	1 (20)	All groups 9.0±17.8 1 (1-21)	All groups 8.8±18 1 (0.5-21)	All groups p=0.317	All groups 1.4±0.8 1 (1-2)	All groups 0.8±1.7 0 (0-2)	All groups p=0.180
Total (n=51)		4.9 (1-19)	23 (45)	All groups 11.3±11.2 10 (1-13)	All groups 9.2±12.1 2 (1-13)	All groups p=0.093	All groups 2.7±1.6 3 (2-4)	All groups 2.2±2.0 2 (0-4)	All groups p=0.012
				Among survivors 7.75±8.26 6 (1-11.8)	Among survivors 1.57±2.68 1 (0-2)	Among survivors p=0.001	Among survivors 1.89±1.34 2 (1-2.8)	Among survivors 0.82±1.18 0 (0-2)	Among survivors p=0.001

*: Wilcoxon test, TPE: Therapeutic plasma exchange, SD: Standard deviation, TAMOF: Thrombocytopenia-associated multiorgan failure, HUS: Hemolytic uremic syndrome, DIC: Disseminated intravascular coagulation, MAS: Macrophage activation syndrome, HLH: Hemophagocytic lymphohistiocytosis

J Acad Res Med 2026;16(2):106-112

Only Group 1 showed a significant increase in mean platelet count after TPE procedures compared to pre-TPE values (90435 vs. 101478 K/uL, p=0.011). The serum albumin levels significantly increased among the survivors of Group 1 after TPE (p=0.025) 3.01 vs. 3.78 g/L). The decrease in serum AST levels after TPE was associated with better survival in all patients and in Group 1 (767 vs. 152 IU/L, p=0.002 in all patients, and 442 vs. 53 IU/L, p=0.008 in Group 1); the same was observed for decreases in LDH (2149 vs. 852 IU/L, p=0.002 in all surviving patients, and 2049 vs. 934 IU/L, p=0.012 in Group 1) and in PELOD scores (7.75 vs. 1.57, p=0.001 in all patients, and 9.11 vs. 2.89, p=0.007 in survivors of Group 1). Serum ALT and ALP levels, as well as prothrombin time and INR, decreased after TPE in all patients, but none of these reductions were associated with improved survival rates. The serum electrolyte levels did not change significantly with TPE treatment in any of the patients (Table 3).

The complications attributable to TPE included rash (4.9%), electrolyte imbalances (4.4%), including hypocalcemia, hypokalemia, and hypernatremia, nausea and vomiting (1.2%), fever (1.2%), hypotension (1.2%), altered consciousness (0.8%), hypertension (0.4%), epistaxis (0.4%), and catheter occlusion (0.4%).

DISCUSSION

Indications for TPE in pediatric patients are often based on adult data, and the ASFA guidelines do not include age-specific recommendations regarding indications. Rather, they are categorized according to strength, ranging from 1 to 4 (3). Category I disorders are those in which TPE may be used as first-line therapy, either alone or in conjunction with other therapeutic options, such as thrombotic thrombocytopenic purpura (TTP), dialysis-independent or alveolar hemorrhage-associated Wegener granulomatosis or Goodpasture syndrome, recurrent focal segmental glomerulosclerosis, Guillain-Barre syndrome, chronic inflammatory demyelinating polyradiculopathy, cryoglobulinemia, hemolytic uremic syndrome related to autoantibody to factor H, pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections and antibody-mediated rejection after renal transplantation (8). Many other diseases may benefit from TPE (8,9). Although peripheral access may be used in older children, central venous access is usually required. The major risks of TPE are related to catheters, including infections, bleeding, thrombosis, pneumothorax, hemothorax, or cardiac arrhythmias (10). A common problem during TPE is paresthesia due to hypocalcemia caused by citrate chelation. Citrate is used as an anticoagulant to prevent the blood from clotting in the

Table 2. The SMR of patients

	Observed mortality (n)	Expected mortality (n)	SMR (95% CI)
Group 1	14	18	0.78
Group 2	7	7	1
Group 3	1	6	0.17
Group 4	1	2	0.5
Total	23	33	0.73 (0.44-1.06)

SMR: Standardized mortality rates, CI: Confidence interval

Table 3. Changes in laboratory parameters before and after TPE

	Pre-TPE mean ± SD (min-max)	Post-TPE mean ± SD (min-max)	p-value*
AST (U/L)	1828±5510 (14-28831)	584±1866 (11-12166)	0.001
ALT (U/L)	823±2254 (8-9960)	234±588 (5-3369)	0.003
ALP (U/L)	164±131 (23-686)	103±88 (18-615)	<0.001
LDH (U/L)	2579±2763 (98-12000)	1988±3957 (425-21626)	0.007
aPTT (s)	36±10 (17.6-63.0)	32±12 (1.1-82.1)	0.011
INR (-)	2.12±1.60 (1.02-9.31)	1.50±0.76 (0.88-5.31)	0.001
Sodium (mmol/L)	137.5±6.0 (126-158)	140.4±5.5 (128-156)	0.055
Potassium (mmol/L)	3.99±0.81 (2.20-5.70)	3.66±0.91 (2.02-6.45)	0.459
Calcium, total (mg/dL)	8.58±1.55 (4.20-13.20)	9.36±1.45 (6.20-13.90)	0.258
Calcium, ionized (mmol/L)	1.06±0.22 (0.16-1.43)	1.12±0.17 (0.69-1.47)	0.414
Phosphorus (mg/dL)	4.40±1.80 (0.35-10.86)	4.27±1.69 (1.39-10.67)	0.903
Glucose (mg/dL)	133±85 (48-601)	136±60 (53-334)	0.453

*: Paired comparisons were performed using the Wilcoxon signed-rank test, TPE: Therapeutic plasma exchange, AST: Aspartate aminotransferase, ALT: Alanine aminotransferase, ALP: Alkaline phosphatase, LDH: Lactate dehydrogenase, aPTT: Activated partial thromboplastin time, INR: International normalized ratio, SD: Standard deviation

J Acad Res Med 2026;16(2):106-112

TPE device and the risk for hypocalcemia is higher in patients with underlying renal or hepatic impairment (11,12). This risk is eliminated if heparin is used instead of citrate. Another common complication during TPE is hypotension, which may be related to hypocalcemia or hypovolemia (10). The coagulation factors decrease by one volume of plasma exchange and factors VIII, IX, and vWF return to normal within 4 hours, whereas the remaining coagulation factors achieve pre-TPE levels by the 24th hour, excluding fibrinogen, which takes 72 hours to return to normal levels after one TPE session (11). Anti-coagulation factors also decrease after TPE, including anti-thrombin III, and there is data on the increased risk of thrombosis after this procedure (11). A decrease in pseudocholinesterase levels after TPE may prolong neuromuscular junction blockage (13).

The replacement fluids include either albumin in physiologic saline or plasma. Plasma poses a risk of transfusion reactions or transfusion-related infections. However, it is the only choice for patients who undergo TPE due to TTP or DIC (11,14).

Previous studies have reported that earlier initiation of TPE is associated with better survival outcomes. Qu et al. (15) analyzed the TPE results in patients with fulminant sepsis aged 8 months to 14 years and reported that all patients who received TPE within 2-10 hours survived. They also reviewed 10 studies including 60 patients with fulminant sepsis and noted that delayed TPE initiation was related to increased mortality (15). Chong et al. (16) analyzed children with critical heart diseases and a clinical diagnosis of TAMOF and reported that patients who received TPE early (within <1 day of ECMO cannulation) had significantly improved modified organ failure index scores and platelet counts compared to patients with late TPE initiation. In the study of Kawai et al. (17) on 14 children with sepsis-induced TAMOF, initiating TPE early (within <30 hours from admission) was associated with greater improvement in organ dysfunction and decreased requirement for vasoactive and/or inotropic agents. In our study, the mean time to initiation of TPE was 2.2 days. In Group 1, earlier TPE initiation appeared to be associated with more favorable outcomes; however, this observation should be interpreted cautiously because of the small sample size. In our patients with MAS and HLH, the timing of TPE was not correlated with improved survival; however, this may be attributed to the small sample size. Group 3 included patients with intoxications; all underwent TPE within 24 hours, and only one patient died. No clear association between earlier initiation and improved outcomes was observed in Group 4.

Among all our patients, there were adverse events that may be attributed to TPE in 14.9%, of which rash and electrolyte imbalance were the most common, similar to the literature (18-20). No TPE-related mortality occurred. Michon et al. (21) examined 1632 apheresis procedures performed on 186 pediatric patients and observed adverse events in 55% of procedures and 82% of patients. Hypotension and symptomatic hypocalcemia were the leading adverse events; two patients died from TPE complications. The risk factors associated with complications were low body weight, low hemoglobin level, apheresis in the PICU, and an increased number of procedures per patient (21). In a similar study, 4.3% serious adverse effects had been observed (22). In a study by Sık et al. (4), minor adverse effects were seen in 16.5% of the sessions and no patients died because of TPE. They stated that circuit clotting and access malfunction were the most frequent adverse events. In a study conducted among European pediatric nephrology units, the rate of minor adverse effects observed with plasma exchange was 6.9%, hypotension, and premature disconnection being the most common (23).

In our cohort, 35 patients required ventilation support, of whom 22 (62.8%) died before extubation. Patients in Groups 1 and 3 had the shortest intubation periods. Similarly, in a study of 48 critically ill children, Cortina et al. (24) stated that the rate of ventilated patients among the survivors and non-survivors were 42.5%, and 75%, respectively.

In a prospective study on thrombocytopenia-associated multiple organ failure, Fortenberry et al. (25) showed that 28-day all-cause mortality was lower in children treated with TPE compared with those who were not. A retrospective study conducted in 43 hospitals over 9 years reported mortality rates of 22% among TPE patients without MODS and 44.4% among those with MODS. Children receiving TPE had higher mortality because of greater comorbidities and MODS (26). In our current study, the mortality rate among patients with sepsis and TAMOF was 60,% with an SMR of 0.78. An SMR of <1 indicates lower-than-expected mortality among Group 1 patients who underwent TPE. In the HLH/MAS group, observed mortality was similar to expected mortality, and no clear survival advantage associated with TPE could be demonstrated. In a previous study from our center, we reported lifesaving TPE applications in three severe amitriptyline intoxication cases (27). Group 3 had the lowest SMR scores, and only 1 out of 12 patients (8%) who was intoxicated with a lethal dose of colchicine (>0.8 mg/kg) died. The greatest benefit of TPE for survival was observed in this group. TPE has limited benefit in colchicine intoxication because of its large volume of distribution.

The GCSs of patients in Group 3 significantly increased after TPE, indicating rapid toxin removal that led to resolution of the comatose state. PELOD scores significantly decreased after TPE in all survivors, with a larger decrease observed in Group 1. It is reasonable to observe the greatest improvement in PELOD scores among Group 1 patients, since this group also included patients who had TAMOF. The decrease in PELOD scores after TPE in survivors indicated a beneficial effect of TPE on the progression of organ failure. Similar to PELOD scores, the scoring by Goldstein et al. (6) revealed a decrease in evidence of organ failure after TPE.

Among survivors in all groups, platelet counts increased post-TPE. The significant increase in platelet counts with TPE in Group 1 patients may serve as an indicator of treatment response. Chong et al. (16) had already reported a similar trend of increase in platelet counts with TPE in critically ill cardiac children with TAMOF.

J Acad Res Med 2026;16(2):106-112

Study Limitations

One of the major strengths of this study is its reliance on a large number of TPE procedures performed in a single PICU over an 11-year period. The inclusion of different disease groups allowed for a broader perspective on the effects of TPE in various clinical presentations. The comparative analysis of clinical parameters [e.g., pre- and post-TPE organ failure scores (PELOD) and GCS] and biochemical parameters (laboratory values) offers a significant methodological advantage. Furthermore, standardization of the devices used in TPE (COMTEC, OPTIA, Fresenius AS 204) is valuable for ensuring data consistency.

However, the study has several limitations. First, its retrospective design precludes establishing a cause-and-effect relationship. Furthermore, changes in clinical practices and treatment protocols over the long period of data collection may have influenced the results. The small number of patients in some subgroups (especially Groups 2 and 4) limits the generalizability of findings. In addition, the relatively small sample size limited the study's statistical power and precluded the use of advanced modeling approaches, such as logistic or Cox regression analyses.

CONCLUSION

The clinical and laboratory responses to TPE varied by underlying indication. The most favorable responses were observed in selected patients in Group 1 and in some cases of intoxication. Among survivors, increased platelet counts and improvements in PELOD scores were the main indicators of treatment response in Group 1. Earlier initiation may be clinically important, although definitive conclusions regarding treatment timing cannot be drawn because of the retrospective design and limited sample size. No clear additional survival benefit was demonstrated in the HLH/MAS group. Prospective multicenter studies are needed to clarify patient selection, optimal timing, and clinical effectiveness.

Ethics

Ethics Committee Approval: The study protocol was approved by the Institutional Ethics Committee of Hacettepe University (approval number: LUT 12/92-46, date: 11.09.2012).

Informed Consent: Written informed consent for the TPE procedure was obtained from the parents or legal guardians as part of routine clinical practice.

Footnotes

Author Contributions: Surgical and Medical Practices - E.A.A., S.K., Ş.Ü.C., S.A., B.B.; Concept - E.A.A.; Design - E.A.A., B.B.; Data Collection and/or Processing - E.A.A.; Analysis and/or Interpretation - E.A.A., B.B.; Literature Search - E.A.A.; Writing - E.A.A.

Conflict of Interest: The authors have no conflict of interest to declare.

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

REFERENCES

1. Nguyen TC, Kiss JE, Goldman JR, Carcillo JA. The role of plasmapheresis in critical illness. Crit Care Clin. 2012; 28: 453-68.
2. Goto H, Matsuo H, Nakane S, Izumoto H, Fukudome T, Kambara C, et al. Plasmapheresis affects T helper type-1/T helper type-2 balance of circulating peripheral lymphocytes. Ther Apher. 2001; 5: 494-6.
3. Connelly-Smith L, Alquist CR, Aqui NA, Hofmann JC, Klingel R, Onwuemene OA, et al. Guidelines on the use of therapeutic apheresis in clinical practice - evidence-based approach from the writing committee of the American Society for Apheresis: the ninth special issue. J Clin Apher. 2023; 38: 77-278.
4. Sık G, Demirbuga A, Annayev A, Akcay A, Çıtak A, Öztürk G. Therapeutic plasma exchange in pediatric intensive care: indications, results and complications. Ther Apher Dial. 2020; 24: 221-9.
5. Leteurtre S, Duhamel A, Grandbastien B, Lacroix J, Leclerc F. Paediatric logistic organ dysfunction (PELOD) score. Lancet. 2006; 367: 897; author reply 900-2. Erratum for: Lancet. 2003; 362: 192-7.
6. Goldstein B, Giroir B, Randolph A; International Consensus Conference on Pediatric Sepsis. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. Pediatr Crit Care Med. 2005; 6: 2-8.
7. Miranda M, Nadel S. Pediatric sepsis: a summary of current definitions and management recommendations. Curr Pediatr Rep. 2023; 11: 29-39.
8. Szczepiorkowski ZM, Winters JL, Bandarenko N, Kim HC, Linenberger ML, Marques MB, et al.; Apheresis Applications Committee of the American Society for Apheresis. Guidelines on the use of therapeutic apheresis in clinical practice--evidence-based approach from the Apheresis Applications Committee of the American Society for Apheresis. J Clin Apher. 2010; 25: 83-177.
9. Kintzel PE, Eastlund T, Calis KA. Extracorporeal removal of antimicrobials during plasmapheresis. J Clin Apher. 2003; 18: 194-205.
10. Kim YA, Sloan SR. Pediatric therapeutic apheresis: rationale and indications for plasmapheresis, cytapheresis, extracorporeal photopheresis, and LDL apheresis. Pediatr Clin North Am. 2013; 60: 1569-80.
11. Winters JL. Plasma exchange: concepts, mechanisms, and an overview of the American Society for Apheresis guidelines. Hematology Am Soc Hematol Educ Program. 2012; 2012: 7-12.
12. Bolan CD, Yau YY, Cullis HC, Horwitz ME, Mackall CL, Barrett AJ, et al. Pediatric large-volume leukapheresis: a single institution experience with heparin versus citrate-based anticoagulant regimens. Transfusion. 2004; 44: 229-38.
13. Naik B, Hirshhorn S, Dharnidharka VR. Prolonged neuromuscular block due to cholinesterase depletion by plasmapheresis. J Clin Anesth. 2002; 14: 381-4.
14. McLeod BC. Therapeutic apheresis: use of human serum albumin, fresh frozen plasma and cryosupernatant plasma in therapeutic plasma exchange. Best Pract Res Clin Haematol. 2006; 19: 157-67.
15. Qu L, Kiss JE, Dargo G, Carcillo JA. Outcomes of previously healthy pediatric patients with fulminant sepsis-induced multisystem organ failure receiving therapeutic plasma exchange. J Clin Apher. 2011; 26: 208-13.
16. Chong M, Lopez-Magallon AJ, Saenz L, Sharma MS, Althouse AD, Morell VO, et al. Use of therapeutic plasma exchange during extracorporeal life support in critically Ill cardiac children with thrombocytopenia-associated multi-organ failure. Front Pediatr. 2017; 5: 254.
17. Kawai Y, Cornell TT, Cooley EG, Beckman CN, Baldridge PK, Mottes TA, et al. Therapeutic plasma exchange may improve hemodynamics and organ failure among children with sepsis-induced multiple organ dysfunction syndrome receiving extracorporeal life support. Pediatr Crit Care Med. 2015; 16: 366-74.
18. Fateen T, Sultana N, Sarwar M, Saqlain N. Complications of therapeutic plasma exchange in pediatric patients: an experience at a tertiary care hospital. Pak J Med Sci. 2023; 39: 994-8.
19. Dalkiran T, Mercan M, Ipek S, Güllü UU, Kandur Y, Acipayam C, Dilber C. Therapeutic plasma exchange in pediatric patients: results from a single center. J Pediatr Intensive Care. 2022; 13: 282-5.
20. Aplenc LM, Wood N, Iqbal NT, George C, Burroughsscanlon C. Therapeutic plasma exchange in tandem with other types of extracorporeal circuits: the experience of a pediatric center and a review of other pediatric center reports. J Clin Apher. 2025; 40: e70007.
21. Michon B, Moghrabi A, Winikoff R, Barrette S, Bernstein ML, Champagne J, et al. Complications of apheresis in children. Transfusion. 2007; 47: 1837-42.

J Acad Res Med 2026;16(2):106-112

22. Balasubramanian KK, Venkatachalapathy P, Margabandhu S, Natraj R, Sridaran VK, Lakshmanan C, et al. Scope, safety, and feasibility of therapeutic plasma exchange in pediatric intensive care unit: a single-center experience. Indian J Crit Care Med. 2023; 27: 766-70.
23. Paglialonga F, Schmitt CP, Shroff R, Vondrak K, Aufricht C, Watson AR, et al. Indications, technique, and outcome of therapeutic apheresis in European pediatric nephrology units. Pediatr Nephrol. 2015; 30: 103-11.
24. Cortina G, McRae R, Chiletti R, Butt W. Therapeutic plasma exchange in critically Ill children requiring intensive care. Pediatric Critical Care Medicine. 2018; 19: e97-104.
25. Fortenberry JD, Nguyen T, Grunwell JR, Aneja RK, Wheeler D, Hall M, et al.; Thrombocytopenia-associated multiple organ failure (TAMOF) network study group. Therapeutic plasma exchange in children with thrombocytopenia-associated multiple organ failure: the thrombocytopenia-associated multiple organ failure network prospective experience. Crit Care Med. 2019; 47: e173-81.
26. Lima LM, McCracken CE, Fortenberry JD, Hebbar KB. Use of plasma exchange in pediatric severe sepsis in children's hospitals. J Crit Care. 2018; 45: 114-20.
27. Tekerek NÜ, Dursun A, Akyildiz BN. Çocuk yoğun bakım ünitesinde takip edilen zehirlenme olgularının geriye dönük değerlendirilmesi [Retrospective analysis of poisoning cases our followed in pediatric intensive care unit]. J Pediatr Emerg Intensive Care Med. 2016; 3: 21-6. Turkish.

DOI: 10.4274/jarem.galenos.2026.27870
J Acad Res Med 2026;16(2):113-120

Second Victim Phenomenon and Support Strategies Among Orthopedic Surgeons in Türkiye: A Cross-sectional Study on Mental Health Symptoms

Yağmur Sever Fidan[1], Arda Kızılsert[2], Sümeyye Yasemin Çallı[1], İkbal Hümay Arman[3], Fırat Fidan[4]

[1]University of Health Sciences Türkiye, Erenköy Mental and Nervous Diseases Training and Research Hospital, Department of Psychiatry, İstanbul, Türkiye
[2]University of Health Sciences Türkiye, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, Department of Psychiatry, İstanbul, Türkiye
[3]İstanbul Medipol University, Department of Family Medicine, İstanbul, Türkiye
[4]İstanbul Aydın University, Department of Orthopedics and Traumatology, İstanbul, Türkiye

Cite this article as: Fidan YS, Y, Kızılsert A, Çallı SY, Arman İH, Fidan F. Second victim phenomenon and support strategies among orthopedic surgeons in Türkiye: a cross-sectional study on mental health symptoms. J Acad Res Med. 2026;16(2):113-120

ABSTRACT

Objective: The second victim phenomenon describes the psychological and professional distress experienced by healthcare workers following adverse events or complications. Despite growing international recognition, empirical data on this phenomenon among surgical specialties remain scarce. This study examined the frequency of second victim experiences and related exposures among orthopedic and traumatology physicians in Türkiye, assessed available support systems, and explored associations with depression, anxiety, and stress symptoms.

Methods: This descriptive cross-sectional study was conducted between May and July 2025 involving 121 orthopedic surgeons and residents in Türkiye. Data were collected through an online questionnaire comprising a sociodemographic form, the Turkish version of the Second Victim Experience and Support Tool (T-SVEST), and the Depression Anxiety Stress Scale-21 (DASS-21). Analyses employed non-parametric methods, including Spearman's correlation and Kruskal-Wallis tests, with Bonferroni correction.

Results: Over three-quarters of participants reported exposure to at least one adverse event that could potentially trigger second victimization, and nearly 90% had no prior awareness of the concept. Approximately one-quarter met the predefined T-SVEST threshold for psychological distress, and one-fifth met the threshold for physical distress. DASS-21 scores showed strong positive correlations with T-SVEST subscales, particularly psychological distress, physical distress, loss of professional self-efficacy, and turnover intention. Turnover intention was higher among physicians with more than 15 years of experience, whereas associate professors showed greater physical distress and higher absenteeism compared with residents. A considerable proportion of participants perceived institutional support as inadequate.

Conclusion: Second victimization among orthopedic surgeons in Türkiye has multidimensional effects, encompassing psychological and physical symptoms, reduced professional self-efficacy, turnover intentions, and absenteeism. These findings highlight the insufficiency of existing support mechanisms and the urgent need for preventive, systematic institutional strategies to protect physician well-being and ensure workforce sustainability.

Keywords: Second victim phenomenon, orthopedic surgery, psychological distress, depression, healthcare workers

INTRODUCTION

The concept of the "second victim" was first described by Wu (1). In the context of medical errors, complications, or patient safety incidents, the first victim is the patient, whereas the second victim refers to the healthcare professional—such as a physician, nurse, pharmacist, etc.—who may have been involved in the event and subsequently suffers negative consequences (2,3).

Often, these professionals perceive themselves as personally accountable for the unexpected patient outcome, leading to self-doubt regarding their clinical skills and knowledge (4,5). Medical errors or adverse events can therefore have profound and lasting psychological impacts on healthcare workers (6). Reported symptoms among second victims include intrusive recollections, anxiety, self-directed anger, feelings of remorse, emotional distress, fear of repeating mistakes, shame, guilt, and disturbances in sleep (7). If such reactions are not identified promptly and sufficient support systems are not provided, they may progress into more serious mental health conditions, including depression, anxiety disorders, or post-traumatic stress disorder. Although

ORCID IDs of the authors: Y.S.F. 0000-0002-7507-2290; A.K. 0000-0001-7167-5740; S.Y.Ç. 0000-0002-7037-0782; İ.H.A. 0000-0003-1492-6763; F.F. 0000-0003-3490-7340

Corresponding Author: Yağmur Sever Fidan, MD;
E-mail: dr.yagmurfidan@gmail.com
Received Date: 20.06.2026 **Accepted Date:** 02.08.2026
Publication Date: 07.08.2026

second victim reactions may partially mirror PTSD-like symptoms, the phenomenon is conceptually distinct: it originates from clinical accountability, moral distress, and professional consequences that follow adverse events, complications, or unexpected patient outcomes. Rather than being tied to a single traumatic episode, second victimization develops within ongoing medical practice and therefore represents a broader, profession-bound construct (3,7). A related but conceptually distinct construct is "moral injury", referring to the psychological distress that arises when clinicians perpetrate, fail to prevent, or witness acts that transgress their deeply held moral beliefs, often due to systemic constraints rather than a single adverse event (8). Unlike second victim reactions, which center on a discrete triggering incident, moral injury reflects a broader, cumulative toll of being unable to act according to one's professional values.

Recognition of the second-victim phenomenon allows healthcare workers' emotional burdens to be acknowledged and their need for psychological support to be addressed. In "punitive" clinical environments with insufficient support structures, healthcare workers often feel apprehensive about reporting medical errors, which in turn negatively impacts the prevention of future errors (9). Studies have shown that many healthcare professionals face difficulties in obtaining support after a medical error (10,11).

In the literature, some studies have reported that surgeons, due to the nature of their work in which the link between technical skills and patient outcomes is more pronounced, experience deeper reactions as second victims (12,13). Orthopedics and traumatology is one of the surgical specialties in which malpractice claims are frequently encountered (14). This study aimed to investigate the second victim phenomenon among orthopedic and traumatology specialists in Türkiye, the support systems they have to cope with it, and the depression, anxiety, and stress symptoms that may result from being a second victim. This study represents the first empirical investigation of the second victim phenomenon among orthopedic surgeons in Türkiye. The type of situation leading to the second-victim phenomenon (medical error, adverse patient-safety event, or complication) was not an objective of this study.

METHODS

Study Design

This cross-sectional, descriptive study was designed to investigate the frequency of second victim experiences and exposure to potentially triggering events, the support strategies utilized, and the relationship between second victim experiences and mental health symptoms such as depression, anxiety, and stress among orthopedic physicians in Türkiye.

Participants were recruited from among orthopedic surgeons and residents in Türkiye. Data were collected between 23 May 2025 and 23 July 2025 through an online questionnaire designed using Google Forms. The survey link was disseminated via professional communication channels, including orthopedic societies, academic and clinical WhatsApp groups, and e-mail networks. The invitation message clearly stated that participation was voluntary and informed consent was required. The purpose of the study was explained, and participants had the option to decline. Confidentiality and anonymity were maintained throughout the process, and no identifying personal information was collected. Participants received no financial compensation for their participation.

Sample Size Calculation

The primary sample size calculation was based on the study's primary objective: to examine associations between scores on the Turkish version of the Second Victim Experience and Support Tool (T-SVEST) and scores on the Depression Anxiety Stress Scale-21 (DASS-21). An a priori power analysis was performed using G*Power version 3.1.9.7 (Heinrich-Heine-Universität Düsseldorf, Germany). For a two-tailed correlation analysis, assuming a medium effect size (r=0.30), an alpha level of 0.05, and a statistical power of 80%, a minimum of 84 participants was required. As no previous study provided a directly transferable effect size estimate for the primary association examined in the present study, a conventional medium effect size was adopted. To account for potential exclusions and incomplete data, the target sample size was increased to at least 95 participants. A total of 126 orthopedic physicians completed the questionnaire; five participants who reported receiving psychiatric treatment were excluded before analysis, resulting in a final sample of 121. Comparisons according to sociodemographic and occupational characteristics were considered secondary analyses.

The study population comprised orthopedic surgeons and residents working in public, university, or private hospitals in Türkiye. Inclusion criteria were: age 18-65, active or past practice in orthopedics and traumatology, and provision of informed consent. Incomplete responses, entries from other specialties, and participants reporting psychiatric treatment were excluded. Participants who were receiving psychiatric treatment were excluded to minimize the confounding effect of pre-existing psychiatric conditions on assessed distress symptoms.

The study was approved by the Ethics Committee of İstanbul Medipol University (decision no: 606, date: 22 May 2025) and conducted in accordance with the Declaration of Helsinki.

Data Collection Tools

Data were collected using a structured online questionnaire comprising three sections:

Sociodemographic and Clinical Form: Developed for this study, including age, gender, education, years of experience, academic title, history of psychiatric diagnosis and treatment. In addition, participants were asked two questions: "Have you previously been involved in an adverse patient safety event?" (defined as harm, or the risk of harm, to a patient arising from the healthcare provided rather than from the patient's underlying clinical condition) and "Have you previously heard of the 'second victim phenomenon?"

T-SVEST: The SVEST, originally developed by Burlison et al. (15), was designed to facilitate the implementation and evaluation

J Acad Res Med 2026;16(2):113-120

of support resources for second victims within healthcare organizations. The Turkish version comprises 24 items across nine domains: psychological distress (4 items), physical distress (4 items), colleague support (2 items), supervisor support (4 items), institutional support (2 items), non-work-related support (2 items), professional self-efficacy (2 items), turnover intention (2 items), and absenteeism (2 items). Additionally, seven items assess healthcare workers' preferences regarding potential support resources, yielding a total of 31 items. All items are rated on a five-point Likert scale, and the scoring and reverse-coding procedures are described in the Statistical Analyses section. The Turkish version of the scale has been validated and shown to be reliable (16).

DASS-21: The DASS-21, developed by Lovibond and Lovibond (17,18), is a self-report instrument widely used to assess negative emotional states. Its Turkish adaptation and psychometric validation were carried out by Yılmaz et al. (19). The tool comprises three subscales—depression, anxiety, and stress—each including seven items rated on a four-point Likert scale. The reliability and validity of the Turkish version across both non-clinical and clinical populations were later confirmed in further analyses by Sarıçam (20).

Statistical Analyses

All analyses were performed using IBM SPSS Statistics 26. Continuous variables were presented as mean ± standard deviation when normally distributed and as median (25^{th}-75^{th} percentiles) when the normality assumption was not met. Categorical variables were presented as frequencies and percentages. The T-SVEST employs a five-point Likert response format ranging from 1 (strongly disagree) to 5 (strongly agree). Most items in the colleague, supervisor, institutional, and non-work-related support domains are worded positively (e.g., "Sharing what happened with my colleagues helps me feel better"); responses for these items were reverse-coded during data entry so that a response of "strongly agree" was recorded as 1 and "strongly disagree" was recorded as 5. This ensured that, as with the remaining domains, a higher recorded score consistently reflects a worse outcome—in this case, a greater perceived lack of support—rather than reflecting greater support. A mean score of ≥4 therefore reflects the presence of second victimization across all domains, including the support domains once reverse-coded; higher scores throughout the instrument correspond to greater severity, whether in terms of psychological/physical distress, professional consequences, or lack of support. The data distribution was assessed using Kolmogorov-Smirnov and Shapiro-Wilk tests. Because the normality assumptions were not met, non-parametric methods were used.

Spearman's correlation was used to assess associations between continuous variables. Group differences in T-SVEST and DASS-21 scores by age, academic title, and years of practice were analyzed with the Kruskal-Wallis H test. Pairwise comparisons with Bonferroni correction were performed for significant results. The internal consistency of the T-SVEST and DASS-21 scales in the study sample was evaluated using Cronbach's alpha coefficients. T-SVEST subscale scores were compared between participants with and without prior experience of an event related to the second victim phenomenon using the Mann-Whitney U test. A two-tailed $p<0.05$ was considered statistically significant.

RESULTS

A total of 121 orthopedic and traumatology physicians participated in the study. The sociodemographic characteristics of the study sample are presented in Table 1.

The most preferred source of support was a respected colleague. Detailed results for T-SVEST and DASS-21 scores are shown in Tables 2 and 3.

In the present sample, the T-SVEST demonstrated excellent internal consistency, with a Cronbach's alpha coefficient of 0.948 for the total scale; the alpha coefficients of its subscales ranged from 0.837 to 0.908. The DASS-21 also showed excellent internal consistency, with Cronbach's alpha coefficients of 0.936 for the depression subscale, 0.920 for the anxiety subscale, 0.916 for the stress subscale, and 0.967 for the total scale.

The strongest positive correlations between total DASS scores and T-SVEST subscales were observed for professional self-efficacy, turnover intention, psychological distress, and physical distress. In particular, depression correlated strongly with turnover intention

Table 1. Sociodemographic characteristics of the study sample

Sociodemographic variables	n (%)
Gender	
Male	118 (97.5)
Female	3 (2.5)
Age	
<25	4 (3.3)
25-35	65 (53.7)
36-45	43 (35.5)
46-55	5 (4.1)
>55	4 (3.3)
Academic title	
Resident physician	58 (47.9)
Specialist physician	40 (33.1)
Assistant professor	2 (1.7)
Associate professor	17 (14)
Professor	4 (3.3)
Professional experience	
<2 years	13 (10.7)
2-5 years	38 (31.4)
6-10 years	26 (21.5)
11-15 years	23 (19)
>15 years	21 (17.4)
Exposure to an event potentially associated with the second victim phenomenon	
Yes	93 (76.9)
No	28 (23.1)
Awareness of the second victim phenomenon	
Yes	13 (10.7)
No	108 (89.3)

n: Number of participants, %: percentage

J Acad Res Med 2026;16(2):113-120

Table 2. Median T-SVEST subscale scores with percentage of agreement, and perceived desirability of potential support options among study participants

	Median (25th-75th percentile)	% of agreement (scores ≥4)
Psychological distress	3.00 (2.00-3.75)	23.9
Physical distress	2.50 (1.75-3.50)	18.18
Colleague support	3.00 (2.00-4.00)	25.61
Supervisor support	3.00 (2.50-3.50)	20.66
Institutional support	4.00 (3.00-4.50)	54.54
Non-work-related support	3.00 (2.00-4.00)	33.88
Professional self-efficacy	2.00 (1.50-3.00)	14.04
Turnover intentions	2.50 (1.50-3.50)	22.31
Absenteeism	2.00 (1.00-3.00)	15.70
	Median (25th-75th percentile)	**Neutral %, desired %**
The ability to take time away	3.00 (2.00-4.00)	18.2-38.8
A specified peaceful location	3.00 (2.00-4.00)	19.8-48.8
A respected peer to discuss	4.00 (3.00-4.00)	18.2-58.6
An employee assistance program	3.00 (2.00-4.00)	23.1-42.2
A discussion with manager or supervisor	3.00 (2.00-4.00)	19.8-48.7
The opportunity to schedule a counselor	3.00 (2.00-4.00)	27.3-34.7
A confidential discussion is available 24 h a day	3.00 (1.00-4.00)	25.6-27.3

T-SVEST: Turkish version of the Second Victim Experience and Support Tool, Neutral: 3 score, Desired: scores ≥4

Table 3. Median scores of the DASS-21 scale

	Median (25th-75th percentile)
DASS-depression	7.00 (3.00-12.00)
DASS-stress	8.00 (5.00-13.00)
DASS-anxiety	6.00 (1.00-8.00)
DASS-total	21.00 (9.00-31.00)

DASS-21: Depression Anxiety Stress Scale-21

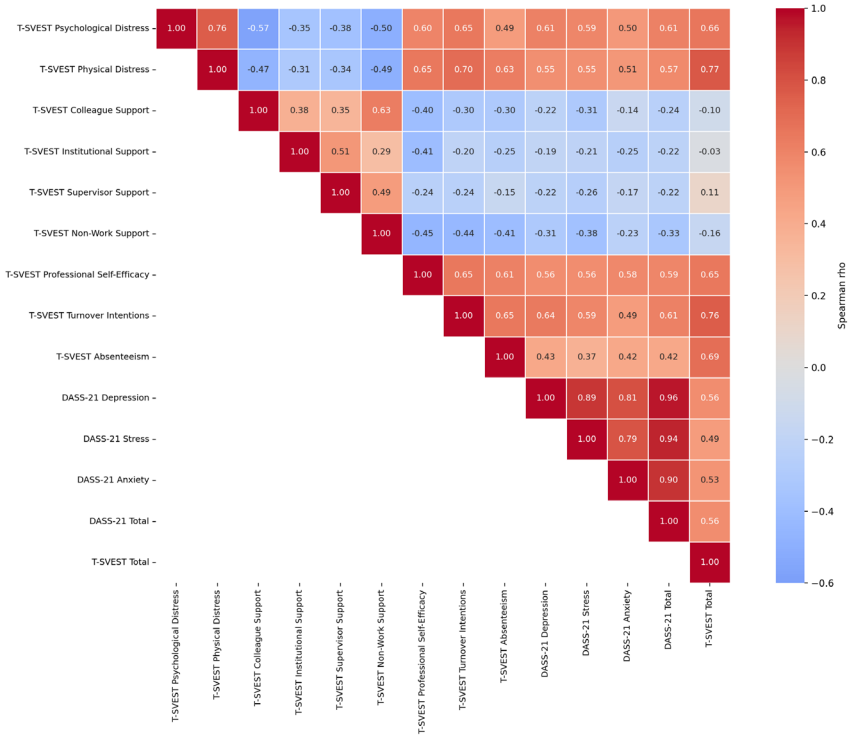

Figure 1. Correlations among participants' subscale scores

T-SVEST: Turkish version of the Second Victim Experience and Support Tool, DASS-21: Depression Anxiety Stress Scale-21

Color intensity reflects the magnitude of Spearman's rho values; red indicates positive and blue indicates negative correlations

J Acad Res Med 2026;16(2):113-120

(Spearman's rho=0.64) and psychological distress (Spearman's rho=0.61) (Figure 1).

Group differences in T-SVEST and DASS-21 subscale scores by age group, academic titles, and years of practice were analyzed using the Kruskal-Wallis H test. When significant differences were found, pairwise comparisons with Bonferroni correction were performed, and adjusted p-values were reported. A significance level of $p<0.05$ was accepted.

Kruskal-Wallis analysis by age group showed significant differences in psychological distress ($H=11.11$, $p=0.025$), physical distress ($H=10.75$, $p=0.029$), and absenteeism ($H=9.62$, $p=0.047$). However, pairwise comparisons did not reveal significant differences between specific age groups.

When comparing years of medical practice, the Kruskal-Wallis test showed significant differences in professional self-efficacy ($H=10.58$, $p=0.032$) and turnover intention ($H=10.73$, $p=0.030$). For turnover intention, pairwise comparisons revealed that physicians with >15 years of experience had a significantly higher mean rank (75.38) than those with <2 years of experience (36.96) (Bonferroni-adjusted $p=0.017$). For professional self-efficacy, no individual pairwise comparison remained significant after Bonferroni correction, despite the significant omnibus test.

When comparing T-SVEST subscale scores across academic title groups, the Kruskal-Wallis test revealed statistically significant differences in the physical distress ($H=11.98$, $p=0.017$) and absenteeism ($H=9.65$, $p=0.047$) subscales. No significant differences were found in the other subscales (all $p>0.05$). Pairwise comparisons demonstrated that associate professors had significantly higher mean rank scores on the physical distress subscale than resident physicians (Bonferroni-adjusted $p=0.010$; mean ranks: associate professors=83.18, residents=51.72). No significant differences were detected between other pairs of groups. Similarly, on the absenteeism subscale, resident physicians and associate professors differed significantly (Bonferroni-adjusted $p=0.028$; mean ranks: associate professors=81.06, residents=52.82).

Kruskal-Wallis analyses showed no significant differences in DASS-21 subscales or total scores across sociodemographic variables. In the Mann-Whitney U test, physicians with second-victim experiences scored significantly higher on psychological and physical distress, professional self-efficacy, and turnover intention (all $p<0.05$). Those without such experiences reported higher scores on colleague support, supervisor support, and non-work-related support ($p<0.05$). No differences in institutional support or absenteeism were found ($p>0.05$) (Table 4).

DISCUSSION

In this study, the multidimensional effects of the second-victim phenomenon were examined in a sample of 121 orthopedic and trauma physicians. Most participants were male and from younger age groups, consistent with previous reports (21). A significant proportion experienced at least one event potentially leading to second victimization, consistent with earlier findings. Surgeons frequently face complications and adverse events throughout their careers, which are among the most common triggers of second victim experiences (4,12,22). These results suggest that such complications are inherent to surgical practice but can still profoundly affect physicians both psychologically and professionally.

Nearly 90% of participants were not previously aware of the second victim phenomenon, indicating very low awareness among orthopedic and traumatology physicians in Türkiye. Although the concept was defined over two decades ago and systematic support programs have since been developed in many countries (4,5), this lack of awareness reflects institutional and educational shortcomings in our country.

In our study, about one-fourth of participants reported psychological distress, and about one-fifth reported physical distress, indicating that the second victim phenomenon involves both emotional and physical symptoms. These proportions were considerably lower than those reported by Alishaq et al. (23), who found psychological distress in 63% and physical distress

Table 4. Comparison of T-SVEST subscales between groups with and without exposure to an event potentially associated with the second victim phenomenon

T-SVEST	Exposure to an event potentially associated with the second victim phenomenon		
	Yes (n=93) Median (25^{th}-75^{th} percentile)	No (n=28) Median (25^{th}-75^{th} percentile)	p-value
Psychological distress	3.25 (2.50-4.00)	2.00 (1.00-3.00)	<0.001
Physical distress	3.00 (1.75-3.75)	2.00 (1.44-3.00)	0.007
Colleague support	2.50 (2.00-3.00)	4.00 (3.00-5.00)	<0.001
Supervisor support	3.00 (2.50-3.50)	3.50 (2.94-4.00)	0.011
Institutional support	4.00 (3.00-4.50)	4.00 (3.00-5.00)	0.356
Non-work-related support	3.00 (1.50-4.00)	4.00 (3.00-4.62)	0.001
Professional self-efficacy	2.50 (1.50-3.00)	2.00 (1.00-2.50)	0.014
Turnover intentions	3.00 (1.50-4.00)	2.00 (1.00-3.00)	0.013
Absenteeism	2.00 (1.00-3.00)	1.75 (1.00-3.00)	0.122

T-SVEST: Turkish version of the Second Victim Experience and Support Tool

in 68% of participants, and by Nosanov et al. (21), who reported anxiety, guilt, and shame in 84% of surgeons; this discrepancy may reflect differences in study population, measurement approach, or cultural context. Nonetheless, consistent with the broader literature, intrusive thoughts, sleep disturbances, and fatigue have also been highlighted as typical symptoms of the second victim phenomenon (5,24).

A considerable proportion of orthopedic physicians reported insufficient institutional support, which is consistent with previous studies. Nydoo et al. (5) found that 53% of healthcare workers received no support, while Alishaq et al. (23) reported that 31% felt their institutions did not care about their well-being. The rates of perceived insufficiency in colleague and supervisor support were relatively low; however, the fact that one in five physicians stated that they did not receive adequate support from their managers is still noteworthy. The literature emphasizes that an unfair and unsupportive institutional culture exacerbates second victimization, whereas structured and multi-layered support models—such as peer support teams, counseling programs, and open communication—play a critical role in protecting physician well-being (25-27). Taken together, our findings underline that improving perceived support systems remains one of the most urgent needs in Türkiye.

Findings regarding the loss of professional self-efficacy indicate that events potentially leading to second victimization undermine physicians' confidence in their professional identities. Consistent with our findings, Alishaq et al. (23) similarly reported that 39% of participants expressed low professional self-efficacy following adverse events. In addition to reducing self-confidence, this may contribute to long-term outcomes such as burnout, decreased performance, and disengagement from work. The observed turnover intentions and absenteeism further suggest that the second victim phenomenon threatens not only individual well-being but also institutional workforce sustainability, aligning with Burlison et al. (15), who reported a strong association between SVEST scores and turnover intention.

In our study, the most prominent outcomes of the second victim phenomenon among orthopedists were psychological and physical distress, both of which showed strong correlations with DASS-21 subscales. These results align with typical second victim symptoms reported in the literature, including depression, anxiety, and sleep disturbances (5,12,21,24). Across support dimensions, very weak to weak negative correlations were observed between high psychological distress and perceived insufficient support. As reported in some studies, this may suggest that support mechanisms are generally activated in the post-event process, and that those experiencing distress may receive more support at the time of the triggering event itself (15,25). This finding may also be interpreted as evidence that support systems in Türkiye, when available, tend to operate in a more reactive manner, or that differences in measurement timing may account for this pattern. Loss of professional self-efficacy correlated strongly with depression, anxiety, and stress, consistent with literature describing "moral injury" in physicians (22,28). Turnover intention and absenteeism were likewise associated with psychological distress, once again indicating that the second victim phenomenon threatens both physician well-being and institutional workforce sustainability (15,23).

In our study, age was associated with differences in psychological distress, physical distress, and absenteeism, while years of practice were linked to professional self-efficacy and turnover intention. With respect to turnover intention, orthopedists with more than 15 years of experience had a significantly higher mean rank than those with less than two years of experience, suggesting that later-career orthopedists may be more prone to professional disengagement. This finding warrants cautious interpretation, as the cross-sectional design and the absence of data on event frequency preclude drawing conclusions about the accumulation of adverse events over time. Although the omnibus test for professional self-efficacy was significant, no individual pairwise comparison remained significant after Bonferroni correction; therefore, this finding should be interpreted with caution. Given that senior physicians constituted a smaller proportion of the sample, this result merits confirmation in larger and more balanced samples. While the literature often emphasizes that younger surgeons report stronger feelings of guilt, anxiety, and inadequacy after complications (12,22,28), our findings suggest that turnover intention may also persist or even intensify among senior physicians over the course of their careers (23).

Among academic title groups, associate professors reported higher levels of physical distress and absenteeism than residents, suggesting that senior responsibilities may exacerbate the second victim phenomenon. The literature also notes that academic surgeons experience considerable distress after complications (21). Our finding that associate professors reported higher physical distress and absenteeism suggests that seniority alone does not protect against second victimization, consistent with the broader recognition that adverse events are linked to moral injury regardless of career stage (22).

Physicians with experience of adverse events reported greater psychological and physical distress, reduced self-efficacy, and increased turnover intention, indicating that second victimization is most severe among those directly exposed. Conversely, those without such experiences scored higher on colleague, supervisor, and non-work support, suggesting a perceived lack of support in potential incidents. It should be noted that our questionnaire assessed direct personal involvement in an adverse event rather than indirect or team-level exposure; therefore, elevated scores among participants without an identifiable adverse event cannot be firmly attributed to team-level second victimization as described in the literature (25,27). Instead, these scores may reflect broader perceptions of workplace distress or support, representing a potential confounding factor in experience-based comparisons rather than direct evidence of second victimization itself.

Overall, our findings, consistent with the international literature, show that the second-victim phenomenon among orthopedists

J Acad Res Med 2026;16(2):113-120

has multidimensional effects at both the individual and institutional levels (psychological and physical distress, loss of self-efficacy; turnover intention, absenteeism, insufficient support). Limited studies in Türkiye have similarly highlighted the vulnerability of younger physicians and inadequate support systems (29-31). As the first empirical study of orthopedists in Türkiye, this research fills a significant gap and underscores the need to address second victimization more seriously in surgical specialties.

Study Limitations

This study has several limitations. Participants were not asked how often they experienced adverse events, and differences in timing between the DASS-21 (past week) and T-SVEST (post-event) may have influenced the results. Although over 75% reported at least one adverse event, the inclusion of participants without such experiences remains a limitation, as the questionnaire did not distinguish between the complete absence of exposure and indirect or team-level exposure that was not directly assessed. Additionally, personal history or other occupational stressors were not assessed, meaning that part of the distress observed may reflect additional emotional burdens rather than second victimization alone. Another limitation is the lack of documentation regarding the type of triggering event, which may limit the interpretability of prevalence findings. Comparisons by institution type were not reported, as several institution categories contained very small subgroup sizes (e.g., n=2), which would have limited the reliability and interpretability of such comparisons. Finally, the sample, being predominantly male and composed of junior residents/specialists, further limits generalizability, especially in comparison to senior physicians. Furthermore, this study employed a convenience-based online sampling strategy, and the response rate could not be calculated because the number of physicians who received the survey link was unknown. This limits the generalizability of the findings, and the reported frequencies should therefore be interpreted as estimates of exposure and self-reported experience rather than as formal prevalence figures.

Second victimization is better conceptualized as a psychosocial stress response rather than a clinical psychiatric disorder, and future research may help clarify whether this phenomenon could be situated within broader health-related classifications. Future research may also expand upon these findings by distinguishing between event types (e.g., medical error, complication, adverse event) to better understand how the nature of the incident shapes second victim responses. Additionally, extending similar research to other healthcare professional groups may increase conceptual visibility and contribute to a broader understanding of second victim experiences across clinical disciplines.

CONCLUSION

This study is the first to examine the second-victim phenomenon among orthopedists in Türkiye and shows that such experiences may lead to intense psychological distress, physical symptoms of stress, loss of self-efficacy, and professional consequences, such as absenteeism and turnover intentions. Turnover intention appeared more pronounced among physicians with longer professional experience, indicating that second victimization and its professional consequences extend well beyond the early-career stage. The findings further suggest that existing support systems were often perceived as inadequate and that preventive strategies and structured post-event support should be prioritized to protect both clinician well-being and workforce sustainability. While the study does not establish causality or explore institutional and cultural mechanisms in depth, identifying the prevalence and the emotional impact provides an important starting point for the field. By increasing awareness of second victimization within orthopedics, this work offers an empirical foundation for future research—particularly studies that differentiate event types, examine contextual triggers, and involve broader healthcare professional groups to strengthen conceptual clarity and guide the development of more supportive clinical environments.

Ethics

Ethics Committee Approval: The study was approved by the Ethics Committee of İstanbul Medipol University (decision no: 606, date: 22 May 2025) and conducted in accordance with the Declaration of Helsinki.

Informed Consent: The invitation message clearly stated that participation was voluntary and informed consent was required.

Footnotes

Author Contributions: Concept - Y.S.F., A.K., S.Y.Ç., İ.H.A., F.F.; Design - Y.S.F., A.K., S.Y.Ç., İ.H.A., F.F.; Data Collection and/or Processing - Y.S.F., S.Y.Ç., F.F.; Analysis and/or Interpretation - Y.S.F., S.Y.Ç., F.F.; Literature Search - Y.S.F., S.Y.Ç., F.F.; Writing - Y.S.F., A.K., S.Y.Ç., İ.H.A., F.F.

Conflict of Interest: The authors have no conflict of interest to declare.

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

REFERENCES

1. Wu AW. Medical error: the second victim. The doctor who makes the mistake needs help too. BMJ. 2000; 320: 726-7.
2. MacLeod L. "Second victim" casualties and how physician leaders can help. Physician Exec. 2014; 40: 8-12.
3. Coughlan B, Powell D, Higgins MF. The second victim: a review. Eur J Obstet Gynecol Reprod Biol. 2017; 213: 11-6.
4. Scott SD, Hirschinger LE, Cox KR, McCoig M, Hahn-Cover K, Epperly KM, et al. Caring for our own: deploying a systemwide second victim rapid response team. Jt Comm J Qual Patient Saf. 2010; 36: 233-40.
5. Nydoo P, Pillay BJ, Naicker T, Moodley J. The second victim phenomenon in health care: a literature review. Scand J Public Health. 2020; 48: 629-37.
6. Xu M, Wang Y, Yao S, Shi R, Sun L. One-year prevalence of perceived medical errors or near misses and its association with depressive symptoms among Chinese medical professionals: a propensity score matching analysis. Int J Environ Res Public Health. 2022; 19: 3286.
7. Busch IM, Moretti F, Purgato M, Barbui C, Wu AW, Rimondini M. Psychological and psychosomatic symptoms of second victims of adverse events: a systematic review and meta-analysis. J Patient Saf. 2020; 16: e61-e74. Erratum in: J Patient Saf. 2020; 16: e211.
8. Dean W, Talbot S, Dean A. Reframing clinician distress: moral injury not burnout. Fed Pract. 2019; 36: 400-2. Erratum in: Fed Pract. 2019; 36: 447.
9. Kim SA, Kim EM, Lee JR. Causes of nurses' second victim distress: an objective analysis. Qual Manag Health Care. 2022; 31: 122-9.
10. Waterman AD, Garbutt J, Hazel E, Dunagan WC, Levinson W, Fraser VJ, et al. The emotional impact of medical errors on practicing physicians in the United States and Canada. Jt Comm J Qual Patient Saf. 2007; 33: 467-76.

11. Ozeke O, Aras S, Baser K, Sen F, Kirbas O, Cay S, et al. Defensive medicine due to different fears by patients and physicians in geriatric atrial fibrillation patients and second victim syndrome. Int J Cardiol. 2016; 212: 251-2.
12. Jain G, Sharma D, Agarwal P, Agrawal V, Yadav SK, Tenzin T, et al. "Second victim" syndrome among the surgeons from South Asia. Indian J Surg. 2022; 84: 40-6.
13. Chong RIH, Yaow CYL, Chong NZ, Yap NLX, Hong ASY, Ng QX, et al. Scoping review of the second victim syndrome among surgeons: understanding the impact, responses, and support systems. Am J Surg. 2024; 229: 5-14.
14. Oktay EA. Ortopedi ve travmatoloji uygulamaları açısından aydınlatılmış onam [Informed consent in orthopedics and traumatology]. TOTBİD Derg. 2020; 19: 32-7. Turkish.
15. Burlison JD, Scott SD, Browne EK, Thompson SG, Hoffman JM. The second victim experience and support tool: validation of an organizational resource for assessing second victim effects and the quality of support resources. J Patient Saf. 2017; 13: 93-102.
16. Koca A, Elhan AH, Genç S, Oğuz AB, Eneyli MG, Polat O. Validation of the Turkish version of the second victim experience and support tool (T-SVEST). Heliyon. 2022; 8: e10553.
17. Lovibond SH, Lovibond PF. Manual for the Depression Anxiety Stress Scales. 2nd ed. Sydney: Psychology Foundation of Australia; 1995.
18. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther. 1995; 33: 335-43.
19. Yılmaz Ö, Boz H, Arslan A. Depresyon Anksiyete Stres Ölçeğinin (DASS 21) Türkçe kısa formunun geçerlilik-güvenilirlik çalışması [The validity and reliability of Depression Stress and Anxiety Scale (DASS21) Turkish short form]. Finans Ekon Sos Araştırmalar Derg. 2017; 2: 78-91. Turkish.
20. Sarıçam H. The psychometric properties of Turkish version of Depression Anxiety Stress Scale-21 (DASS-21) in health control and clinical samples. JCBPR. 2018; 7: 19-30.
21. Nosanov L, Elseth AJ, Maxwell J, Alimi YR, Giri O, Millar JK, et al.; AWS Publications Committee. The things we carry: the scope and impact of second victim syndrome. Am J Surg. 2023; 226: 726-8.
22. Richard RD, Pesante BD, Parry JA, Mauffrey C. The effect of adverse events on orthopaedic surgeons: a review. J Am Acad Orthop Surg. 2024; 32: 771-6.
23. Alishaq M, Kamal Hassan A, Assar AH, Elfaki S, Sobhi Mohamed SM, Abdelaliem SMF, et al. The impact of second-victim syndrome on staff well-being and practice. Medicine (Baltimore). 2025; 104: e41345.
24. Sachs CJ, Wheaton N. Second victim syndrome. [Updated 2023 Jun 20]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2026. Available from: https://www.ncbi.nlm.nih.gov/books/NBK572094/
25. Hall LW, Scott SD. The second victim of adverse health care events. Nurs Clin North Am. 2012; 47: 383-93.
26. Kuris EO, de Kleuver M, Alsoof D, McDonald CL, Daniels AH. Health care professional mental health and supporting the second victim - can organizations across the world do better? Am J Med. 2022; 135: 667-9.
27. Seys D, Panella M, Russotto S, Strametz R, Joaquín Mira J, Van Wilder A, et al. In search of an international multidimensional action plan for second victim support: a narrative review. BMC Health Serv Res. 2023; 23: 816.
28. Gold P, Arshi A. Dealing with complications as a young surgeon. Arthroplast Today. 2024; 27: 101419.
29. Ozeke O, Ozeke V, Coskun O, Budakoglu II. Second victims in health care: current perspectives. Adv Med Educ Pract. 2019; 10: 593-603.
30. Ayvat P, Şehitoğlu Alpağut G, Aşar K, Başer A, Özlü C, Ejderoğlu A. Anestezi ve yoğun bakımda ikinci mağdur (kurban) olgusu [The second victim phenomenon in anesthesia and intensive care]. İzmir Tıp Fak Derg. 2023; 2: 45-50. Turkish.
31. Ayvat P, Ayvat AG, Bilek G, Oner O, Sehitoglu Alpagut G, Sonmez OF. Second-victim experience in anesthesia and intensive care settings: organizational gaps, support preferences, and recommendations for action - a cross-sectional study. Medicine (Baltimore). 2025; 104: e43375.

DOI: 10.4274/jarem.galenos.2026.27879
J Acad Res Med 2026;16(2):121-122

Responsible Healthcare AI Requires Responsible Evidence: Methodological Governance for Clinical Implementation

Ahmet Aykut

University of Health Sciences Türkiye, Van Education and Research Hospital, Department of Emergency Medicine, Van, Türkiye

Cite this article as: Aykut A. Responsible healthcare AI requires responsible evidence: methodological governance for clinical implementation. J Acad Res Med. 2026;16(2):121-122

Keywords: Artificial intelligence, clinical validation, methodological transparency

Dear Editor,

The commentary by Öney Doğanyiğit and Yılmaz Altuntaş (1) provides a timely reminder that artificial intelligence (AI) adoption in healthcare should be governed by transparency, privacy, clinical validation, accountability, and lifecycle monitoring. This position is highly relevant, as generative systems and large language models (LLMs) are increasingly framed not only as administrative tools but also as interactive clinical partners.

However, the central challenge is the responsible transformation of not only healthcare but also evidence. In the absence of structured reporting, external validation, calibration, and post-deployment surveillance, the phrase "clinical validation" may remain too broad to protect patients or guide clinicians. This is particularly important when AI systems are evaluated using convenience datasets, static benchmark questions, single-center records, or unblinded expert panels without sufficient detail on model versioning, prompt design, input preprocessing, missing data handling, reference standards, or error adjudication.

Recent reporting frameworks provide a practical route from principles to implementable evidence. TRIPOD+AI clarifies reporting for diagnostic and prognostic prediction models using regression or machine learning methods (2). DECIDE-AI addresses early-stage clinical evaluation of AI-driven decision-support systems, including workflow interaction and human-system communication (3). For generative models and LLM-based applications, TRIPOD-LLM and MI-CLEAR-LLM highlight aspects that are often invisible in conventional biomedical reporting: model identity, access route, prompt and execution details, stochasticity management, adaptation strategy, independence of test data, and task-specific performance reporting (4,5). These elements are essential because LLM outputs may change with vendor updates, interface modifications, retrieval layers, and prompt variations, even when the nominal clinical task remains the same.

For this reason, healthcare AI governance should include a minimum evidentiary chain before clinical implementation: precise description of the AI system, transparent reporting of the clinical task and intended users, representative test data, clinically meaningful reference standards, external validation when possible, calibration and error analysis, assessment of human-AI interaction, and lifecycle monitoring after deployment. Such a chain would make regulatory and ethical principles auditable rather than merely aspirational.

The commentary rightly emphasizes that AI should augment, not replace, clinicians (1). Meaningful human oversight, however, requires access to interpretable and reproducible evidence. Clinicians cannot responsibly supervise systems whose development, evaluation, and updating are insufficiently reported. Therefore, responsible healthcare AI depends not only on trust, empathy, and regulation but also on methodological transparency. Without robust evidence, responsible transformation may remain incomplete.

Acknowledgements

The author acknowledges the use of artificial intelligence-based language assistance during the preparation of the manuscript. The final content, interpretation, and conclusions are the sole responsibility of the author.

ORCID ID of the author: A.A. 0009-0001-3173-8994

Corresponding Author: Ahmet Aykut, MD;
E-mail: ahmet.aykut@gmail.com
Received Date: 07.06.2026 **Accepted Date:** 25.06.2026
Epub: 26.06.2026
Publication Date: 07.08.2026

J Acad Res Med 2026;16(2):121-122

Footnotes

Financial Disclosure: The author declared no financial support.

REFERENCES

1. Öney Doğanyiğit S, Yılmaz Altuntaş E. Artificial intelligence in healthcare: a critical moment for responsible transformation. J Acad Res Med. 2026; 16: 1-3.
2. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024; 385: e078378. Erratum in: BMJ. 2024; 385: q902.
3. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al.; DECIDE-AI Expert Group. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. BMJ. 2022; 377: e070904.
4. Gallifant J, Afshar M, Ameen S, Aphinyanaphongs Y, Chen S, Cacciamani G, et al. The TRIPOD-LLM reporting guideline for studies using large language models. Nat Med. 2025; 31: 60-9.
5. Park SH, Suh CH, Lee JH, Tejani AS, You SC, Kahn CE, et al. Minimum reporting items for clear evaluation of accuracy reports of large language models in healthcare (MI-CLEAR-LLM): 2025 updates. Korean J Radiol. 2025; 26: 1123-32.

DOI: 10.4274/jarem.galenos.2026.c001
J Acad Res Med 2026;16(2):123

Öney Doğanyiğit S, Yılmaz Altuntaş E. Artificial intelligence in healthcare: a critical moment for responsible transformation. J Acad Res Med. 2026;16(1):1-3.

DOI: 10.4274/jarem.galenos.2026.68926

Following publication of the above article, the authors identified an error in Reference 17, which had been inadvertently cited incorrectly. The reference has been corrected, and the corresponding text under the **"Human Oversight and Clinical Expertise"** section has been updated accordingly.

Reference 17 was originally published as:

17. Naing SY, Severin J, Zomer A, Kuntaman K, Suandy I, Sunandar S, et al. Impact of reducing colistin use on colistin resistance in *Escherichia coli* isolated from humans and poultry in Indonesia (COINCIDE): a protocol for a multisectoral, transdisciplinary One Health study. One Health. 2026; 22: 101347.

The correct Reference 17 is:

17. Powell MR. Empathy in digital healthcare (dissertation). Orlando (FL): University of Central Florida. 2024.

Accordingly, the following text under the **"Human Oversight and Clinical Expertise"** section:

"According to Schneiders and van der Graaf (17), traditional empathy is a process where one person senses another's feelings, but it is resource-limited (requiring time and energy) and can be tiring (burnout). Scalable empathy, on the other hand, is the simulation of this process at a behavioral level through AI, enabling it to be delivered to millions of patients simultaneously and seamlessly."

Has been replaced by the following:

"Scalable empathy is the simulation of this process at a behavioral level through artificial intelligence, enabling it to be delivered to millions of patients simultaneously and seamlessly. In Powell's (17) study, attention is drawn to the pivotal role of AI and digital interfaces in striving to sustain empathy, demonstrating how human-centered digital infrastructures can actively bridge the communication gap between clinicians and patients. These studies highlight that within the realm of digital health solutions, integration of artificial intelligence and digital tools serves as a strategic means to preserve and cultivate empathy a cornerstone of the physician-patient relationship."

The authors apologize for this error and state that this correction does not affect the scientific conclusions of the article.