ANKARA MEDICAL JOURNAL

Vol. 26/ Num. 3, 2026
www.ankaramedj.com
e-ISSN: 2148-4570

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

Managing Editor and Owner

- Ahmet KESKİN, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Dept. of Family Medicine

Editor in Chief

- Ahmet KESKİN, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Dept. of Family Medicine

Editors

- Hameed AKLAN, Sana University of Science and Technology, Faculty of Medicine, Dept. of Radiology, Yemen
- Cüneyt ARDIÇ, Recep Tayyip Erdoğan University Faculty of Medicine, Department of Family Medicine, Turkey
- La AVALIANI, Tbilisi State Medical University, Dept. of Propedeutics, Georgia
- Aylin BAYDAR ARTANTAŞ, Ankara City Hospital, Family Medicine Clinic, Turkey
- Mahcube ÇUBUKÇU, Samsun TRH Family Medicine Clinic, Turkey
- Memet Taşkın EGİCİ, University of Health Sciences, İstanbul Haydarpaşa ERH, Clinic of Family Medicine, Turkey
- Eiad A. AL-FARIS, King Saud University, College of Medicine, Dept. of Family and Community Medicine, Saudi Arabia
- Umut GÖK BALCI, İzmir Tepecik TRH, Family Medicine Clinic, Turkey
- Didem KAFADAR, *Istanbul University-Cerrahpaşa, Cerrahpaşa Faculty of Medicine, Department of Family Medicine, Turkey*
- İrep KARATAŞ ERAY, Ankara City Hospital, Family Medicine Clinic, Turkey
- Abdulsattar KHAN, King Faisal University, College of Medicine, Dept. of Family and Community Medicine, Saudi Arabia
- Erhan ŞİMŞEK, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Dept. of Family Medicine, Turkey
- Ebru UĞRAŞ, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Dept. of Family Medicine, Turkey
- Güzin ZEREN ÖZTÜRK, Şişli TRH, Family Medicine Clinic, Turkey

Biostatistics Editor

- Yavuz SANİSOĞLU, Ankara Yıldırım Beyazıt University Faculty of Medicine, Dept. of Biostatistics, Turkey

Web Editor

- Erhan ŞİMŞEK, Ankara Yıldırım Beyazıt University Faculty of Medicine, Dept. of Family Medicine, Turkey

English Editing

- Gamze BOZCUK GÜZELDEMİRCİ, Ankara Provincial Health Directorate, Public Health Services, Education Sub-Unit, Turkey

Editorial Advisory Board

- Akın AKTAŞ, Ankara Yıldırım Beyazıt University Faculty of Medicine, Department of Dermatology
- İsmail ARSLAN, Ankara Training and Research Hospital, Department of Family Medicine
- Ümit AYDOĞAN, University of Health Sciences, Ankara Gülhane ERH, Dept. of Family Medicine
- Okay BAŞAK, Adnan Menderes University, Faculty of Medicine, Dept. of Family Medicine
- Uğur BİLGE, Eskişehir Osmangazi University, Faculty of Medicine, Dept. of Family Medicine
- Işıl İrem BUDAKOĞLU, Gazi University, Faculty of Medicine, Dept. of Medical Education
- Mehmet Fatih CEYLAN, Ankara Yıldırım Beyazıt University, Yenimahalle TRH, Dept. of Pediatric Psychiatry
- Fatma Gökşin CİHAN, Konya NEÜ, Meram Faculty of Medicine, Dept. of Family Medicine
- Bekir ÇAKIR, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Dept. of Internal Medicine
- Ali ÇAYKÖYLÜ, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Dept. of Psychiatry
- Mustafa ÇELİK, University of Health Sciences, Ankara ERH, Clinic of Family Medicine
- Reşat DABAK, Lütfi Kırdar Kartal ERH, Clinic of Family Medicine
- Reyhan ERSOY, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Dept. of Internal Medicine
- Nazan KARAOĞLU, Konya NE University, Faculty of Medicine, Dept. of Family Medicine
- Ljiljana Cvejanov KEZUNOVIC, University of Montenegro, Medical Faculty Podgorica, Dept. of Family Medicine
- Ruhuşen KUTLU, Konya NEÜ, Meram Faculty of Medicine, Dept. of Family Medicine
- Valentina Christova MADJOVA, Medical University of Varna, Dept. of General Medicine
- Kamile MARAKOĞLU, Selçuk University Faculty of Medicine, Department of Family Medicine
- M. Mümtaz MAZICIOĞLU, Kayseri Erciyes University, Faculty of Medicine, Dept. of Family Medicine
- Salih MOLLAHALİLOĞLU, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Dept. of Public Health
- Adem ÖZKARA, University of Health Sciences, Ankara City Hospital, Dept. of Family Medicine
- Danica Rotar PAVLIČ, University of Ljubljana, Faculty of Medicine, Dept. of Family Medicine
- Richard ROBERTS, University of Wisconsin, Faculty of Medicine, Dept. of Family Medicine, Past President of WONCA and AAFP
- Esra SAATÇİ, Çukurova University, Faculty of Medicine, Dept. of Family Medicine
- Mehmet SARGIN, İstanbul Medeniyet University, Faculty of Medicine, Dept. of Family Medicine
- Oktay SARI, University of Health Sciences, Gülhane Medical Faculty, Department of Family Medicine
- Ljubin ŠUKRIEV, President of AGP/FM SEE
- Zeynep TUZCULAR VURAL, İstanbul Haydarpaşa Numune ERH, Dept. of Family Medicine
- İlhami ÜNLÜOĞLU, Eskişehir Osmangazi University, Faculty of Medicine, Dept. of Family Medicine
- Chris van WEEL, Radboud University Nijmegen, The Netherlands, Emeritus Professor of Family Medicine / General Practice / Australian National University, Professor of Primary Health Care Research / Past President of WONCA
- Abdussamed YALÇIN, Ankara Yıldırım Beyazıt University Faculty of Medicine, Department of General Surgery
- Bülent YALÇIN, Ankara Yıldırım Beyazıt University Faculty of Medicine, Department of

Internal Medicine
- Ayşe Filiz YAVUZ, Ankara Yıldırım Beyazıt University Faculty of Medicine, Department of Obstetrics & Gynecology
- Ömer Hınç YILMAZ, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Dept. of Public Health

Editorial Secretary
- Ebru UĞRAŞ, Ankara Yıldırım Beyazıt University Faculty of Medicine, Dept. of Family Medicine, Turkey
- Erhan ŞİMŞEK, Ankara Yıldırım Beyazıt University Faculty of Medicine, Dept. of Family Medicine, Turkey

Organization, Preparation and Correspondence
- Ankara Yıldırım Beyazıt University, Faculty of Medicine, Department of Family Medicine, 06800 Bilkent / Ankara / TURKEY

Date of Issue: 22.09.2026

Ankara Medical Journal is an international peer-reviewed journal and is published quarterly. The responsibility of the published articles belongs to the authors.

Ankara Medical Journal is indexed / abstracted in Scopus, DOAJ, ULAKBIM TR Index, Google Scholar, Index Copernicus, ROAD, ResearchBib, SHERPA/RoMEO, Turkish Citation Index and Turkish Medline.

From the Editor

Dear Readers,

It is our pleasure to share the third issue of 2026 with you. This issue brings together a diverse collection of scholarly contributions focusing on current topics, emerging challenges, and important developments in primary care and healthcare delivery. We are pleased to present 13 original research articles offering perspectives that we hope will be of value to healthcare professionals, particularly those working in primary care.

As Türkiye's leading journal in the field of primary care, we continue our commitment to providing high-quality, timely, and evidence-based scientific content. We believe that sharing research and fostering scientific dialogue are essential to supporting professional development and contributing to the continuous improvement of healthcare services. We sincerely appreciate the efforts of our authors, reviewers, readers, and all those who contribute to the ongoing development of our journal.

We hope that the articles featured in this issue will provide useful insights for clinical practice, stimulate further discussion and research, and contribute to the advancement of primary care. The continued interest and engagement of our readers remain an important source of motivation for us as we work to strengthen scientific communication and support innovation in healthcare.

We are delighted to share this new issue with you and look forward to bringing together another selection of valuable contributions in our forthcoming issue.

Prof. Dr. Ahmet Keskin

ANKARA MEDICAL JOURNAL

Contents

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):319-336 // 10.5505/amj.2026.72929

ADAPTATION OF A STRUCTURED DIGITAL 5C CONSULTATION MODEL AND ITS EFFECT ON PATIENT MANAGEMENT INTERVALS IN THE EMERGENCY DEPARTMENT

Furkan Altaş[1], Çağdaş Yıldırım[2], Oğuzhan Zengin[3], Fatih Ahmet Kahraman[4], Gulhan Kurtoglu Celik[2], Mehmet Yılmaz[1]

[1]Department of Emergency Medicine, Ankara Bilkent City Hospital, Ankara, Türkiye
[2]Department of Emergency Medicine, Ankara Yildirim Beyazit University Faculty of Medicine, Ankara, Türkiye
[3]Department of Internal Medicine, Ankara Bilkent City Hospital, Ankara, Türkiye
[4]Department of Emergency Medicine, Lösante Children and Adult Hospital,Ankara, Türkiye

Correspondence:
Furkan Altaş (e-mail: furkanaltasmd@gmail.com)

Submitted: 18.03.2026 // Accepted: 05.07.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: Consultation-related delays contribute to emergency department (ED) crowding, especially for internal medicine referrals. The 5C consultation model provides a structured communication framework that may improve consultation efficiency when embedded in digital systems.

Materials and Methods: This single-center quasi-experimental pre-post study was conducted in the ED of a tertiary-care hospital with approximately 450,000 annual visits. Internal medicine consultations requested from yellow-zone observation units were compared before (July 1-10, 2023) and after (July 20-30, 2023) implementation of a structured digital consultation form based on four 5C elements (Contact, Communicate, Core Question, Closing the Loop) following a 40-minute resident training session. The primary outcome was consultation-to-decision time.

Results: Fifty-four emergency medicine residents generated 194 pre-intervention and 141 post-intervention consultations. Median consultation-to-decision time decreased from 65 minutes (IQR 40-112) to 44 minutes (IQR 28-69; $p<0.001$). Consultation response time and final decision time also decreased significantly (both $p<0.001$). Total ED LOS remained unchanged (median, 15 hours in both periods; $p=0.772$). Repeat consultations were less frequent after the combined intervention ($p=0.035$).

Conclusion: The four-component 5C-derived model digital consultation form, combined with targeted resident training, was associated with shorter consultation-related decision intervals and fewer repeat internal medicine consultations, without changing overall ED LOS. These preliminary findings suggest that digitizing structured communication may improve workflow efficiency, though multicenter randomized trials are required to confirm these effects.

Keywords: Emergency department, referral and consultation, communication, electronic health records, patient care management.

Introduction

The emergency department (ED) is one of the hospital settings in which specialist consultations are requested most frequently.[1] Consultations initiated from the ED are generally sought to support key aspects of patient care, including determining the need for hospital admission, refining diagnostic and therapeutic decision-making, performing specialized procedures, and coordinating safe follow-up and post-discharge care.[1,2]

The duration between the consultant's arrival at the ED, the completion of the consultation, and final decision-making accounts for approximately 33% of the ED length of stay (LOS) in discharged patients and 54% in hospitalized patients.[3] Studies have demonstrated that prolonged ED LOS is associated with decreased patient satisfaction, increased rates of hospital admission, and elevated mortality.[4] Conversely, implementing specific communication systems has been shown to improve consultation response times and reduce ED LOS.[4-6]

To improve consultation efficiency in emergency settings, Kessler developed the 5C Consultation Model, which aims to standardize the consultation process. This model provides a structured framework for comprehensive patient assessment, identification of optimal treatment options, and coordination of interdepartmental transfers. The 5Cs represent the following sequential steps: Contact, Communicate, Core Question, Collaborate, and Closing the Loop.[7]

Due to the broad scope of internal medicine and the high frequency of comorbid internal diseases in patients presenting to various specialties, it remains one of the most frequently consulted departments in emergency medicine.[8] Previous studies evaluating delays in the ED have identified consultation-related factors as the third most common cause of delay among patients requiring internal medicine evaluation.[9,10]

We hypothesized that using a structured digital format based on the 5C model would significantly reduce consultation decision times compared with standard unstructured

request methods. We aimed to integrate the four-component 5C-derived model into the hospital digital consultation form and evaluate the effect of internal medicine consultations requested via this digital format on patient management.

Materials and Methods

Study design and setting

We conducted a single-center quasi-experimental pre-post intervention study in the Emergency Medicine Department of Ankara Bilkent City Hospital, a tertiary-care center with an annual ED census of approximately 450,000 visits. The study was carried out from June 1, 2023, to September 1, 2023. The ED is organized using a triage-based workflow with eight observation units: two high-acuity (red) units, four urgent (yellow) units, and two non-urgent (green) units. Each unit contains 24 stretcher bays. This manuscript was revised with reference to the TREND statement for the reporting of nonrandomized evaluations of behavioral and public health interventions.

The study was conducted in accordance with the ethical principles of the Declaration of Helsinki and was approved by the Ankara Bilkent City Hospital Ethics Committee (approval no. E1-23-3508; approval date: May 10, 2023). Written informed consent was obtained from all participating physicians and patients. No formal a priori sample size calculation was performed because this study was designed as a pilot quasi-experimental study to evaluate the feasibility and preliminary effect of the structured digital consultation intervention. Therefore, all eligible consultations during the predefined study periods were included in the analysis.

Participants

We analyzed internal medicine consultation requests initiated by emergency medicine resident physicians for patients managed in the ED yellow-zone observation units.

Because prior exposure to a structured consultation curriculum can influence consultation performance, residents who had previously received formal training in the 5C consultation model were excluded.

For all participating residents, we collected demographic and training characteristics, including age (years), sex, and residency training duration (months).

Inclusion criteria

The ED was the patient's initial point of care.

The patient was being monitored in a yellow-zone observation unit.

The consultation request was directed to the Department of Internal Medicine.

Exclusion criteria

The time metrics of repeat internal medicine consultations for the same patient encounter after the initial request. Repeat consultations were not included as separate observations in the primary time-based analyses; only the first internal medicine consultation for each ED encounter was considered the index consultation.

Patients initially managed in red or green zones, or transferred to these zones during follow-up. Incomplete or unfinalized consultations (e.g., no documented consultant response or decision). Patients who left the ED before treatment or evaluation was completed. Consultation requests in which the study-specific digital consultation form was not used during July 20-30, 2023.

Intervention

In the study hospital, consultations are initiated and completed via the Hospital Information Management System (HIMS). The system records the following timestamps: time of

consultation request, time of the consultant's first evaluation, and time the consultant completes and records the response.

The original 5C consultation model was developed for verbal consultation requests. Because real-time verbal collaboration could not be incorporated into the digital workflow, we developed a four-component 5C-derived digital consultation form containing Contact, Communicate, Core Question, and Closing the Loop. This digital form, together with the resident training session, constituted the combined intervention evaluated in this study (Table 1).

Table 1. Four-component 5C-derived digital consultation form adapted for use in the emergency department.

Component	Template content
Contact	Dear Colleague,
Communicate	A ...-year-old male/female presented to the emergency department with the chief complaint of.... Symptoms have been present for ... days and are associated with.... Past Medical History: ... Current Medications: ... Pertinent Physical Examination Findings: ... Vital Signs: Temperature ... °C; Heart rate ... bpm; Blood pressure .../... mmHg; SpO2 ...%. Laboratory Findings: ... Imaging Findings: ...
Core Question	In the emergency department, the patient was started on ... for a working diagnosis of.... We kindly request your evaluation and recommendations regarding ... at your earliest convenience.
Closing the Loop	Observation unit: ... / Dr. ...

All consultations included in the study were extracted from HIMS during July 2023. The internal medicine department at the study hospital operates through various subspecialty

units, and emergency consultant physicians rotate monthly across clinics. To standardize the consultant clinic and minimize variability, data collection was restricted to July 2023.

From July 1-10, 2023, no intervention was made in the routine consultation request process (pre-intervention). A 40-minute didactic lecture explaining the 5C model and the use of the digital consultation form was delivered to participating emergency medicine residents via Zoom on July 15, 2023. On July 20, the digital consultation form was installed on computers in the observation units. Residents were instructed to use the form when requesting internal medicine consultations from July 20-30, 2023. Internal medicine consultants and the researchers conducting the data analysis were blinded to the study intervention.

The 40-minute Zoom-based training session included: (1) the rationale for structured consultation in the ED, (2) an overview of the original 5C consultation model, (3) the adaptation of four 5C components to the digital consultation workflow, (4) step-by-step instruction on completing the electronic consultation form, (5) example consultation scenarios, and (6) a brief question-and-answer session. During the intervention period, the study-specific consultation template was installed on observation-unit computers and residents were instructed to use it for all eligible internal medicine consultations. Compliance was assessed retrospectively through HIMS records, and consultation requests that did not use the study-specific digital form during the intervention period were excluded from analysis.

Measurements and outcomes

Consultation response time (minutes): the interval between the time the consultation request was entered into HIMS and the time of the consultant's first documented evaluation.

Consultation-to-decision time (minutes): the interval between the time the consultation request was entered into HIMS and the time the consultant completed and documented the consultation response in the system.

Final decision time (hours): the interval between the initial consultation request and the time the final ED disposition decision (hospitalization or discharge) was recorded.

ED length of stay (ED LOS): the interval between the patient's registration or admission to the ED and physical departure from the ED.

Repeat consultation frequency: The total count of any subsequent internal medicine consultation requests generated for the same patient following the initial consultation. The number of included initial index consultations served as the denominator for this outcome.

The primary outcome was consultation-to-decision time. Secondary outcomes were consultation response time, final decision time, ED LOS, and repeat consultation frequency.

Statistical Analysis

Statistical analyses were performed using SPSS for Windows, Version 20.0. Resident demographic characteristics were summarized descriptively. Categorical variables were reported as frequencies and percentages and were compared using the chi-square test or Fisher's exact test, as appropriate. The distribution of continuous variables was assessed using the Shapiro-Wilk test. Normally distributed continuous variables were reported as mean +/- standard deviation (SD), whereas non-normally distributed variables were reported as median and interquartile range (IQR). The independent samples t-test was used for normally distributed variables, and the Mann-Whitney U test was used for non-normally distributed variables. In addition to p-values, absolute between-group differences were reported to improve clinical interpretability. For non-normally distributed outcomes, median differences were presented. Effect sizes were also calculated for between-group comparisons where appropriate. A p-value of <0.05 was considered statistically significant.

Results

Consultations requested by 54 emergency medicine residents were included in the study. Of the participating residents, 33 (61.1%) were male, and 21 (38.9%) were female. The mean age was 27 ± 2 years, and the mean duration of residency was 10 ± 4 months.

A total of 194 pre-intervention and 141 post-intervention consultations were included in the study over the one month. Data regarding excluded consultations are presented in the flowchart (Figure 1).

Baseline encounter and requesting-physician characteristics before and after implementation are presented in Table 2. Requesting physician age and residency duration were similar between the two periods, whereas patient sex distribution and consultation request time differed significantly.

Table 2. Baseline encounter and requesting-physician characteristics before and after intervention.

Characteristic	Pre-intervention (n=194)	Post-intervention (n=141)	p value
Patient sex, male/female, n	115 / 79	63 / 78	0.011
Consultation request time, n (%)			0.043
00:00-08:00	57 (29.4)	34 (24.1)	0.043
08:00-16:00	40 (20.6)	18 (12.8)	0.651
16:00-00:00	97 (50.0)	89 (63.1)	
Requesting physician age, years, median (IQR)	27 (26-29)	27 (25-29)	
Residency duration, months, median (IQR)	10 (7-12)	10 (7-12)	0.495

p values were calculated using the chi-square test for categorical variables and the Mann-Whitney U test for continuous variables. The sex comparison p-value was calculated from the available counts. IQR, interquartile range.

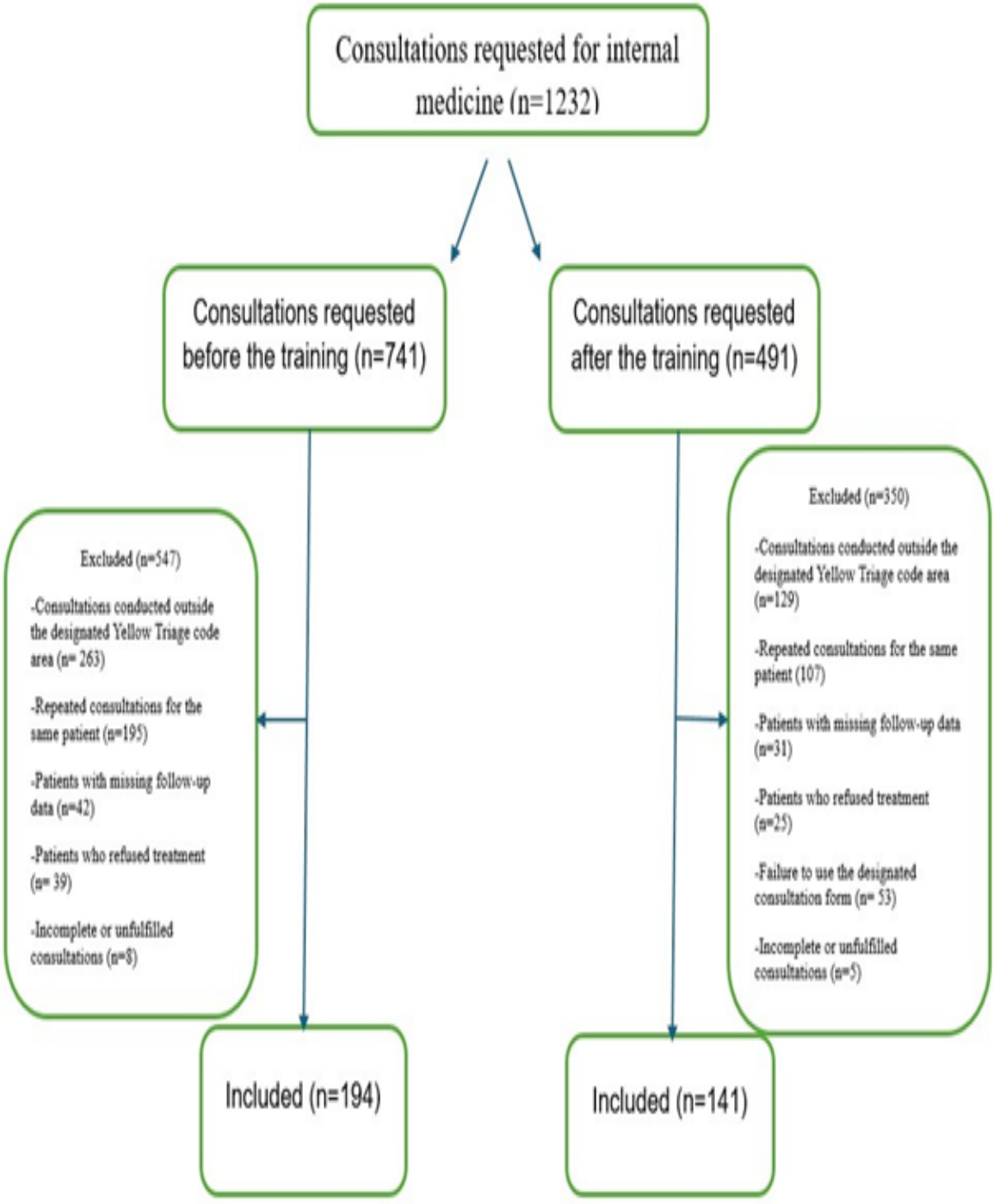


Figure 1. Flow diagram of included and excluded consultations.

The median consultation-to-decision time, which was the primary outcome, decreased from 65 minutes (IQR 40-112) in the pre-intervention period to 44 minutes (IQR 28-69) in the post-intervention period, corresponding to an absolute median reduction of 21 minutes ($p<0.001$) (Figure 2). Consultation response time was also significantly shorter in the post-intervention period, with an absolute reduction of 14 minutes ($p<0.001$). Final decision time also decreased significantly, from 8 hours to 5 hours, corresponding to an absolute median reduction of 3 hours ($p<0.001$). There was no significant difference in total ED LOS ($p=0.772$) (Table 3). Effect-size estimates are presented in Table 3. The largest effect was observed for consultation response time, while consultation-to-decision time and final decision time showed moderate effect sizes in favor of the post-intervention period. In contrast, the effect-size estimate for total ED LOS was negligible, consistent with the absence of a meaningful difference in overall ED stay between the two periods.

Data regarding secondary endpoints (consultation response time, final decision time, ED LOS, and repeat consultations) are presented in Figure 2. Repeat consultations were significantly lower in the post-intervention period. This difference was attributed to the distribution shown in Figure 3.

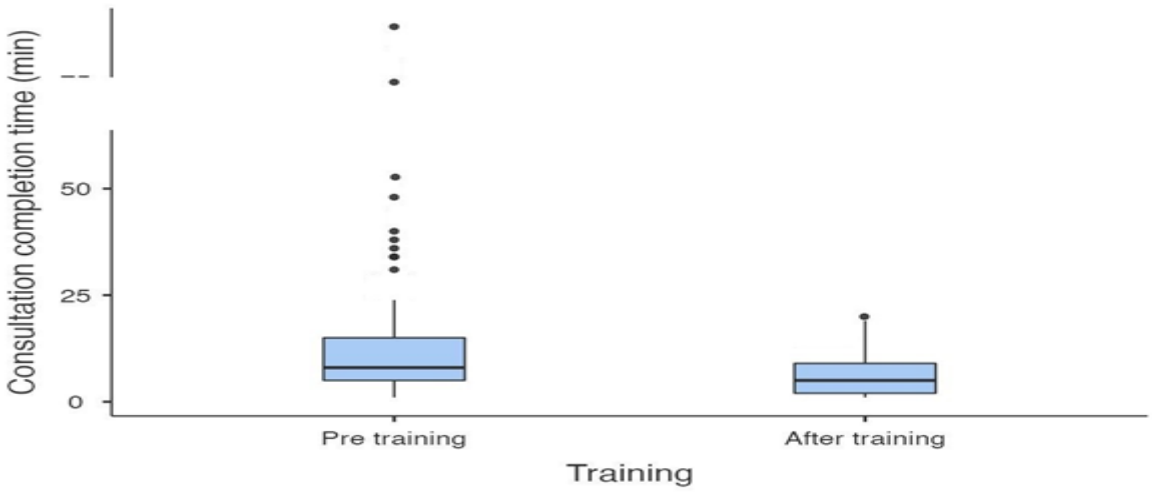


Figure 2. Consultation-related time outcomes before and after implementation of the digital 5C form.

Table 3. Main consultation outcomes before and after intervention.

Outcome	Pre-intervention (n=194)	Post-intervention (n=141)	Median difference (95% CI)	Effect size (95% CI)	p value
Consultation response time, min, median (IQR)	23 (12-43)	9 (5-18)	-14.0 (-17.73 to -10.27)	-0.503 (-0.605 to -0.397)	<0.001
Consultation-to-decision time, min, median (IQR)	65 (40-112)	44 (28-69)	-20.5 (-31.38 to -9.62)	-0.315 (-0.430 to -0.200)	<0.001
Final decision time, h, median (IQR)	8 (5-15)	5 (2-9)	-3.0 (-4.39 to -1.61)	-0.348 (-0.465 to -0.235)	<0.001
ED length of stay, h, median (IQR)	15 (9-26)	15 (9-24)	0.5 (-2.64 to 3.64)	-0.019 (-0.143 to 0.111)	0.772
Internal medicine reconsultation count per encounter, median (IQR)	1 (0-2)	1 (0-2)	0.0 (0.0 to 0.0)	-0.130 (-0.251 to -0.008)	0.035

Note: Repeat consultation was defined as any additional internal medicine consultation request recorded for the same patient during the same ED encounter after the initial consultation. The denominator was all eligible consultation encounters in each period: 194

before and 141 after implementation. Standardized effect sizes were calculated using rank-biserial correlation.

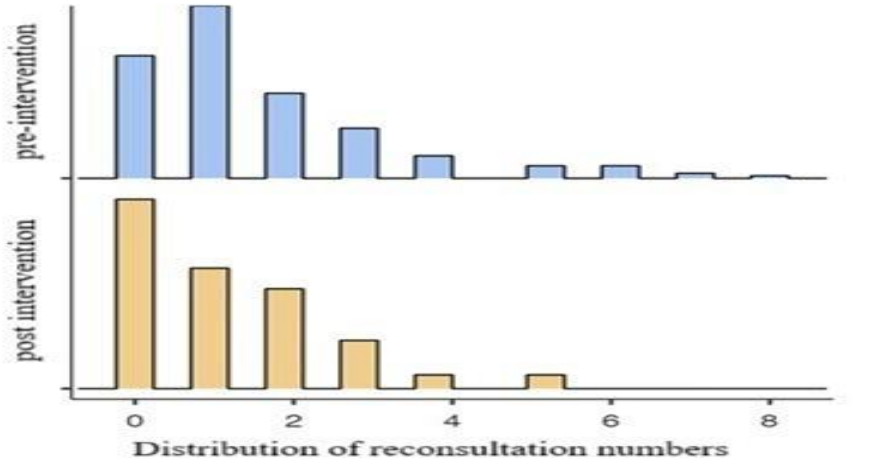


Figure 3. Distribution of repeat internal medicine consultations before and after implementation

Discussion

In this single-center quasi-experimental pre-post study, the combined intervention consisting of a four-component 5C-derived digital consultation form and targeted resident training was associated with shorter internal medicine consultation-related intervals. Median consultation-to-decision time decreased from 65 to 44 minutes, consultation response time decreased by 14 minutes, and final decision time decreased from 8 to 5 hours. Repeat consultations were also less frequent after the intervention. However, overall ED LOS remained unchanged. These findings suggest that the combined intervention may support consultation-related process efficiency, although causal conclusions cannot be drawn from the present study design.

The primary outcome of this study was the significant reduction in median consultation-to-decision time, which decreased from 65 minutes to 44 minutes following implementation of the four-component 5C-derived model. This approximately 32% improvement represents a clinically meaningful gain in ED efficiency. Our results are consistent with a recent systematic review by Beckerleg et al., which analyzed nine studies and concluded that workflow interventions-ranging from SMS messaging to process restructuring-consistently reduce consultation-to-decision times.[11] Specifically, our digital intervention aligns with findings by Ravi et al., who demonstrated that integrating consult orders directly into the electronic health record reduced consult response time by 15 minutes.[12] The 5C model was originally developed by Kessler et al. to optimize verbal communication; our findings suggest that a digital form derived from the 5C framework, when combined with resident training, may support more timely consultation processes.[3] However, because the digital form and resident training were implemented together, their individual contributions to the observed associations could not be determined. By embedding the 5C framework into a structured digital consultation template, we created a forcing function that promoted clear articulation of the core question at the earliest point of contact, reducing delays caused by back-and-forth clarification.

Beyond speed, the reduction in repeat consultations (p=0.035) suggests a potential improvement in communication clarity, likely by reducing the need for back-and-forth questioning. Previous educational interventions, such as those by Turner et al., failed to show significant improvements in consultation quality, suggesting that training alone may not change complex behaviors.[13] In contrast, our study mandated the structured entry of critical information, including Contact and Core Question elements. This supports the view of Baylis et al. that standardization mitigates ambiguity.[14] By digitally enforcing these parameters, our model minimized the information asymmetry that often leads to redundant re-evaluations and callbacks. Interestingly, despite shorter consultation and decision-making intervals, total ED LOS remained unchanged (median, 15 hours in both periods; p=0.772). This finding contrasts with Ravi et al., who reported a 52-minute reduction in ED LOS following implementation of an integrated consultation order.[12] More recent studies

have also reported improvements in ED LOS or other time-sensitive care intervals after implementing digital consultation and EHR-based consultation workflows.[15,16] The absence of a significant ED LOS difference in our study suggests that shorter consultation-related intervals may not translate into earlier physical departure when downstream factors, such as inpatient bed availability, boarding, and institutional admission or discharge processes, remain unchanged. Consultation delay represents only one component of the multifactorial process underlying ED crowding.[5]

Educational interventions may be affected by reliance on memory and possible skill decay. Previous studies involving medical students demonstrated the educational value of the original 5C consultation model.[17,18] In the present study, resident training was combined with a four-component 5C-derived digital consultation form embedded in HIMS. This combined approach may provide point-of-care prompts for structured consultation requests; however, its long-term sustainability and independent effects require further evaluation. From a family medicine perspective, a structured digital consultation form may support clearer documentation of acute interventions and follow-up needs during transitions from the ED to outpatient care. For family medicine residents undergoing ED rotations, a four-component 5C-derived digital consultation form combined with consultation training may also help standardize interdepartmental referral communication.

Limitations

Our study has several limitations inherent to its design. First, it was conducted in a single tertiary-care center and included only internal medicine consultations, which limits the generalizability of our findings to other ED settings and specialties. Second, the study used a nonrandomized pre-post intervention design rather than a randomized controlled trial. Because participants were not randomly assigned to intervention and control conditions, causal inference should be made cautiously, and unmeasured confounders may have influenced the observed associations. Therefore, several potential sources of bias should be considered when interpreting the results. A learning effect may have contributed to improved consultation performance after the training session, independent of the digital

form itself. In addition, a Hawthorne effect is possible, as residents who were aware of the intervention may have modified their consultation behavior during the post-intervention period. The pre-post design is also vulnerable to temporal bias, including differences in patient case mix, workload, staffing patterns, consultant availability, and inpatient bed capacity across the study periods. Furthermore, because the study primarily focused on consultation process metrics and physician workflow, only a limited set of patient-level baseline characteristics was available. Although patient sex was recorded, data on patient age, presenting diagnosis, comorbidities, clinical severity, and physiological parameters were not collected. Therefore, residual differences in patient case mix or complexity between the pre- and post-intervention periods cannot be excluded. Although we restricted data collection to a single month (July) to reduce seasonal variation, consultant rotation-related variability, and the potential effects of resident maturation, residual confounding cannot be excluded. Third, although repeat consultation frequency differed between the two periods, this variable was an indirect process measure, and we did not qualitatively assess consultation content, verbal interactions, or communication quality. Finally, our primary outcomes were process-based time measures, and we did not assess the direct impact of the intervention on patient-centered clinical outcomes such as mortality, readmission, or other downstream measures of care quality.

Conclusion

The combined intervention consisting of a four-component 5C-derived digital consultation form and targeted resident training was associated with shorter consultation-related intervals and may support consultation workflow efficiency in the ED. However, it was not associated with a significant improvement in overall ED LOS, indicating that optimizing consultation-related processes alone may be insufficient to address ED overcrowding. Multicenter studies, preferably using randomized or controlled designs, are needed to evaluate the independent and combined effects of such interventions on consultation processes and patient-centered clinical outcomes.

Ethical Considerations: The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Ethical approval was obtained from the Ankara Bilkent City Hospital Ethics Committee (approval no. E1-23-3508; approval date: May 10, 2023). Written informed consent was obtained from all participating physicians and patients.

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

References

1. Woods RA, Lee R, Ospina MB, et al. Consultation outcomes in the emergency department: exploring rates and complexity. CJEM. 2008;10(1):25-31 doi:10.1017/S1481803500009970.
2. Ozcete E, Kacmaz O, Damar O, Ustundag M, Orak M, Guloglu C. Evaluation of consultations requested from patients admitted to the emergency department. Eurasian J Emerg Med. 2015;14(3):113-7 doi:10.5152/jaem.2015.59672.
3. Kessler CS, Tadisina KK, Saks M, et al. The 5Cs of consultation: training medical students to communicate effectively in the emergency department. J Emerg Med. 2015;49(5):713-21 doi:10.1016/j.jemermed.2015.05.012.
4. Chalfin DB, Trzeciak S, Likourezos A, et al. Impact of delayed transfer of critically ill patients from the emergency department to the intensive care unit. Crit Care Med. 2007;35(6):1477-83 doi:10.1097/01.CCM.0000266585.74905.5A.
5. Voaklander B, Gaudet LA, Kirkland SW, et al. Interventions to improve consultations in the emergency department: a systematic review. Acad Emerg Med. 2022;29(12):1475-95 doi:10.1111/acem.14520.
6. Brick C, Lowes J, Lovstrom L, et al. The impact of consultation on length of stay in tertiary care emergency departments. Emerg Med J. 2014;31(2):134-8 doi:10.1136/emermed-2012-201908.
7. Kessler CS, Afshar Y, Sardar G, Yudkowsky R, Ankel F, Schwartz A. A prospective, randomized, controlled study demonstrating a novel, effective model of transfer of

care between physicians: the 5 Cs of consultation. Acad Emerg Med. 2012;19(8):968-74 doi:10.1111/j.1553-2712.2012.01412.x.

8. Holliman CJ. The art of dealing with consultants. J Emerg Med. 1993;11(5):633-40 doi:10.1016/0736-4679(93)90323-y.
9. Kessler CS, Asrow A, Beach C, et al. The taxonomy of emergency department consultations: results of an expert consensus panel. Ann Emerg Med. 2013;61(2):161-6 doi:10.1016/j.annemergmed.2012.07.122.
10. Donmez SS, Durak VA, Torun G, Koksal O, Aydin S. Acil Serviste Gerceklestirilen Konsultasyon Surecinin Incelenmesi [An Examination of the Consultation Process in the Emergency Department]. Uludag Univ Tip Fak Derg. 2017;43(1):23-8. Turkish.
11. Beckerleg W, Wooller K, Hasimjia D. Interventions to reduce emergency department consultation time: a systematic review of the literature. CJEM. 2020;22(1):56-64 doi:10.1017/cem.2019.435.
12. Ravi A, Shochat G, Wang RC, Khanna R. Improvements to emergency department length of stay and user satisfaction after implementation of an integrated consult order. J Am Coll Emerg Physicians Open. 2023;4(2):e12922 doi:10.1002/emp2.12922.
13. Turner J, Litzau M, Morgan ZS, Pollard K, Cooper DD. Failure of an educational intervention to improve consultation and implications for healthcare consultation. Cureus. 2019;11(2):e4105 doi:10.7759/cureus.4105.
14. Baylis J, Miloslavsky EM, Woods R, Chan TM. Conquering consultations: a guide to advances in the science of referral-consultation interactions for residency education. Ann Emerg Med. 2019;74(1):119-25 doi:10.1016/j.annemergmed.2018.12.004.
15. Alemseged NM, Getachew M, Woldasilase D, Wondimneh F, Weber EJ. Effectiveness of digital consultation in reducing emergency department length of stay. Front Emerg Med. 2025;9(3):e23 doi:10.18502/fem.v9i3.20019.
16. Iguidbashian J, Lun Z, Bata K, et al. Novel electronic health records-based consultation workflow improves time to operating room for vascular surgery patients in an acute setting. Ann Vasc Surg. 2023;97:139-46 doi:10.1016/j.avsg.2023.07.101.

17. Carter K, Golden A, Martin S, et al. Results from the first year of implementation of CONSULT: consultation with novel methods and simulation for UME longitudinal training. West J Emerg Med. 2015;16(6):845-50 doi:10.5811/westjem.2015.9.25520.
18. Lee-Miller C, Navarre B, Lorimer D, et al. Calling consults: a workshop to teach trainees using both didactic and small group-based learning. ATS Scholar. 2020;1(3):301-6 doi:10.34197/ats-scholar.2020-0021.

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):337-357 // 10.5505/amj.2026.01033

BEYOND THE INFECTION: DECODING THE BIOCHEMICAL AND HEMATOLOGICAL FINGERPRINT OF POST-COVID-19 HAIR LOSS

Hayriye Ortak Bulur[1], Özlem Güç Suvak[2], Muhammed İnan[2]

[1]Department of Family Medicine, Dışkapı Yıldırım Beyazıt Training and Research Hospital, University of Health Sciences, Ankara, Türkiye

[2]Department of Family Medicine, Ankara Etlik City Hospital, University of Health Sciences, Ankara, Türkiye

Correspondence:
Ozlem Güç Suvak (e-mail: ozlemsuvak@gmail.com)

Submitted: 06.04.2026 // Accepted: 11.07.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: Post-viral alopecia is a prevalent sequela of SARS-CoV-2 infection, yet its biochemical correlates remain poorly defined. This study aimed to elucidate the hematological and metabolic parameters associated with hair loss in patients with a history of COVID-19 compared to unexposed controls.

Materials and Methods: This retrospective case-control study included 366 adult outpatients presenting with alopecia (101 with prior SARS-CoV-2 infection; 265 without). Comprehensive clinical and laboratory data, encompassing blood counts, biochemical panels, and thyroid indices, were evaluated. Data were subjected to univariate and multivariable logistic regression models. A stringent Bonferroni correction was applied.

Results: Univariate comparisons revealed significantly elevated serum iron, AST, and MCV levels, alongside altered thyroid profiles in the COVID-19 (+) cohort ($p < 0.05$). However, following multivariable adjustment for baseline confounders, notably a paradoxically higher chronic comorbidity burden in the control group, the initial significance of these parameters was attenuated. Serum iron remained a robust, strictly hypothesis-generating, associative marker ($p < 0.001$), whereas TSH showed a marginal trend ($p = 0.058$).

Conclusion: The biochemical footprint of post-COVID-19 hair loss highlights altered iron homeostasis rather than sustained, broad-spectrum systemic inflammation. While our findings are hypothesis-generating rather than mechanically definitive due to the retrospective design, they underscore the clinical utility of targeted metabolic screening in primary care to optimize post-viral alopecia management.

Keywords: Post-COVID-19, hair loss, telogen effluvium, biomarkers, iron, TSH.

Introduction

The COVID-19 pandemic has generated a broad spectrum of lingering sequelae, with non-scarring hair loss (HL), predominantly telogen effluvium (TE)

emerging as a frequent complication. Although the precise pathophysiology linking SARS-CoV-2 to HL remains to be fully elucidated, stress, systemic inflammation, hormonal dysregulation, and metabolic shifts are hypothesized as main factors.[1,2] Non-scarring alopecias include alopecia areata, androgenetic alopecia, traction alopecia, tinea capitis, anagen effluvium, and telogen effluvium (TE), are those that do not cause permanent damage to the hair follicle.[3]

Despite the surge in recent literature, the mechanisms underlying the persistence of post-COVID-19 alopecia remain partially defined.[3,4] While hematological alterations have been documented in acute COVID-19, the role of pre-existing systemic factors in exacerbating post-viral hair loss lacks thorough investigation.[5] This study addresses this gap by evaluating specific biomarkers to elucidate the biological stressors involved, potentially informing more targeted clinical management strategies.

The biomarkers evaluated in this analysis serum iron, aspartate aminotransferase (AST), and thyroid hormones were selected based on their established associations with follicular homeostasis. While iron is fundamental for hair growth, aberrant elevations may reflect systemic oxidative stress, which is hypothesized to impair follicular function.[6] Furthermore, AST was included as a potential surrogate for systemic or hepatic stress—states frequently documented in SARS-CoV-2 infection, which may indirectly correlate with hair health.[3] Similarly, thyroid-stimulating hormone (TSH) and thyroxine (T4) were monitored to evaluate the impact of thyroid dysregulation, a recognized physiological trigger for hair cycle disruption.[2]

This study evaluates these biomarkers in a cohort of hair loss patients with a prior history of SARS-CoV-2 infection, compared to a control group of HL without such exposure. By

examining these parameters, we aim to elucidate the biological associations between prior infection and subsequent hair loss dynamics, potentially informing future clinical interventions and patient care.

Materials and Methods

Study participants

The study was conducted with data from patients who presented to the dermatology outpatient clinics with complaints of HL, esp. TE between April 1, 2020, and August 31, 2022.

Study population

Assuming a moderate effect size (Cohen's d = 0.6), a 5% significance level (two-tailed), and 80% power, with a 1:2 case-to-control ratio, a minimum of 38 cases and 76 controls (total n = 114) was required for adequate statistical power using the Mann-Whitney U test. Our final sample included 101 cases and 265 controls (total n = 366), and the final sample (n = 366) exceeded this requirement.

The study population comprised patients aged 18 years and older who presented to the dermatology outpatient clinics with primary complaints of HL between April 1, 2020, and August 31, 2022. Participants were categorized into two groups:

1. Case Group: Individuals with a documented history of SARS-CoV-2 infection, confirmed either by a positive PCR test or clinical diagnosis (ICD-10 codes U07.1, U07.2) within the six months preceding their dermatological consultation.

2. Control Group: Individuals presenting with HL who had no documented history of SARS-CoV-2 infection.

Comparative analysis of the study cohorts demonstrated no statistically significant differences regarding age, sex, or alopecia subtype ($p > 0.05$). Notably, the prevalence of

chronic comorbidities was marginally higher in the control group (p = 0.046) (Table 1). Exclusion criteria comprised age under 18 years, incomplete medical records, history of chemotherapy or radiotherapy within the preceding year, diagnosis of acute tinea capitis, and concurrent treatment for alopecia at the time of enrollment. The systematic flowchart detailing the patient selection process is presented in Figure 1.

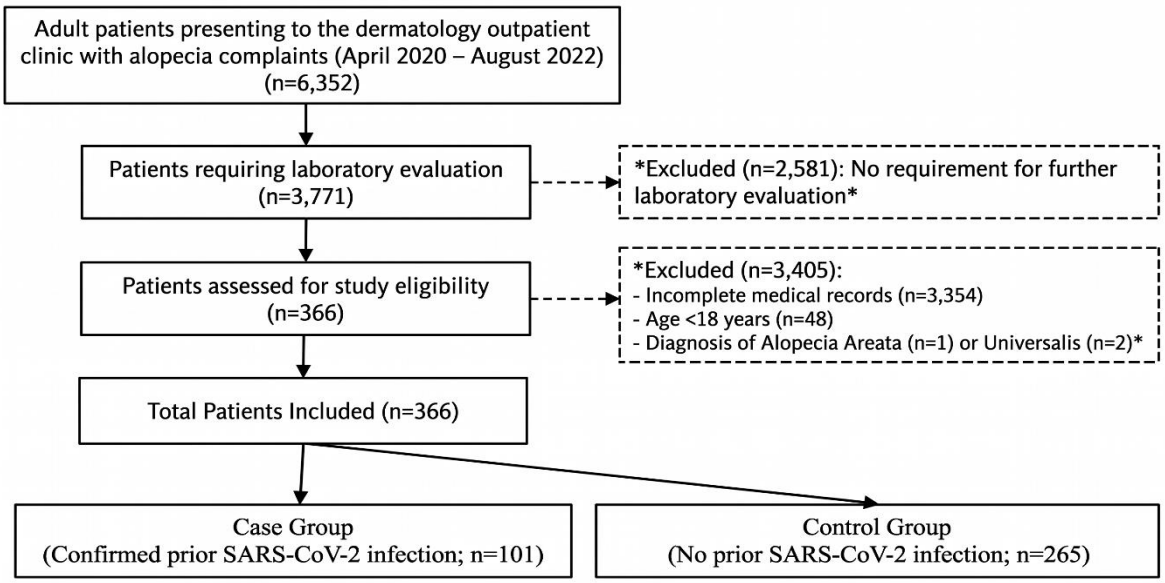


Figure 1. Systematic patient selection flowchart according to study criteria

Study design and Measurements

This study was designed as a single-center, retrospective case-control study to assess the association between prior SARS-CoV-2 infection and various HL etiologies, focusing on telogen effluvium. Reporting followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for case-control studies.

The outcome of interest was standardized by categorizing HL based on ICD-10 codes (L65.0, L63, L64), primarily focusing on Telogen Effluvium, which is the expected post-viral manifestation. Demographic information, comorbidities, and laboratory variables were extracted from the hospital's electronic health records and cross-referenced with national health databases. The recorded data included hemogram parameters (WBC, HGB, HCT, MCV,

RDW, PLT, MPV, NEU, LYM) and biochemical markers (urea, creatinine, AST, ALT, TSH, fT4, vitamin B12, folate, ferritin, iron, and total-binding iron capacity). All laboratory parameters were evaluated according to the standard reference intervals established by our institutional central laboratory. Specifically, the reference ranges were defined as follows: serum iron (60–170 µg/dL), AST (0–35 U/L), TSH (0.4–4.0 mIU/L), and MCV (80–100 fL). In the case group, blood samples were generally obtained within 1–6 months following COVID-19 diagnosis. The onset of HL was typically reported within 2–12 weeks prior to the dermatology visit, consistent with the expected latency period of telogen effluvium HL onset, and blood sampling varied among patients. To reduce variability, the laboratory data closest to the initial dermatological consultation were prioritized for analysis.

Ethical Approval and Informed Consent

The study was approved by the University of Health Sciences Ankara Dışkapı Yıldırım Beyazıt Training and Research Hospital ethics committee (Protocol No 140/34, 20.06.2022) and conducted in accordance with the 1964 Helsinki Declaration and its subsequent amendments. Written informed consent was obtained from all participants before their enrollment.

Statistical analysis

Statistical analyses were performed using the IBM SPSS 20.0 statistical software package. Data normality was assessed via the Shapiro–Wilk normality test and Q–Q plots; group variance homogeneity was checked via Levene's homogeneity test. Power analysis (calculated using G*Power) indicated that a total sample of 114 participants (1:2 ratio) was required to detect a moderate effect size (Cohen's d = 0.6) with 80% power at alpha = 0.05. Continuous variables are presented as mean ± standard deviation or median (interquartile range [IQR]) based on normality; categorical variables are expressed as frequencies and percentages. Missing values for TSH (n=1), B12 (n=1), folate (n=6), and ferritin (n=1) within the 366-patient cohort were addressed via median imputation to preserve statistical power and mitigate outlier influence in multivariate modeling. Furthermore, laboratory parameters

were standardized to ensure analytical consistency. A p-value less than 0.05 was considered statistically significant. For univariate comparisons, the Mann-Whitney U test and Chi-square test were utilized. Predictors of prior SARS-CoV-2 infection were assessed across 22 clinical and laboratory parameters using univariate logistic regression. To counteract Type I error inflation from multiple comparisons, a Bonferroni correction was applied, establishing a strictly adjusted significance threshold of $p < 0.00227$ (0.05/22). To ensure comprehensive adjustment for potential confounders within the multivariate logistic regression framework, variable selection was guided by the Hosmer-Lemeshow criterion. Specifically, candidate predictors that demonstrated a potential association in the preliminary univariate analysis at a liberal significance threshold of $p < 0.25$ were selected for inclusion. Consequently, age, serum iron, and TSH were incorporated into the final multivariate model.

Results

A total of 366 participants (101 cases and 265 controls) were included in the study. The descriptive analysis is outlined in Table 1. Preliminary univariate analysis (Table 1) revealed no statistically significant differences between the case and control groups regarding gender, age distribution, or specific HL subtypes ($p > 0.05$ for all), except for chronic diseases in the case group ($\chi 2$: 3.984; $p = 0.046$). To rigorously address the potential confounding introduced by different alopecia phenotypes, a dedicated subgroup analysis was restricted to patients explicitly diagnosed with Telogen Effluvium (L65, n=296). Within this homogeneous TE cohort, serum iron levels remained robustly and significantly elevated in the prior SARS-CoV-2 infection group (n=81) compared to the uninfected controls (n=215) [$p < 0.001$ (Mann-Whitney U test), confirming that this biochemical alteration is intrinsically linked to the post-viral TE presentation rather than diagnostic heterogeneity.

Table 1. Baseline Demographic and Clinical Characteristics of the Study Population

		Control Group (COVID-19 Negative)		Case Group (COVID-19 Positive)		$\chi 2$	p-value
Variables		n	%	n	%		
Sex	Female	207	71,6	82	28,4	0,252	0,616
	Male	58	75,3	19	24,7		
Age group	≤25	102	75,0	34	25,0	1,039	0,595
	26-35	77	72,6	29	27,4		
	> 35	86	69,4	38	30,6		
Comorbidities	Yes	62	64,6	34	35,4	**3,984**	**0,046**
	No	203	75,2	67	24,8		
Alopecia Subtype	L63*	35	71,4	14	28,6	0,097	0,953
	L64*	14	70,0	6	30,0		
	L65*	216	72,7	81	27,3		

**L63: Alopecia Areata, L64: Androgenetic Alopecia, L65: Non-Scarring Hair Loss, χ2: Chi-square, p<0.05 is considered statistically significant.*

A prior diagnosis of COVID-19 was significantly associated with higher levels of AST and TSH, as well as lower levels of T4 ($p < 0.05$ for each). Additionally, the study discovered that serum iron measurements were notably higher in HL patients who had experienced COVID-19 ($p < 0.001$). A comparative analysis of biochemical and hematological markers is detailed in Tables 2 and 3.

ANKARA MEDICAL JOURNAL

Table 2. Comparison of Laboratory Parameters Between COVID-19 (+) and Control Groups

Parameter		n	%	Control Group (COVID-19 Negative)		Case Group (COVID-19 Positive)		χ2	p
				n	%	n	%		
Urea (17 – 43 mg/dL)	Low	37	10,1	29	78,4	8	21,6	1,947	0,378
	Normal	320	87,4	231	72,2	89	27,8		
	High	9	2,5	5	55,6	4	44,4		
Creatinine (0.5 – 1.2 mg/dL)	Low	14	3,8	11	78,6	3	21,4	0,423	0,809
	Normal	343	93,7	247	72,0	96	28,0		
	High	9	2,5	7	77,8	2	22,2		
AST (0–35 U/L)	Low	-	-	-	-	-	-	**5,403**	**0,020**
	Normal	356	97,3	261	73,3	95	26,7		
	High	**10**	**2,7**	**4**	**40,0**	**6**	**60,0**		
ALT (0 – 35 U/L)	Low	-	-	-	-	-	-	0,060	0,806
	Normal	339	92,6	246	72,6	93	27,4		
	High	27	7,4	19	70,4	8	29,6		
TSH (0.4–4.0 mIU/L)	Low	1	0,3	0	0,0	1	100,0	**6,215**	**0,045**
	Normal	349	95,4	247	74,0	87	26,0		
	High	**5**	**1,4**	**18**	**58,1**	**13**	**41,9**		
fT4 (0.93 – 1.70 ng/dL)	**Low**	**12**	**3,3**	**4**	**33,3**	**8**	**66,7**	**9,579**	**0,008**
	Normal	349	95,4	257	73,6	92	26,4		
	High	5	1,4	4	80,0	1	20,0		
Vitamin B12 (197 – 771 pg/mL)	Low	24	6,6	20	83,3	4	16,7	1,542	0,463
	Normal	331	90,4	237	71,6	94	28,4		
	High	11	3,0	8	72,7	3	27,3		
Folate (3.1 – 17.5 ng/mL)	Low	11	3,0	6	54,5	5	45,5	2,704	0,259
	Normal	348	95,1	255	73,3	93	26,7		
	High	7	1,9	4	57,1	3	42,9		

**AST: Aspartate Aminotransferase Test, ALT: Alanine Aminotransferase Test, TSH: Thyroid-stimulating hormone, χ2: Chi-square, p<0.05 is considered statistically significant*

Regarding hematological indices, although the median MCV was similar between the case group (86.9, IQR: 83.3–89.7) and the control group (86.6, IQR: 83.9–88.9) (p = 0.554). However, a categorical analysis of MCV (low, normal, high) revealed significant differences between groups ($\chi 2$=11.080, p=0.004)(Table 3).

Table 3. Comparison of Laboratory Parameters Between COVID-19 (+) and Control Groups-2

Parameters		**n**	**%**	**Control Group (COVID-19 Negative)**		**Case Group (COVID-19 Positive)**		**χ^2**	**p**
				n	**%**	**n**	**%**		
Ferritin (13-400 ng/mL)	Low	107	29,2	78	72,9	29	27,1		
	Normal	250	68,3	181	72,4	69	27,6	0,161	0,923
	High	9	2,5	6	66,7	3	33,3		
Serum Iron (60–170 µg/dL)	Low	202	55,2	168	83,2	34	16,8		
	Normal	164	44,8	97	59,1	67	40,9	**26,141**	**<0.001**
	High	-	-	-	-	-	-		
Total Iron–Binding Capacity (TIBC) (250-450 µg/dL)	Low	3	0,8	2	66,7	1	33,3		
	Normal	290	79,2	212	73,1	78	26,9	0,356	0,837
	High	73	19,9	51	69,9	22	30,1		
WBC (4.5-11.0 x 10^3/µL)	Low	-	-	-	-	-	-		
	Normal	354	96,7	259	73,2	95	26,8	3,177	0,077
	High	12	3,3	6	50,0	6	50,0		
HGB (12.0-16.0 g/dL)	Low	28	7,7	22	78,6	6	21,4		
	Normal	322	88,0	232	72,0	90	28,0	0,660	0,719
	High	16	4,4	11	68,8	5	31,2		
HTC (36-46%)	Low	13	3,6	10	76,9	3	23,1		
	Normal	312	85,2	223	71,5	89	28,5	0,922	0,631
	High	41	11,2	32	18,0	9	22,0		
MCV (80-100 fL)	Low	8	2,2	5	62,5	3	37,5		
	Normal	354	96,7	260	73,4	94	26,6	**11,080**	**0,004**
	High	4	1,1	0	0,0	4	100,0		
	Low	66	18,0	46	69,7	20	30,3	0,312	0,856

RDW (11.5 – 14.5%)	Normal	292	79,8	213	72,9	79	27,1		
	High	8	2,2	6	75,0	2	25,0		
	Low	-	-	-	-	-	-		
PLT (150-400x10^3/µL)	Normal	335	91,5	244	72,8	91	27,2	0,368	0,544
	High	31	8,5	21	67,7	10	32,3		
	Low	-	-	-	-	-	-		
MPV (7.4 – 10.4fL)	Normal	338	92,3	245	72,5	93	27,5	0,014	0,904
	High	28	7,7	20	71,4	8	28,6		
	Low	-	-	-	-	-	-		
NEU (2.0-7.0 x10^3/µL)	Normal	358	97,8	261	72,9	97	27,1	0,224*	
	High	8	2,2	4	50,0	4	50,0		
	Low	-	-	-	-	-	-		
LYM (1.0-4.0 x10^3/µL)	Normal	300	82,0	218	72,7	82	27,3	0,008	0,930
	High	66	18,0	47	71,2	19	18,8		
Total (n)		366	100,0	265	72,4	101	27,6		

**Fisher's Exact Test, χ2: Chi-square, p<0.05 is considered statistically significant. WBC: White Blood Cell, HGB: Hemoglobin, HTC: Hematocrit, MCV: Mean Corpuscular Volume, RDW: Red Cell Distribution Width, PLT: Platelet, MPV: Mean Platelet Volume, NEU: Neutrophil, LYM: Lymphocyte*

Univariate analysis revealed that serum iron was the sole parameter significantly associated with prior SARS-CoV-2 infection. Notably, this association retained its significance even after the application of the highly conservative Bonferroni correction (p=0.000054), thereby demonstrating the profound statistical robustness of this relationship. (Table 4)

Table 4. Univariate Logistic Regression Analysis of Selected Parameters Associated with Prior SARS-CoV-2 Infection

Variable	Odds Ratio (OR)	95% CI	p-value
Serum Iron	1.001	1.0005 – 1.0015	**< 0.001***
TSH	1.158	0.9914 – 1.3529	0.064
Age	1.011	0.9927 – 1.0313	0.224
AST	1.015	0.9880 – 1.0442	0.267
Urea	1.017	0.9834 – 1.0527	0.317
Sex (Male)	0.826	0.4640 – 1.4736	0.519
Ferritin	1.001	0.9970 – 1.0053	0.577

OR: Odds Ratio; CI: Confidence Interval; TSH: Thyroid-Stimulating Hormone; AST: Aspartate Aminotransferase; $p < 0.05$ *is considered statistically significant.*

**Indicates the variable that retained statistical significance following Bonferroni correction (adjusted threshold: $p < 0.00227$). All other hematological and biochemical indices failed to reach statistical significance ($p > 0.50$)*

The independent biomarkers demonstrating clinical and statistical relevance based on prior SARS-CoV-2 infection were illustrated in Figure 2.

In the multivariable logistic regression model, serum iron remained independently associated with prior SARS-CoV-2 infection after adjusting for age and TSH (aOR = 1.001; 95% CI, 1.0005–1.0015; $p < 0.001$). TSH exhibited a trend toward marginal significance ($p = 0.058$). Although the odds ratio for serum iron reflects a per-unit increment, it signifies a clinically relevant cumulative effect, as a 50-unit discrepancy corresponds to an approximate 5% shift in relative risk.

ANKARA MEDICAL JOURNAL

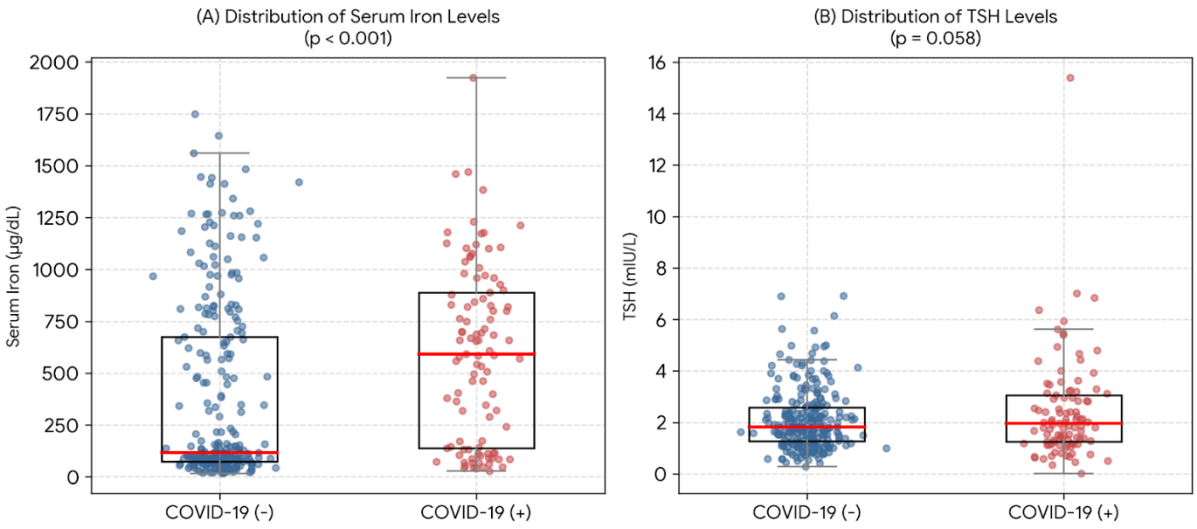


Figure 2. Box-and-strip plots of laboratory parameters demonstrating clinical and statistical relevance based on prior SARS-CoV-2 infection. Red horizontal bars represent median values, while box boundaries denote the interquartile range (IQR; 25th and 75th percentiles). (A) Serum iron levels were significantly elevated in the COVID-19 (+) participants (B) TSH measurements exhibited a marginal trend toward differentiation between the groups

Discussion

This study aimed to investigate the relationship between COVID-19 infection and HL. The present study suggests that serum iron levels may serve as an independent biochemical correlate in HL patients with a prior history of SARS-CoV-2 infection. This observed association demonstrated notable statistical robustness, persisting not only through stringent Bonferroni correction for multiple comparisons but also remaining significant within multivariable models adjusted for potential confounders such as age and TSH.

Hair loss following COVID-19 is a widely reported phenomenon that encompasses a heterogeneous spectrum of clinical diagnoses. While it frequently presents with features consistent with TE, the viral prodrome has also been linked to the exacerbation of androgenetic alopecia and the relapse of alopecia areata.[2,7,8] Diffuse HL and especially TE is typically observed within weeks to months following the infectious episode in this context.[9,10] Given this diagnostic heterogeneity, post-viral alopecia is more accurately

conceptualized as a broad clinical manifestation rather than a singular, uniformly driven entity.[7,8] The inclusion of diverse phenotypes, such as Alopecia Areata (L63) and Androgenetic Alopecia (L64) alongside Telogen Effluvium (L65), mirrors the inherently heterogeneous nature of post-viral alopecia. SARS-CoV-2-precipitated systemic inflammation and metabolic stress do not merely induce acute TE; they can also exacerbate alternative hair loss modalities via immune dysregulation and follicular microinflammation. Consequently, investigating this expansive spectrum provides a more comprehensive framework for capturing the real-world clinical landscape encountered in primary care.

Consequently, any accompanying biochemical or hematological alterations may largely reflect transient, generalized physiological responses to the preceding infection, rather than functioning as independent, causative drivers of the HL process.

Serum higher iron measurements were demonstrated to have a significant correlation with prior SARS-CoV-2 infection. A notable limitation of our study design is the utilization of a control group consisting of individuals with pre-existing alopecia rather than healthy, non-alopecic subjects. Because the biomarkers under investigation, such as iron and thyroid indices, are intrinsically linked to follicular dysregulation, accurately discerning the specific pathogenetic impact of the virus is inherently complex. Despite this confounding baseline, the associative risks identified within our logistic regression framework underscore a potential virus-driven metabolic shift that warrants careful clinical consideration. Rather than indicating improved iron status, these elevations likely reflect inflammation-induced dysregulation, where pro-inflammatory cytokines impair iron sequestration. This shift may trigger a premature transition of hair follicles into the telogen phase.[10,11] However, since laboratory assessments in our cohort were conducted at highly variable intervals (1 to 6 months) post-infection, biological inferences regarding iron metabolism must be interpreted with caution. Serum iron is profoundly time-dependent and heavily influenced by fluctuating acute-phase responses during post-infectious recovery.[12] While previous studies have reported an association with iron deficiency and HL in COVID-19.[13,14] consequently, the

observed elevations may largely represent transient acute-phase reactions rather than a persistent or direct mechanistic driver of follicular disruption.

Another notable finding was the elevated TSH levels observed in patients with a prior SARS-CoV-2 infection. This aligns with literature suggesting that thyroid dysfunction may develop following COVID-19 due to immune-mediated thyroid inflammation.[2] Nevertheless, the interpretation of thyroid panel alterations is intrinsically limited by our study's temporal heterogeneity. Thyroid hormones are sensitive to non-thyroidal illness syndrome and undergo significant fluctuations during acute and convalescent phases. Given the variable timeline between viral clearance and blood sampling, the observed TSH elevations might reflect disparate stages of post-viral endocrine recovery rather than a uniform pathophysiological trigger for premature telogen transition. [17,18]

Post-viral sequelae, systemic inflammation, and oxidative stress have been hypothesized to influence proteoglycan metabolism, which plays a supportive role in maintaining the anagen phase of the hair cycle. It is plausible that such metabolic disturbances may indirectly contribute to the prolonged HL observed following SARS-CoV-2 infection.[19] Concurrently, the marginal TSH elevation observed in the multivariable model warrants cautious interpretation. Given the heterogeneous timing of laboratory assessments, this subtle endocrine fluctuation likely reflects transient post-infectious systemic stress rather than a direct mechanistic trigger for HL. Consequently, this relationship remains strictly associative and necessitates prospective investigation within the long COVID-19 framework.[1]

Elevated AST levels in the case group, while not directly attributable to HL in extant literature, suggest possible liver dysfunction or systemic inflammation. AST is a marker of liver stress, which has been reported in COVID-19 patients due to widespread inflammatory damage.[20,21] Nonetheless, AST is a highly dynamic enzyme with a time-dependent serum half-life. Given the broad window of laboratory assessment relative to the onset of SARS-CoV-2 infection and subsequent HL, these elevated levels might represent recovering acute hepatic injury from the viral prodrome.[21] Therefore, projecting a definitive mechanistic link between transient transaminitis and delayed HL remains methodologically challenging.[22,23]

Although serum iron demonstrated a robust independent association and TSH exhibited a marginal trend in our models, these retrospective observations remain strictly hypothesis-generating and necessitate prospective validation. Furthermore, our cohort was predominantly composed of mild-to-moderate COVID-19 cases, a clinical demographic in which profound hematological dysregulation is intrinsically less frequent.[24] This milder disease trajectory, coupled with the previously discussed temporal heterogeneity in laboratory sampling, likely accounts for the absence of significant variations in established systemic inflammatory indices, such as the neutrophil-to-lymphocyte (NLR) and platelet-to-lymphocyte (PLR) ratios.[25]

Another notable methodological constraint is the wide temporal variability (1 to 6 months) concerning laboratory sampling following the initial SARS-CoV-2 infection. Because biomarkers such as serum iron and thyroid indices are highly dynamic and exquisitely sensitive to the resolving phases of systemic inflammation (e.g., convalescent acute-phase responses or non-thyroidal illness syndrome), this lack of temporal standardization precludes the establishment of definitive causality. The observed biochemical fluctuations may therefore represent disparate stages of post-viral physiological recovery rather than a stable, uniform pathogenic trigger for alopecia.

Additionally, a nominal elevation in mean corpuscular volume (MCV) was observed among patients with a history of SARS-CoV-2. Rather than indicating a persistent hematological defect, this observation might reflect transient, inflammation-induced alterations in erythropoiesis.[26] In a parallel context, systemic immune activation following vaccination has also been implicated in episodes of HL.[9,10] Although the literature suggests that infection-induced alopecia is generally more severe than its vaccine-induced counterpart, presumably mirroring a more intense inflammatory cascade, we could not adjust for this variable. Since our data collection predated widespread immunization initiatives, vaccination status constitutes an unassessed confounder in the present analysis.

Furthermore, the absence of significant differences across the remaining biochemical parameters is likely a direct consequence of the aforementioned temporal heterogeneity in

laboratory sampling. Because these systemic markers are inherently time-dependent and highly sensitive to varying stages of post-infectious recovery, future investigations must strictly standardize evaluation timelines while robustly controlling for confounding variables, including specific pharmacotherapies (e.g., corticosteroids, antivirals), baseline comorbidities, and psychosocial stress.

Importantly, the multivariate logistic regression model, adjusted for age, sex, and chronic comorbidities, demonstrated that no clinical or biochemical variables remained as independent predictors of previous SARS-CoV-2 infection. While initial univariate analyses highlighted significant intergroup differences across several biochemical parameters, the complete attenuation of these associations in the multivariate model is a critical methodological finding. It strongly suggests that the laboratory abnormalities observed univariately were predominantly driven by confounding demographic and clinical factors. Consequently, our results preclude any definitive causal conclusions. The initial univariate variations should not be interpreted as specific mechanistic drivers of conditions such as post-COVID-19 telogen effluvium; rather, they are strictly hypothesis-generating. These laboratory deviations might merely reflect a non-specific surrogate for broader physiological perturbations, potentially suggesting a loose association with generalized systemic inflammation or metabolic stress rather than a direct etiologic pathway.

Regarding the baseline characteristics of our study population, a paradoxical imbalance was observed. Chronic comorbidities were more prevalent in the control group, introducing potential confounding for markers like AST and TSH. However, this baseline imbalance underscores the robustness of our findings. Since pre-existing conditions typically elevate these stress biomarkers, their significant differentiation in the COVID-19 (+) cohort suggests the post-viral impact supersedes baseline metabolic stressors. Nonetheless, this clinical heterogeneity warrants cautious interpretation.

Consequently, while these biochemical and hematological deviations cannot be independently attributed to a specific post-COVID-19 etiology, their recognition holds substantial value in everyday clinical practice. The data tentatively support the broader

concept that a milieu of systemic inflammation and generalized stress may parallel episodes of HL. In the realm of family medicine, proactively monitoring iron stores, thyroid indices, and liver enzymes in affected individuals is highly advisable. Such assessments serve not to establish a definitive causal link to prior infection, but rather to ensure a well-rounded, multidimensional approach to patient evaluation and overall management.

Limitations

Limitations of this study include its retrospective, single-center design, and the lack of adjustment for unmeasured confounders, such as disease severity, dietary shifts, acute COVID-19 pharmacotherapies, and psychosocial stressors, which may independently alter hair cycle dynamics. A core methodological constraint is the utilization of a control group consisting of individuals with pre-existing alopecia rather than healthy, non-alopecic subjects; because the biomarkers under investigation are intrinsically linked to baseline follicular dysregulation, isolating the virus-specific pathogenetic impact is inherently complex. Furthermore, the significant timing heterogeneity among SARS-CoV-2 infection, the clinical onset of hair loss, and the subsequent laboratory evaluations fundamentally constrains biological interpretability. Since key biomarkers (e.g., serum iron, ferritin, AST, and thyroid hormones) are strictly time-dependent and highly susceptible to post-infectious recovery dynamics, this temporal variability precludes the establishment of a definitive pathogenetic timeline. Consequently, although a stringent Bonferroni correction mitigated Type I error inflation during univariate analyses, the complete attenuation of independent predictors within the multivariable framework necessitates a cautious, strictly hypothesis-generating interpretation of the findings.

Conclusion

In conclusion, this retrospective analysis identifies notable univariate associations between prior SARS-CoV-2 infection and specific hematological and biochemical alterations, particularly serum iron and TSH dynamics, in patients presenting with alopecia. Nevertheless, the lack of independent predictors within the multivariate logistic regression

framework underscores the complex, confounding interplay of pre-existing comorbidities and baseline physiological variations. Therefore, rather than establishing definitive mechanistic pathways, our findings are strictly hypothesis-generating. The subtly altered biochemical footprint delineated here suggests a potential post-viral metabolic shift that warrants further investigation. From a Family Medicine perspective, primary care serves as the crucial gateway for the early detection and longitudinal management of such post-viral sequelae. Since alopecia frequently prompts the initial clinical visit, family physicians hold a pivotal role in the holistic evaluation and triage of these patients. Integrating targeted metabolic assessments into first-line evaluations can optimize early intervention and alleviate the diagnostic burden. Ultimately, future prospective, multicenter studies with rigorous control for temporal and demographic confounders are warranted to fully elucidate the specific dynamics of post-viral alopecia and refine clinical management strategies.

Ethical Considerations: Approval was received from the Ethics Committee of the University of Health Sciences Dışkapı Yıldırım Beyazıt Training and Research Hospital (Protocol No 140/34, 20.06.2022) and conducted in accordance with the 1964 Helsinki Declaration and its subsequent amendments.

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

Acknowledgements: We would like to thank our clinical training coordinators of Family Medicine and Dermatology for supporting the study.

References

1. Su S, Zhao Y, Zeng N, et al. Epidemiology, clinical presentation, pathophysiology, and management of long COVID: an update. Mol Psychiatry. 2023;28(10):4056-69. https://doi.org/10.1038/s41380-023-02171-3
2. Popescu M, Berteanu M, Beiu C, et al. Complementary Strategies to Promote Hair Regrowth in Post-COVID-19 Telogen Effluvium. Clin Cosmet Investig Dermatol. 2022;15:735-43. https://doi.org/10.2147/CCID.S359052
3. Gadzhigoroeva A, Guzman-Sanchez D, Firooz A, et al. COVID-19 can exacerbate pattern hair loss and trigger telogen effluvium: the role of proteoglycan replacement therapy with Nourkrin® in clinical treatment of COVID-19 associated hair loss. J Dermatol Res Ther. 2021;7:103. https://doi.org/10.23937/2469-5750/1510103
4. Abushukur Y, Mansour M, Rehman R, et al. A systematic review of hair loss as a consequence of COVID-19 infection. Int J Dermatol. 2023;62(8):e436-39. https://doi.org/10.1111/ijd.16542
5. Shreewastav RK, Singh G, Jha KK, et al. Study on status of serum biochemical and hematological parameters in COVID-19 positive patients attending a tertiary care hospital. Kathmandu Univ Med J. 2022;20(79):295-300. https://doi.org/10.3126/kumj.v20i3.53934
6. Holzer I, Ott J, Beitl K, et al. Iron status in women with infertility and controls: a case-control study. Front Endocrinol (Lausanne). 2023;14:1173100. https://doi.org/10.3389/fendo.2023.1173100
7. Abrantes T, Artounian K, Falsey RR, et al. Time of onset and duration of post-COVID-19 acute telogen effluvium. J Am Acad Dermatol. 2021;85(4):975-76. https://doi.org/10.1016/j.jaad.2021.07.021
8. Yıldırım SK, Erbağcı E, Öğüt ND. Evaluation of patients with telogen effluvium during the pandemic: May the monocytes be responsible for post COVID-19 telogen effluvium? J Cosmet Dermatol. 2022;21(4):1339-44. https://doi.org/10.1111/jocd.14883

9. Alharbi M. Telogen effluvium after COVID-19 vaccination among public in Saudi Arabia. J Family Med Prim Care. 2022;11(10):6056-60. https://doi.org/10.4103/jfmpc.jfmpc_377_22
10. Ramya Sree A, Sethumadhavan K, Pullakanam SPT, et al. Evaluation of COVID-19 infection among vaccinated and unvaccinated individuals using biochemical markers. Bioinformation. 2024;20(3):223-28. https://doi.org/10.6026/973206300200223
11. Aryanian Z, Balighi K, Hatami P, et al. The role of SARS-CoV-2 infection and its vaccines in various types of hair loss. Dermatol Ther. 2022;35(6):e15433. https://doi.org/10.1111/dth.15433
12. Henry BM, de Oliveira MHS, Benoit S, et al. Hematologic, biochemical and immune biomarker abnormalities associated with severe illness and mortality in COVID-19: a meta-analysis. Clin Chem Lab Med. 2020;58(7):1021-28. https://doi.org/10.1515/cclm-2020-0369
13. Park SY, Na SY, Kim JH, et al. Iron Plays a Certain Role in Patterned Hair Loss. J Korean Med Sci. 2013;28(6):934-38. https://doi.org/10.3346/jkms.2013.28.6.934
14. Edeas M, Saleh J, Peyssonnaux C. Iron: Innocent bystander or vicious culprit in COVID-19 pathogenesis? Int J Infect Dis. 2020;97:303-5. https://doi.org/10.1016/j.ijid.2020.05.110
15. Ali SY, Fatima U. Serum ferritin levels and Diffuse hair loss—A correlation. IP Indian J Clin Exp Dermatol. 2020;6(3):212-14. https://doi.org/10.18231/j.ijced.2020.055
16. Guarnieri G, Marchi LBD, Marcon A, et al. Relationship between hair shedding and systemic inflammation in COVID-19 pneumonia. Ann Med. 2022;54(1):904-10. https://doi.org/10.1080/07853890.2022.2054026
17. Gao W, Guo W, Guo Y, et al. Thyroid hormone concentrations in severely or critically ill patients with COVID-19. J Endocrinol Invest. 2021;44(7):1031-40. https://doi.org/10.1007/s40618-020-01460-w

18. Wang W, Su X, Ding Y, et al. Thyroid function abnormalities in COVID-19 patients. Front Endocrinol (Lausanne). 2020;11:623792. https://doi.org/10.3389/fendo.2020.623792
19. Khalid A, Jaffar MA, Khan T, et al. Hematological and biochemical parameters as diagnostic and prognostic markers in SARS-CoV-2 infected patients of Pakistan: a retrospective comparative analysis. Hematology. 2021;26(1):529-42. https://doi.org/10.1080/16078454.2021.1950898
20. Effenberger M, Grander C, Grabherr F, et al. Systemic inflammation as fuel for acute liver injury in COVID-19. Dig Liver Dis. 2021;53(2):158-65. https://doi.org/10.1016/j.dld.2020.08.004
21. Bertolini A, van de Peppel IP, Bodewes F, et al. Abnormal liver function tests in patients with COVID-19: relevance and potential pathogenesis. Hepatology. 2020;72(5):1864-72. https://doi.org/10.1002/hep.31480
22. Farías JP, Codes L, Vinhaes D et al. Impact of baseline abnormal liver enzymes in the outcome of COVID-19 infection. Transl Gastroenterol Hepatol. 2021;8:41. https://doi.org/10.21037/tgh-22-41
23. Liakina V, Stundienė I, Milaknytė G, et al. Effects of COVID-19 on the liver: the experience of a single center. World J Gastroenterol. 2022;28(39):5735-49. https://doi.org/10.3748/wjg.v28.i39.5735
24. Shahcheraghi S, Hamidfar M, Heidari F, et al. Albumin, WBC, RBC, hemoglobin, platelet, MCV and MCH levels in COVID-19 patients with mucormycosis opportunistic infection. J Med Res. 2022;8(3):115-19.
25. Pujani M, Raychaudhuri S, Singh M, et al. An analysis of hematological, coagulation and biochemical markers in COVID-19 disease and their association with clinical severity and mortality: an Indian outlook. Am J Blood Res. 2021;11(6):580–91.
26. Berzuini A, Bianco C, Paccapelo C, et al. Red cell distribution width and mean corpuscular volume: a simple parameter with a complex significance in COVID-19 patients. Blood Adv. 2020;4(7):2040–42. https://doi.org/10.1182/bloodadvances.2020002657

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):358-380 // 10.5505/amj.2026.74271

VACCINATION OF HIGH-RISK CHILDREN FROM THE PERSPECTIVE OF PEDIATRIC SUBSPECIALISTS: A QUALITATIVE STUDY

Ayşe Esra Tapcı[1], İzzet Fidancı[2]

[1]Department of Pediatrics, University of Health Sciences Ankara Training and Education Hospital, Ankara, Türkiye

[2]Department of Family Medicine, Hacettepe University, Faculty of Medicine, Ankara, Türkiye

Correspondence:
İzzet Fidancı (e-mail: izzetfidanci@hacettepe.edu.tr)

Submitted: 29.03.2026// Accepted: 19.07.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: This study aimed to examine pediatric subspecialists' experiences, perceptions, and clinical decision-making processes regarding the vaccination of high-risk children using a qualitative approach, thereby identifying the individual and systemic factors influencing this process.

Materials and Methods: In this qualitative study, semi-structured in-depth interviews were conducted with physicians from different pediatric subspecialties working at a tertiary education and research hospital in Ankara, Türkiye. The interviews were audio-recorded, transcribed, and the data were analyzed using thematic analysis methods.

Results: Seventeen pediatric subspecialists were included in this study. The analysis identified seven main themes: (1) risk group vaccination perception and clinical framework, (2) vaccine prioritization and clinical priorities, (3) clinical decision-making process and uncertainties, (4) institutional structures and systemic functioning, (5) communication with families and vaccine hesitancy, (6) access, reimbursement, and logistical issues, and (7) training needs and development recommendations. Participants defined the risk group as a dynamic and specialty-specific concept and emphasized influenza, pneumococcal, and meningococcal vaccines as vital priorities. Differences between guidelines and patient-based clinical variability create significant uncertainties in decision-making.

Conclusion: The results reveal that the vaccination of at-risk children should not be left to individual physician initiative; rather, a structured healthcare system approach is necessary, coordinated by units specializing in preventive healthcare services, such as social pediatrics, and supported by standard algorithms, free access policies, and integrated electronic record systems.

Keywords: High-risk children, vaccination, pediatric subspecialists, clinical decision-making, health system barriers.

Introduction

Vaccination is one of the most effective and cost-effective interventions in preventing infectious diseases in children's health. However, children in risk groups are at a higher risk of vaccine-preventable diseases than the general pediatric population and require additional and specific vaccination strategies beyond standard vaccination schedules. Vaccination decisions, particularly for children with chronic diseases and those receiving immunosuppressive therapy, require a multidisciplinary and individualized approach owing to factors such as vaccine timing, efficacy, and safety.[1] Immunization programs promoting vaccination of these patient groups have been implemented worldwide because of the increased risk of morbidity and mortality associated with vaccine-preventable diseases in individuals with chronic conditions, such as cardiovascular, metabolic, and neoplastic diseases.[2-4] Immunization strategies, vaccination coverage, and the timely administration of vaccines for these individuals must be monitored.[5] Although high national vaccination coverage rates have been reported for healthy infants and children, definitive and up-to-date data on vaccination rates among infants and children with chronic diseases are limited, both in Turkey and worldwide.[6-7] However, vaccination rates among these children are lower than those in the general population.[6,8] This situation poses an additional risk to this vulnerable population, which is already more susceptible to infection. Therefore, it is recommended that children in special risk groups receive not only the vaccines recommended for healthy children, but also additional vaccines based on disease-specific indications.[7] In this context, physicians' clinical experience, risk perceptions, and opportunities offered by institutional structures play a decisive role in shaping vaccination decisions.[8] While the literature contains quantitative studies focusing on vaccination rates among children in risk groups, barriers to vaccine access, and parental attitudes[8-9], qualitative studies that examine in depth how pediatric subspecialists experience this process, the reasons behind their decisions, and the difficulties they encounter are limited. However, pediatric subspecialists hold a key position in the vaccination process, both clinically and ethically, as they are responsible for the long-term follow-up of this patient group. To date, most studies in this field have focused on vaccination coverage, parental

attitudes, vaccine hesitancy, or adherence to immunization recommendations. However, little is known about how pediatric subspecialists make vaccination decisions in everyday clinical practice, how they interpret uncertainty, and how institutional and health system factors influence these decisions. Exploring these perspectives is particularly important in the Turkish healthcare context, where the coordination of vaccination services for high-risk children involves multiple clinical disciplines and levels of care.

This study aimed to examine pediatric subspecialists' experiences, risk perceptions, and clinical decision-making processes regarding the vaccination of children in risk groups using a qualitative approach, thereby revealing the individual and systemic determinants that shape this process. This study aims to contribute to a better understanding of the clinical and structural difficulties encountered in vaccinating high-risk groups and the development of more effective, standardized, and accessible immunization policies.

Materials and Methods

This study was designed within an interpretive qualitative research paradigm to explore how pediatric subspecialists construct meaning around vaccination decision-making in children belonging to risk groups. A qualitative approach was chosen because the research question focuses on understanding complex clinical decision-making processes, which are shaped by contextual, experiential, and systemic factors that cannot be adequately captured through quantitative methods.

The interpretive approach enables an in-depth exploration of physicians' perceptions, reasoning, and experiences, acknowledging that clinical decisions are not solely based on objective guidelines but are influenced by individual judgment and contextual realities. Semi-structured in-depth interviews were preferred as the data collection method to allow flexibility in exploring participants' perspectives while ensuring consistency across key thematic domains.

This qualitative study was conducted to thoroughly examine the experiences, perceptions, and decision-making processes of pediatric subspecialists regarding the vaccination of children in risk groups. The research was designed based on an interpretive paradigm, and data were collected through semi-structured in-depth individual interviews. This study was reported based on the Consolidated Criteria for Reporting Qualitative Research (COREQ) and Standards for Reporting Qualitative Research (SRQR) guidelines.[10-11]

The study was conducted at a teaching and research hospital in Ankara, Türkiye, which provides tertiary care services. All physicians participating in the study were actively working in pediatric subspecialties and were directly involved in the follow-up and treatment processes of children in the risk group.

The study sample consisted of 17 physicians working in different pediatric subspecialties (pediatric neurology, hematology, cardiology, nephrology, gastroenterology, endocrinology, rheumatology, and allergy-immunology). Participants were selected using purposive sampling, and experience in monitoring children in the risk group was the basic inclusion criterion. Care was taken to ensure diversity in terms of clinical experience and academic title. Interviews were terminated when data saturation was achieved. Participants were assigned anonymous identifiers (P1–P17) to ensure confidentiality. These identifiers were used solely for anonymization purposes and did not reflect participants' professional rank or academic position. The sample size (n=17) was considered adequate for achieving depth and thematic richness, which are primary goals in qualitative research rather than statistical generalizability. Participants were selected using purposive sampling to ensure inclusion of information-rich cases with direct experience in managing children in risk groups. Maximum variation was sought across pediatric subspecialties and levels of professional experience to capture diverse perspectives. Participants were selected using purposive sampling to ensure inclusion of information-rich cases with direct experience in managing children in risk groups. Maximum variation sampling was employed across pediatric subspecialties and levels of professional experience to capture diverse perspectives on vaccination practices.

Data collection and analysis were conducted concurrently. After each interview, transcripts were reviewed, and preliminary coding was performed. Emerging codes and themes were continuously compared with data obtained from subsequent interviews. Data saturation was considered achieved when no new codes, themes, or conceptual insights emerged and when additional interviews yielded repetitive information without contributing further thematic development. During the final three interviews, no novel concepts were identified, and existing themes were consistently reinforced, indicating sufficient thematic depth and information power.

Although the sample consisted of 17 pediatric subspecialists, qualitative research prioritizes thematic richness and depth rather than statistical representativeness. Given the focused research question, the specialized expertise of participants, the diversity of subspecialties represented, and the achievement of data saturation, the sample size was considered adequate to provide a comprehensive understanding of the phenomenon under investigation.

All interviews were conducted face-to-face in a quiet and private room within the hospital setting to ensure confidentiality and minimize external interruptions. Each interview lasted approximately 25–40 minutes, depending on the depth of the discussion and participant engagement. The interviews were conducted by a researcher experienced in qualitative research methods and familiar with the clinical context of pediatric care. This facilitated effective communication with participants and enabled in-depth exploration of their perspectives. Field notes were taken during and immediately after the interviews to capture contextual details and non-verbal cues.

Data were collected using a semi-structured interview form developed by researchers. The interview questions covered topics such as risk group definition, vaccine prioritization, clinical decision-making process, guideline use, institutional practices, communication with families, access and reimbursement issues, and training needs. Interviews were conducted face-to-face, audio recordings were made with the informed consent of the participants, data were transcribed verbatim after the interviews were completed, and the audio recordings

were securely destroyed within 24 h. The interviews were conducted by a family physician and associate professor with previous experience in qualitative research and interview-based studies. The interviewer was familiar with the clinical context of pediatric care but had no supervisory, evaluative, or hierarchical relationship with the participants. No prior researcher-participant relationship existed that could influence participants' responses. To enhance reflexivity and reduce potential researcher bias, field notes were maintained throughout the study and analytical decisions were discussed regularly within the research team. Data coding and theme development were conducted independently by two researchers, and discrepancies were resolved through consensus. Investigator triangulation and the use of verbatim quotations were employed to ensure that the findings remained grounded in participants' accounts.

Data analysis was conducted in accordance with Braun and Clarke's thematic analysis approach.[12] Two researchers independently reviewed the transcripts and generated initial inductive codes from meaningful segments of text. A preliminary codebook was developed and refined throughout the analysis. Conceptually related codes were grouped into categories and subsequently organized into themes. Discrepancies in coding and theme allocation were resolved through discussion and consensus. Data were managed manually without the use of qualitative data analysis software because the dataset was of a manageable size.

Data analysis began with the researchers thoroughly reviewing the interview recordings to gain a deep understanding of the data collected. During this process, meaningful units of expression were identified, and preliminary codes were developed based on these units. The codes created were grouped according to their conceptual similarities to form thematic structures, and these themes were re-evaluated in terms of their compatibility with the dataset. In the final stage, the themes were defined in a clear and understandable manner, and the findings were presented systematically and coherently.

The coding process was conducted independently by two researchers, and differences between codes and themes were discussed until a consensus was reached. The final analysis

identified seven main themes reflecting the participants' narratives, along with major and minor themes related to them. To enhance reflexivity and minimize potential researcher bias, coding and theme development were performed independently by two researchers. Differences were discussed until consensus was achieved. Investigator triangulation, regular analytical discussions, and the use of verbatim quotations were employed to ensure that findings reflected participants' perspectives rather than researchers' assumptions.

The reliability and validity of the study were ensured in accordance with the strategies recommended in qualitative research. Data and researcher diversity, the use of direct participant quotations, inter-researcher comparison in the coding process, and continuous review of the consistency of themes with the data were applied. Participant characteristics were presented in detail to increase the findings' transferability.

Ethical Considerations

The study was approved by the Non-Interventional Clinical Research Ethics Committee of the relevant university (Ethics Committee Meeting Date: 01/2026, Decision No: E-25-732). Informed consent was obtained from all physicians participating in the study prior to their interviews. Participant confidentiality was maintained, interviews were coded to ensure anonymity, and the data were used solely for research purposes. The study was conducted in accordance with the principles outlined in the ethics committee application and approval documents.

Reflexivity was considered throughout the research process. The interviewer was a family physician and academic researcher with experience in qualitative health research and an interest in vaccination practices. Recognizing that professional experience could influence data interpretation, reflexive field notes, independent coding by two researchers, investigator triangulation, and regular analytical discussions were used to minimize potential bias and ensure that themes remained grounded in participants' perspectives.

Results

This study examined pediatric subspecialists' experiences, decision-making approaches, and challenges related to vaccinating children in high-risk groups multidimensionally. The participants' narratives revealed that vaccinating high-risk groups is not only a practice based on clinical knowledge; it is also a complex process shaped by uncertainty, institutional structure, shared responsibility, interaction with families, and limitations related to the healthcare system. Seventeen pediatric subspecialists from eight pediatric subspecialties participated in the study. Participants represented a range of professional roles and levels of clinical experience, including clinical fellows, specialists, and academic faculty members. This heterogeneity enabled the inclusion of diverse perspectives regarding the vaccination of high-risk children and strengthened the breadth of the findings. The demographic and professional characteristics of the pediatric subspecialists participating in the study and their general behavioral profiles regarding vaccination decisions in risky situations are shown in Table 1.

Table 1. Demographic and professional characteristics of the participants and their vaccination decision-making approaches in risky situations.

Participant No. (Position)	Years	Speciality*	Primary Vaccination Approach
P-1 (Associate Professor)	13	Neurology	Referral to social pediatrics
P-2 (Specialist Physician)	6	Neurology	Cautious, patient-based
P-3 (Specialist Physician)	5	Nephrology	Delegated responsibility
P-4 (Professor)	24	Hematology	Guideline-driven
P-5 (Specialist Physician)	6	Cardiology	Indication-focused
P-6 (Specialist Physician)	6	Nephrology	Responsibility excluded
P-7 (Specialist Physician)	4	Rheumatology	Safety-oriented
P-8 (Associate Professor)	13	Cardiology	Routine vaccination advocate
P-9 (Associate Professor)	6	Rheumatology	Treatment-prioritized
P-10 (Specialist Physician)	8	Gastroenterology	Shared decision-making
P-11 (Specialist Physician)	7	Endocrinology	Consultation-oriented
P-12 (Specialist Physician)	6	Neurology	Accessibility-sensitive
P-13 (Specialist Physician)	7	Gastroenterology	Ethical counselling
P-14 (Associate Professor)	6	Allergy and Immunology	Risk–benefit assessment
P-15 (Associate Professor)	10	Cardiology	Team-based decision-making
P-16 (Specialist Physician)	4	Endocrinology	Context-sensitive routine practice
P-17 (Specialist Physician)	4	Allergy and Immunology	Vaccine advocacy

P: Participant. *All participants were employed at the same type of institution, a tertiary care training and research hospital, and were working in Ankara, Türkiye.

Participants with varying clinical experiences and working in different departments showed diversity in their approaches to the vaccination of high-risk children. The thematic analysis of the interview data identified seven main themes explaining these individual approaches. These themes were grouped under the following headings: perception of risk group vaccination and clinical framework, vaccine prioritization and clinical priorities, clinical decision-making process and uncertainties, institutional structures and systemic functioning, communication with families and vaccine hesitancy, access, reimbursement and logistical issues, and training needs and development recommendations (Table 2).

Table 2. Theme Structure Related to Pediatric Subspecialty Physicians' Risk Group Vaccination Practices

Main Theme	Major Theme	Minor Themes
Risk Group Vaccination Perception and Clinical Framework	Definition of the risk group	Chronic diseases; Immunodeficiencies; Use of immunosuppressive/biological agents; Organ transplantation/asplenia; Prematurity
	Specialty-specific risk perception	Neurological; Cardiac; Hematological-oncological; Renal; Metabolic/endocrine diseases
Vaccine Prioritization and Clinical Priorities	Vaccines universally considered critical	Influenza; Pneumococcus; Meningococcus; Hepatitis A/B
	Priority vaccines specific to the branch	Palivizumab; Meningococcal (eculizumab/complement deficiency); Completion of live vaccines prior to transplantation; HPV
Clinical Decision-Making Process and Uncertainties	Use of guidelines	National guidelines; International guidelines (WHO/CDC/ACIP); Lack of guidelines; Difficulty accessing current information
	Guideline-based decision-making	Patient-based decision; Approach based on clinical experience; Preference for the safest option
	Dealing with conflicting recommendations	Postponing vaccination; Prioritizing treatment; Consultation within the specialty.
Institutional structures and systemic functioning	Protocols and flow charts	Written protocol exists; No written protocol; Application based on association guidelines
	Area of responsibility	Referral to social pediatrics; Healthy child clinic; Responsibility perceived as outside the specialty
	Recording and monitoring systems	No follow-up in the hospital information system; Follow-up via Personal Health Record System; Manual registration
Communication with Families and Vaccine Hesitancy	General attitude of families	High acceptance in chronic diseases; Rarely encounter vaccine refusal
	Communication strategies	Emphasis on risk; Negative examples; "If it were my child" approach; Leaving the decision to the family
	Written information and consent	No written material; Pre-medication information forms; No standard vaccine consent form
Access, reimbursement, and logistical issues	Vaccine supply	Lack of influenza vaccine; Seasonal supply issues
	Reimbursement issues	Non-reimbursed vaccines despite being in a risk group; Financial constraints; Perception of unfairness
Training Requirements and Development Recommendations	Requirements for physicians	Clear algorithms; Up-to-date guideline summaries; Online training
	System-level recommendations	Free vaccines for high-risk groups; National standard algorithm; Electronic record integration
	Public awareness	Media campaigns; Ministry of Health announcements; Accurate information on social media

1. Risk Group Vaccination Perception and Clinical Framework

Participants defined the concept of risk group as a dynamic framework that includes all clinical and treatment-related factors that increase a child's vulnerability to infections, rather than a narrow classification based on diagnosis. This approach reflects the understanding that the risk group is not a fixed category; it can vary depending on the stage of the disease, medications used, and accompanying conditions. Immunosuppressive treatments, a history of organ transplantation, and prematurity were central to the participants' risk definitions.

"When I hear 'risk group,' immunosuppressed patients immediately come to mind. They cannot be assessed using a standard schedule." (P-4)

“Whether they are premature, have a neurological condition, or are on long-term medication, they are all more susceptible to infection.” (P-1)

“In children with chronic kidney disease, infections tend to be much more severe.” (P-3)

The fact that risk perception is shaped by specialty indicates that there is no common definition of risk. Physicians assess risk based on the complications they frequently encounter in their clinical practice, leading to heterogeneity across specialties in risk group vaccination rates.

2. Vaccine Prioritization and Clinical Priorities

Participants' narratives revealed that not all vaccines are considered equally important for children in the risk group, with some vaccines being positioned as vital and indispensable. The influenza vaccine was the most emphasized vaccine because of its prevalence and potential to cause severe infections. Pneumococcal and meningococcal vaccines are generally associated with the risk of fatal complications.

"We have seen very severe cases when the flu vaccine was not administered." (P-12)

"When pneumococcal and meningococcal vaccines are not administered, irreversible outcomes occur." (P-15)

"For high-risk groups, influenza vaccination is truly lifesaving." (P-8)

However, the importance of some vaccines varies depending on the specialty. In cardiology, palivizumab administration stands out; in immunology and rheumatology, meningococcal vaccines; and in patients undergoing transplantation, the timing of live vaccines is important. This situation shows that vaccine prioritization is determined by clinical context rather than universal principles.

3. Clinical Decision-Making Process and Uncertainties

While acknowledging the existence of national and international guidelines, participants stated that these guidelines cannot always be directly applied to daily clinical practice settings. Differences between guidelines, frequent updates, and patient-specific clinical conditions create uncertainty in the decision-making process. This uncertainty leads physicians to approach the guidelines critically and selectively.

"Guidelines exist, but there is no clear algorithm that fits every patient." (P-7)
"The national guideline says one thing, the CDC says another." (P-10)
"I often have to interpret the guidelines." (P-11)

"When the patient's condition becomes more complicated, the guidelines are no longer sufficient on their own. We often have to make decisions based on clinical judgment." (P-14)

In uncertain situations, physicians often make patient- and experience-based decisions. The "safest course" approach can lead to postponing vaccination, especially during periods of intensive treatment. This situation shows that physicians constantly reevaluate the risk-benefit balance.

4. Institutional Structures and Systemic Functioning

Most participants stated that there was no standard protocol or flowchart for risk group vaccination at their institutions. It was observed that practices were often based on individual experiences and clinical habits. This makes it difficult to ensure consistency in practice within the institution.

"We don't have a written vaccination protocol at our clinic." (P-6)

“Everyone proceeds based on their own experience.” (P-3)

“This task relies somewhat on personal effort.” (P-15)

“I usually refer these patients to social pediatrics because there is no clearly defined responsibility within our clinic.” (P-2)

The uncertainty surrounding vaccination responsibilities often leads subspecialists to perceive themselves as non-decision-makers. The vaccination process is frequently referred to as social pediatrics or well-child units. However, a small number of participants (n=3) positioned decisions regarding the vaccination of at-risk groups outside their own clinical responsibilities and consciously referred this process to other units.

5. Communication with Families and Vaccine Hesitancy

Participants emphasized that vaccine acceptance was generally high among families of children with chronic illnesses. The severity of illness and past negative experiences positively influenced families' attitudes toward vaccination. While vaccine refusal is rare, physicians typically use persuasive and experience-based communication strategies.

“These families are already open to vaccination because they have been through so much.” (P-13)
“Vaccine refusal is very rare in our group.” (P-11)

"Fear of illness convinces families." (P-8)

"Families are often reassured when we explain what we would recommend for our own patients or children." (P-6)

Healthcare providers sharing their personal opinions and experiences has emerged as a key factor in strengthening trust.

6. Access, Reimbursement, and Logistical Issues

Participants identified vaccine supply and reimbursement issues as significant barriers to vaccinating at-risk groups. Difficulties in obtaining influenza vaccines and the fact that some vaccines are not covered by reimbursement directly affect clinical decision-making.

"Asking for the flu vaccine when it's not available is just theoretical." (P-12)
"Some vaccines are not covered by reimbursement despite being part of the risk group." (P-17)
"Even if the family wants it, they cannot afford it financially." (P-9)

"It is frustrating to recommend a vaccine that we know is necessary but that families cannot access because of cost." (P-7)

This situation creates a clear sense of injustice among physicians and highlights the need for systemic reforms.

7. Education Requirements and Development Recommendations

Participants emphasized that systemic and structural reforms, rather than individual knowledge, are needed to increase the effectiveness of group at-risk vaccination. Creating clear, simple, and up-to-date algorithms, making risk group vaccines available free of charge, and strengthening electronic record systems were among the most frequently mentioned recommendations.

"If there were a clear national algorithm, everyone would feel at ease." (P-4)
"We need short and up-to-date guideline summaries." (P-11)
"Risk group vaccines should be free." (P-15)

Discussion

This qualitative study revealed that pediatric subspecialists' decision-making processes regarding the vaccination of children in risk groups are based not only on medical knowledge but also on multidimensional factors such as uncertainty, institutional structure, perception of responsibility, interaction with families, and health system constraints. The findings indicate that risk group vaccination is not a mechanical process carried out according to standard protocols but rather a context-sensitive clinical practice largely based on physician initiative. By focusing on pediatric subspecialists working within a tertiary care setting in Türkiye, the study provides context-specific insights into the clinical and organizational challenges that may not be captured through quantitative surveys or vaccination coverage data alone.

Unlike previous studies that primarily focused on vaccination coverage rates, parental attitudes, or guideline adherence, this study provides an in-depth exploration of how pediatric subspecialists experience vaccination decision-making in high-risk children. The findings demonstrate that vaccination practices are shaped not only by clinical evidence but also by uncertainty management, institutional responsibility, reimbursement policies, and interdisciplinary communication. This broader perspective contributes to understanding why gaps between recommendations and clinical practice may persist despite the availability of guidelines.

Theme 1: Perception of Risk Group Vaccination and Clinical Framework

This study found that physicians do not define the risk group as a fixed and homogeneous category but as a dynamic framework that constantly changes based on the child's clinical condition, treatment process, and functional status of the immune system. In the literature,

children in the risk group are defined as individuals who show increased susceptibility to vaccine-preventable diseases owing to underlying chronic diseases, immunodeficiencies, or immunosuppressive treatments.[1,13] Premature infants, organ transplant recipients, patients with asplenia, those using biological agents, and children with chronic cardiac, pulmonary, renal, neurological, hematological, and immune-compromising diseases are particularly evaluated as "special risk groups" in many studies.[5,13] However, current reviews and field studies emphasize that there is no single, universal definition of a risk group; the level of risk can change over time depending on factors such as disease severity, treatments used, and immune response.[14] This approach demonstrates that the risk group is not a fixed clinical category but a context-sensitive concept that requires regular reassessment, consistent with our finding that participants' risk perception was shaped in a discipline-specific and variable manner. The variation in risk perception among healthcare workers highlights the absence of a common and standardized definition of risk, creating a significant area of heterogeneity that could lead to different vaccination approaches for children with similar clinical conditions.

Theme 2. Vaccine Prioritization and Clinical Priorities

The literature reports that not all vaccines have equal clinical importance in children at risk, with influenza, pneumococcal, and meningococcal vaccines being considered priority and critical vaccines for these groups. The CDC/ACIP guidelines define the influenza vaccine as a basic protective intervention for children with chronic diseases or immunodeficiency. Pneumococcal and meningococcal vaccines are classified as high-priority vaccines because of the risk of invasive and life-threatening infections[15]. Indeed, influenza infection has been shown to significantly increase the need for hospitalization and intensive care in at-risk children, while pneumococcal and meningococcal infections are associated with higher mortality and risk of permanent sequelae in this population[16-17]. However, vaccine prioritization is not independent of the specific clinical context of the specialty; for example, in cardiology practice, palivizumab stands out; in immunology and rheumatology,

meningococcal vaccines are prominent; and in the transplantation process, the timing of live vaccines is emphasized.

Theme 3: Clinical Decision-Making Process and Uncertainties

In the literature, the clinical decision-making process is defined as a dynamic process that inevitably involves uncertainty, particularly in situations involving complex and multiple risk factors. Clinical uncertainty is considered a fundamental factor that makes it difficult for physicians to make decisions based on a single, definitive truth owing to insufficient evidence, inconsistencies between guidelines, and variable clinical conditions on a patient-by-patient basis.[18-19] Indeed, the fact that vaccine guidelines are presented with different content and priorities by different institutions (e.g., national authorities, WHO, and CDC/ACIP) makes it difficult to develop clear algorithms that can be directly translated into clinical practice, especially for children in risk groups. In this context, physicians adopting the "safest course" approach by interpreting guidelines based on their own clinical experience and the patient's specific risk profile, rather than applying them verbatim, is consistent with experience-based and risk-averse clinical decision models defined in the literature.[20] This situation provides a theoretical framework for explaining participants' tendencies to delay vaccination, prioritize treatment, or make individualized decisions in our study, and shows that decision-making processes in risk group vaccination depend not only on guidelines but also on professional judgment based on uncertainty management.

Theme 4: Institutional Structures and Systemic Functioning

The literature indicates that the effective delivery of preventive health services depends not only on clinical knowledge but also on institutional structures with clearly defined areas of responsibility. The World Health Organization defines the coordination of immunization services, carried out in a comprehensive manner through monitoring and registration processes and organized by centralized and specialized units, as a fundamental component of health system effectiveness[21]. In this context, it is recommended in many countries that the vaccination of children in special risk groups be coordinated by units focused on the

preventive aspect of child health, with clinical subspecialties contributing to this process in an advisory and guiding role.[22-23] In our study, the fact that most participants directed vaccination responsibility to social pediatrics or well-child units indicates that central planning by a specialized unit, such as social pediatrics, may be a more functional and sustainable model than conducting risk group vaccination under the individual responsibility of subspecialty clinics. This approach could reduce the clinical burden and strengthen consistency in practice by ensuring that the immunization processes of children in risk groups are carried out in a standardized, traceable, and coordinated manner. An important finding of this study is that responsibility for vaccination of high-risk children is often fragmented across subspecialty services. This highlights the potential coordinating role of primary care and family medicine, which are uniquely positioned to provide continuity of care, monitor vaccination status, and facilitate communication between tertiary care services and families. Strengthening collaboration between subspecialists and primary care providers may help reduce missed vaccination opportunities and improve the consistency of vaccination practices across healthcare settings.

Theme 5: Communication with Families and Vaccine Hesitancy

The literature indicates that attitudes toward vaccination among parents of children with chronic diseases are heterogeneous, with some studies reporting that vaccine hesitancy may be higher in this group than in the general population. In a study conducted by Napolitano et al. with parents of children with chronic diseases in Italy, hesitancy toward vaccines was found in a significant proportion of parents, and concerns about the possible side effects of vaccines were shown to be a determining factor.[24] Similarly, a recent study conducted in Turkey reported that vaccine hesitancy rates among parents of children with chronic diseases were higher than those among parents of healthy children; this trend was more pronounced in certain disease groups, such as type 1 diabetes and autism spectrum disorder.[25] These findings suggest that chronic illness alone is not a factor that increases vaccine acceptance; on the contrary, uncertainty and risk perception in some clinical groups may reinforce vaccine hesitancy. In our study, the fact that participants rarely encountered

vaccine refusal can be explained by the physicians interviewed working primarily with patient groups at high risk for congenital, immunological, and infectious diseases and families having long-term contact with the healthcare system. This situation reveals that vaccine hesitancy is shaped not so much by the presence of chronic disease as by the type of disease, families' sources of information, and the trust established with physicians.

Theme 6. Access, Reimbursement, and Logistical Issues

In the literature, access to vaccines and financial barriers are among the most important systemic factors limiting immunization rates in at-risk populations. The World Health Organization emphasizes that the availability and free provision of vaccines, especially for vulnerable groups, are critical determinants of the effectiveness of immunization programs, noting that disruptions in vaccine access deepen health inequities[2]. Similarly, the non-reimbursement of vaccines or the existence of additional costs that families must cover is identified as a significant barrier that directly affects vaccination decisions for children in risk groups.[26] Indeed, multicenter studies have shown that supply problems and cost barriers in the provision of seasonal vaccines, such as influenza vaccines, render physicians' clinical recommendations impractical and significantly reduce vaccination rates.[27] In this context, the fact that physicians in our study identified vaccine shortages and reimbursement issues as key barriers highlights that vaccinating at-risk groups is not only a clinical issue but also a structural problem directly related to the health system's dimensions of equity, financing, and resource management.

Theme 7. Training Needs and Development Recommendations

Literature emphasizes that structured and sustainable mechanisms at the system level, rather than individual physician knowledge, are decisive in increasing the effectiveness of vaccination programs for high-risk children. The World Health Organization defines clear clinical algorithms, standard guidelines, and the integration of electronic record systems as fundamental health system components for improving the quality of immunization services[21]. Similarly, clinical decision support systems and up-to-date guideline summaries

have been shown to reduce physicians' decision burden and strengthen the consistency of practice, particularly in complex patient groups.[28] The integration of electronic health records with immunization systems increases continuity in patient follow-up and contributes to reducing missed vaccination opportunities.[29] In this context, the recommendations of the participants in our study regarding national standard algorithms, free vaccinations for risk groups, and strengthening electronic record systems reveal that the vaccination of risk groups should be transformed from an individual effort-based clinical practice into a structured and traceable healthcare service model.

One of the most important strengths of this study is that it examines the experiences of physicians working in different pediatric subspecialties in depth, revealing the decision-making processes regarding the vaccination of children in risk groups from a qualitative perspective. The participants' diverse clinical backgrounds ensured that the findings were multidimensional and rich in content.

This study has several limitations that should be considered when interpreting the findings. First, the study was conducted in a single tertiary care center, which may limit the transferability of the findings to other healthcare settings with different institutional structures and patient populations. Second, although the sample size (n=17) is appropriate for qualitative research aiming at depth and thematic richness, it may still be considered limited in terms of capturing the full spectrum of perspectives across all pediatric subspecialties. Third, the study focused exclusively on pediatric subspecialists and did not include other key stakeholders involved in the vaccination process, such as families, primary care physicians, or public health professionals. The absence of these perspectives may limit a comprehensive understanding of the multi-level dynamics influencing vaccination practices in risk groups. However, despite these limitations, the study provides in-depth insights into the clinical decision-making processes of pediatric subspecialists, a group that plays a central and often underexplored role in the management of high-risk pediatric populations. Qualitative design enabled a detailed exploration of contextual and experiential factors that are not easily captured through quantitative approaches.

The findings of this study have direct clinical implications. Variability in decision-making across subspecialties indicates the need for clear and standardized vaccination algorithms for high-risk children. Implementing institutional protocols and integrating vaccination pathways into hospital systems may improve coordination and consistency. Strengthening clinical decision support and interdisciplinary consultation can reduce uncertainty and prevent unnecessary delays in vaccination. In addition, ensuring vaccine availability and reimbursement for high-risk groups is essential to translate recommendations into practice. Targeted training and concise guideline tools may further support physicians in managing complex vaccination decisions. From a Family Medicine perspective, the findings highlight the central role of primary care in coordinating vaccination processes for children in risk groups. Family physicians are uniquely positioned to ensure continuity of care, monitor vaccination status, and bridge communication between subspecialists and families. Strengthening collaboration between tertiary care and primary care, along with integrating vaccination follow-up into primary care systems, may improve coverage, timeliness, and consistency of vaccination practices in high-risk pediatric populations.

Conclusion

This qualitative study suggests that pediatric subspecialists' decision-making regarding the vaccination of high-risk children is influenced not only by clinical considerations but also by institutional structures, guideline-related uncertainties, and health system factors. The findings highlight potential areas for improvement in the organization and coordination of vaccination services for high-risk children, including clearer clinical guidance, enhanced interdisciplinary collaboration, and strengthened support systems. Given the single-center qualitative design, further studies involving different healthcare settings and stakeholder groups are needed to explore the transferability of these findings.

Ethical Considerations: The study was approved by the Non-Interventional Clinical Research Ethics Committee of the Ankara Training and Research Hospital (Ethics Committee Meeting Date: 01/2026, Decision No: E-25-732).

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

References

1. Crawford NW, Bines JE, Royle J, Buttery JP. Optimizing immunization in pediatric special risk groups. *Expert Rev Vaccines.* 2011;10(2):175–86.
2. World Health Organization. Global Vaccine Action Plan 2011–2020. Geneva: World Health Organization; 2013. Available from: https://www.who.int/publications/i/item/global-vaccine-action-plan-2011-2020 (accessed on 24 Jan 2026).
3. Centers for Disease Control and Prevention. Vaccines & Immunizations. Special Situations. Available online: https://www.cdc.gov/vaccines/hcp/imz-best-practices/special-situations.html (accessed on 24 Jan 2026).
4. European Centre for Disease Prevention and Control. Guidance on priority risk groups for influenza vaccination. Stockholm: ECDC; 2008. Available from: https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/0808_GUI_Priority_Risk_Groups_for_Influenza_Vaccination.pdf. (accessed on 24 Jan 2026)
5. Caselli D, Aricò M, Fiasca F, Tafuri S. Policy of vaccination of "fragile children": Results of a survey of 14 Italian children's hospitals. *Hum Vaccin Immunother.* 2023;19(2):2245701.
6. Diallo D, Santal C, Lagrée M, Martinot A, Dubos F. Vaccination coverage of children with chronic diseases is inadequate, especially for specifically recommended vaccines. *Acta Paediatr.* 2020;109(12):2677–84.
7. Sheikh S, Biundo E, Courcier S, Damm O, Launay O, Maes E, et al. A report of the status of vaccination in Europe. *Vaccine.* 2018;36(33):4979–92.
8. Pandolfi E, Graziani MC, Ieraci R, Cavagni G, Tozzi AE. A comparison of populations vaccinated in a public service and in a private hospital setting in the same area. *BMC Public Health.* 2008;8:278.
9. Napolitano F, Della Polla G, Capano MS, Augimeri M, Angelillo IF. Vaccinations and chronic diseases: Knowledge, attitudes, and self-reported adherence among patients in Italy. *Vaccines (Basel).* 2020;8(4):560.

10. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *Int J Qual Health Care.* 2007;19(6):349–357. doi:10.1093/intqhc/mzm042.
11. EQUATOR Network. Enhancing the QUAlity and Transparency Of Health Research. Oxford: EQUATOR Network; 2025. Available from: https://www.equator-network.org/ (accessed on 24 Jan 2026).
12. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol.* 2006;3(2):77–101.
13. Doherty M, Schmidt-Ott R, Santos JI, Stanberry LR, Hofstetter AM, Rosenthal SL, Cunningham AL. Vaccination of special populations: Protecting the vulnerable. *Vaccine.* 2016;34(48 Suppl 5):S106–13.
14. Wang C, Zhang L, Xu B, Fu X, Fu L, Li B, et al. Vaccination recommendation patterns and associated factors among children with special health care needs. *Vaccines.* 2025;13(11):1145.
15. Kroger A, Bahta L, Hunter P. General best practice guidelines for immunization: Best practices guidance of the Advisory Committee on Immunization Practices (ACIP). *MMWR Recomm Rep.* 2022;71(4):1–95.
16. World Health Organization. Influenza vaccines: WHO position paper – March 2022. *Weekly Epidemiological Record*. 2022;97(19):185–208.
17. Esposito S, Principi N. Vaccination of children with chronic diseases. *Hum Vaccin Immunother.* 2015;11(10):2566–73.
18. Han PKJ, Klein WMP, Arora NK. Varieties of uncertainty in health care: A conceptual taxonomy. *Medical Decision Making.* 2011;31(6):828–38.
19. Greenhalgh T, Howick J, Maskrey N; Evidence Based Medicine Renaissance Group. Evidence based medicine: a movement in crisis? BMJ. 2014 Jun 13;348:g3725.
20. Djulbegovic B, Elqayam S, Dale W. Rational decision making in medicine: Implications for overuse and underuse. J Eval Clin Pract. 2018;24(3):655-65.
21. World Health Organization. Everybody's business: Strengthening health systems to improve health outcomes. Geneva: World Health Organization; 2007.

22. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. Milbank Q. 2005;83(3):457-502.
23. Kringos DS, Boerma WGW, Hutchinson A, Saltman RB, editors. Building primary care in a changing Europe [Internet]. Copenhagen (Denmark): European Observatory on Health Systems and Policies; 2015.
24. Napolitano F, Miraglia del Giudice G, Angelillo S, Fattore I, Licata F, Pelullo CP, Di Giuseppe G. Hesitancy towards childhood vaccinations among parents of children with underlying chronic medical conditions in Italy. *Vaccines (Basel).* 2022;10(8):1254.
25. Kömürlüoğlu A, Çelik N, Çiçek AU, Yalçın SS. Vaccine hesitancy among parents of children with chronic diseases of different pathophysiology: a cross-sectional study in Sivas, Türkiye. BMC Public Health. 2025;25(1):1683.
26. Ozawa S, Clark S, Portnoy A, Grewal S, Brenzel L, Walker DG. Return On Investment From Childhood Immunization In Low- And Middle-Income Countries, 2011-20. Health Aff (Millwood). 2016;35(2):199-207.
27. Luyten J, Beutels P. The Social Value Of Vaccination Programs: Beyond Cost-Effectiveness. Health Aff (Millwood). 2016;35(2):212-8.
28. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ. 2005;330(7494):765.
29. Groom H, Hopkins DP, Pabst LJ, Murphy Morgan J, Patel M, Calonge N, et al.; Community Preventive Services Task Force. Immunization information systems to increase vaccination rates: a community guide systematic review. J Public Health Manag Pract. 2015;21(3):227-48.

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):381-405 // 10.5505/amj.2026.28913

ASSOCIATIONS OF URINARY AND BLOOD TRACE ELEMENTS WITH DIABETES MELLITUS IN US ADULTS: A CROSS-SECTIONAL NHANES 2013–2014 ANALYSIS

Melih Gaffar Gözükara[1], Servet Birgin İritaş[2]

[1] Department of Public Health, Faculty of Medicine, Ankara Yıldırım Beyazıt University, Ankara, Türkiye

[2] Department of Forensic Medicine, Faculty of Medicine, Ankara Yıldırım Beyazıt University, Ankara, Türkiye

Correspondence:
Melih Gaffar Gözükara (e-mail: mggozukara@aybu.edu.tr)

Submitted: 17.01.2026 // Accepted: 10.08.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: Environmental trace element exposure has been proposed as a modifiable risk factor for type 2 diabetes mellitus (DM), but findings are inconsistent. We examined associations between urinary and blood trace elements and DM, and their sensitivity to model specification.

Materials and Methods: We analysed 1,486 adults aged 20 years or older from the 2013–2014 NHANES cycle. DM was defined by American Diabetes Association criteria. Seventeen urinary and five blood elements were measured by mass spectrometry. Groups were compared using Mann–Whitney U and chi-square tests with Benjamini–Hochberg correction. Log-transformed concentrations were entered into logistic regression models adjusted for age, sex, body mass index, race/ethnicity, socioeconomic position, smoking, hypertension, and kidney disease, fitted with and without complex survey weights. Odds ratios (OR) are reported per interquartile-range increase.

Results: DM prevalence was 13.7%. In adjusted models (n=1,388), urinary arsenous acid was inversely associated with DM (OR 0.55, 95% CI 0.36–0.86, p=0.008); this persisted under the survey design (OR 0.47, 0.30–0.72, p<0.001) and after creatinine adjustment. Urinary cadmium, strongly associated when age was omitted (OR 1.56, 1.17–2.07, p=0.002), showed no association after adjustment for age (OR 0.89, 0.65–1.23, p=0.481). Blood selenium was associated in the unweighted analysis (OR 1.26, p=0.020) but not in the weighted analysis; urinary tin and molybdenum were only associated under the survey design.

Conclusion: The cadmium–DM association was fully attributable to age confounding, and arsenous acid showed an inverse rather than positive association. Such associations are highly sensitive to model specification; adequate age adjustment and use of the complex survey design are essential in NHANES-based analyses.

Keywords: Trace elements, environmental exposure, diabetes mellitus, nutrition surveys, confounding factors, epidemiologic.

Introduction

Environmental exposure to heavy metals and trace elements has become a prominent concern in the study of metabolic disease.[1] Type 2 diabetes mellitus (DM) continues to increase in prevalence despite advances in clinical management, prompting interest in contributors to metabolic risk that are modifiable at the population level through environmental regulation, occupational protection and tobacco control. Unlike individual behavioural risk factors, exposure to toxic metals is distributed unevenly across populations and is amenable to structural rather than purely clinical intervention.

Trace elements occupy a dual position in human physiology. Selenium (Se) is required for antioxidant defence through its incorporation into selenoproteins, yet its therapeutic window is narrow.[2]

Evidence from the Selenium and Vitamin E Cancer Prevention Trial and from subsequent population studies has suggested a U-shaped or positive association between supranutritional Se status and DM risk, commonly described as the "selenium paradox".[3-5] The strength and direction of this relationship nevertheless vary across populations and exposure ranges, and estimates are sensitive to the extent of confounder adjustment.

Cadmium (Cd), by contrast, is a non-essential toxic metal present in cigarette smoke, industrial emissions and contaminated food crops, with a biological half-life of 10–30 years.[6] Experimental work has implicated Cd in pancreatic beta-cell dysfunction and insulin resistance, and toxic metals have been shown to accumulate in metabolically active tissues following long-term exposure.[7,8] Epidemiological findings are markedly inconsistent, however: systematic reviews report associations ranging from strongly positive to null, and NHANES analyses have described positive, null and even inverse associations depending on the biomarker matrix and covariate set applied.[9-11] This heterogeneity has a structural explanation. Urinary Cd reflects cumulative lifetime burden and rises steeply with age, while age is simultaneously the strongest determinant of DM prevalence; where age adjustment is

incomplete, this configuration can generate associations that do not survive fuller specification.

Similar considerations apply to arsenic. Arsenous acid, the most reactive trivalent inorganic species, is a recognised endocrine disruptor that impairs insulin-stimulated glucose uptake and interferes with GLUT4 translocation in experimental systems, and high chronic exposure through drinking water is associated with elevated diabetes prevalence.[12-14] At the low exposure levels typical of the general US population, the evidence is considerably weaker, and its interpretation is complicated by two features: a substantial proportion of measurements falls below the analytical limit of detection, and concentrations in spot urine depend heavily on urinary dilution, which differs between individuals with and without diabetes because of osmotic diuresis. The metabolic significance of other elements remains less well characterised; reports on tin conflict, and molybdenum has shown inconsistent relationships with glucose homeostasis despite insulin-mimetic effects in experimental models.[15-20]

The National Health and Nutrition Examination Survey (NHANES) permits simultaneous measurement of a broad element panel alongside detailed demographic and clinical data in a nationally representative sample, but realising that advantage requires attention to the complex sampling design and to the confounding structure described above.[21,22] This study had two aims: to describe associations between urinary and blood trace elements and DM among US adults, applying correction for multiple comparisons across the full panel rather than reporting selected elements in isolation; and to evaluate how robust these associations are to model specification, specifically to adjustment for age and socioeconomic position and to incorporation of the survey design. Given the cross-sectional design, all findings are presented as associations and no causal inference is intended.

Materials and Methods

Study design and population

This was a cross-sectional analysis of the 2013–2014 cycle of the National Health and Nutrition Examination Survey (NHANES), a continuous programme conducted by the National Center for Health Statistics using a stratified, multistage probability sampling design to characterise the civilian, non-institutionalised US population. This cycle was selected because it is the most recent in which the full panel of urinary trace elements examined here, including speciated urinary arsenic, was measured concurrently with the blood panel in the same subsample, permitting mutual adjustment within a single model.

Of the 10,175 participants, 5,769 were aged 20 years or older and 5,588 completed the Mobile Examination Centre (MEC) examination. Urinary trace elements were measured in a randomly selected one-third subsample; 1,858 examined adults were selected, of whom 1,811 had a complete urinary panel. After exclusion of participants with missing HbA1c, blood trace element, or covariate data, the descriptive sample comprised 1,486 adults. Multivariable models were restricted to participants with complete data on the family income-to-poverty ratio and urinary creatinine, yielding 1,388 participants with 192 diabetes cases. No imputation was performed. Participant flow is shown in Figure 1.

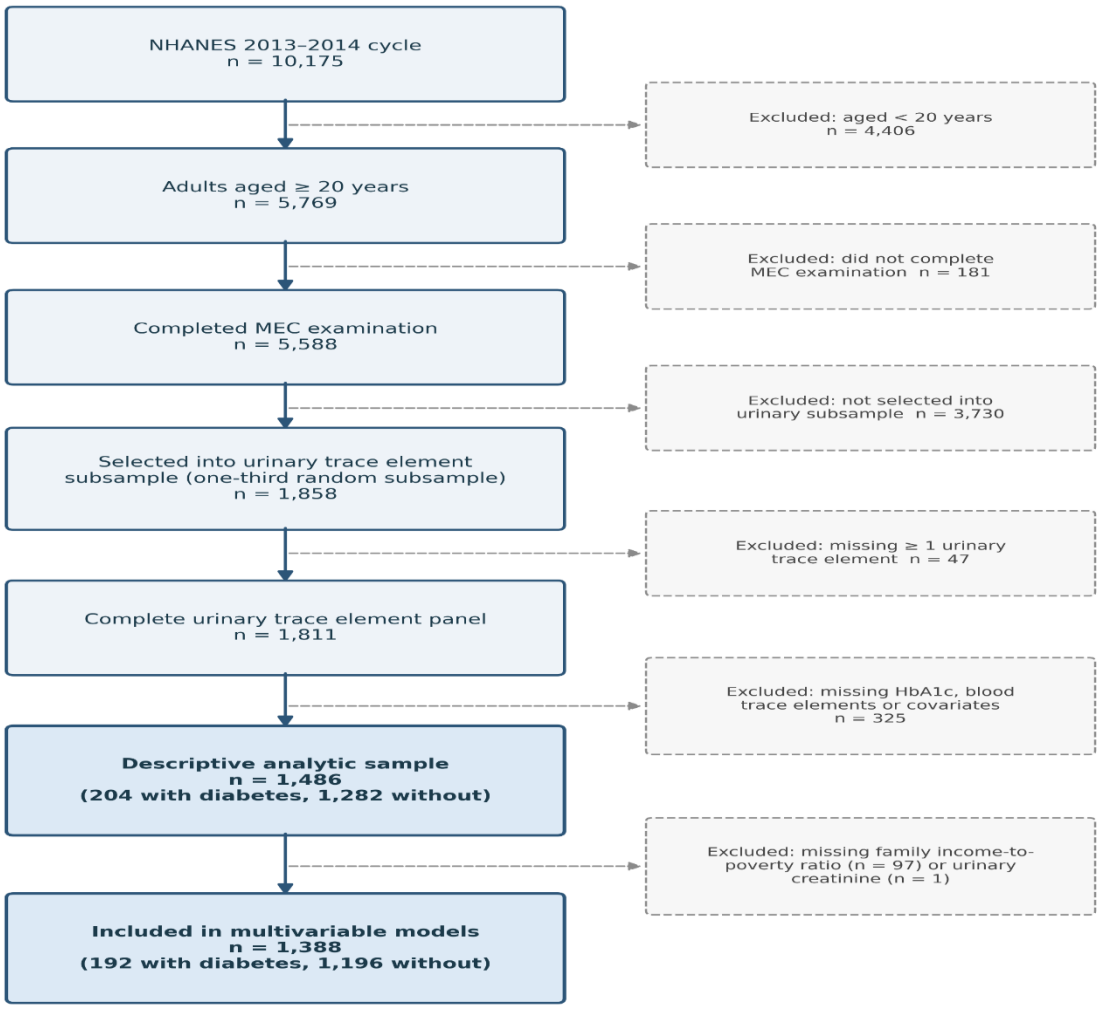


Figure 1. Participant flow diagram

Ascertainment of diabetes

Diabetes was defined according to American Diabetes Association criteria as HbA1c ≥6.5%, a physician diagnosis reported by the participant, or current use of insulin or oral glucose-lowering agents. Definition by HbA1c alone would misclassify pharmacologically treated participants with controlled glycaemia as non-diabetic, and the composite definition was therefore applied throughout.

Laboratory measurements

HbA1c was measured by high-performance liquid chromatography, and urinary creatinine enzymatically. Trace elements in whole blood and urine were quantified by inductively coupled plasma dynamic reaction cell mass spectrometry, with urinary arsenic species separated chromatographically before detection. Concentrations below the limit of detection (LOD) were retained at the LOD/$\sqrt{2}$ value assigned by NHANES; for urinary arsenous acid, this applied to 32.3% of measurements. Serum vitamin B12 and 25-hydroxyvitamin D are reported as descriptive characteristics only and were not entered into regression models. Body mass index was derived from MEC measurements, and smoking status from standardised questionnaires.

Statistical analysis

Analyses were performed in Python 3.11 using statsmodels 0.14. Because element concentrations were right-skewed, continuous variables were summarised as medians with interquartile ranges and compared using the Mann–Whitney U test, and categorical variables using the chi-square test. Because 22 trace elements were compared, the Benjamini–Hochberg false discovery rate procedure was applied across this set of exposure comparisons, and adjusted q values are reported alongside unadjusted p values; no correction was applied to demographic or clinical characteristics, which were not treated as exposure hypotheses.

Associations between element concentrations and diabetes were examined using logistic regression. Element concentrations were natural log-transformed and entered as continuous variables, and odds ratios are reported per interquartile-range increase to render estimates interpretable on the observed exposure scale. Covariates were specified a priori on the basis of established associations with both diabetes and metal exposure, and comprised age, sex, body mass index, race/ethnicity, education, family income-to-poverty ratio, smoking, hypertension, and self-reported kidney disease. Race/ethnicity was dummy coded with non-Hispanic White as reference. Odds ratios with 95% confidence intervals are

reported for every variable entered, including non-significant ones. Variance inflation factors were all below 2.

Two models are presented. Model 1 was unweighted. Model 2 accounted for the complex survey design; in accordance with NHANES analytic guidelines, the weight appropriate to the variable collected on the smallest number of respondents was applied, and because urinary elements were measured in a one-third subsample, this was the subsample weight rather than the MEC examination weight.11 Sampling strata and primary sampling units were incorporated, with variance estimated by Taylor series linearisation and standard errors clustered on primary sampling unit.

Three sensitivity analyses were prespecified. Because spot urinary concentrations depend on dilution, which differs between individuals with and without diabetes owing to osmotic diuresis, both models were refitted with log-transformed urinary creatinine as an additional covariate (Table S1). Second, models were refitted without age to quantify age confounding (Table S5), and urinary arsenous acid was additionally modelled as a binary LOD indicator and in participants with detectable concentrations only (Tables S2a and S2b). Third, models were refitted using an HbA1c-only case definition (Table S3). An exploratory model including all remaining elements differing at $q<0.10$ is presented in Table S4.

Table S1. Multivariable models additionally adjusted for urinary creatinine.

Variable	OR (95% CI)	p	OR (95% CI)	p
	Unweighted		Weighted	
Age (per year)	1.06 (1.04–1.07)	<0.001	1.05 (1.04–1.07)	<0.001
Male sex	1.27 (0.85–1.88)	0.238	1.32 (0.81–2.15)	0.268
BMI (per kg/m^2)	1.11 (1.08–1.14)	<0.001	1.11 (1.07–1.16)	<0.001
Active smoking	1.15 (0.70–1.86)	0.584	0.96 (0.48–1.95)	0.915
Hypertension	0.93 (0.65–1.33)	0.691	1.11 (0.67–1.85)	0.681
Self-reported kidney disease	3.61 (1.69–7.73)	<0.001	1.98 (0.95–4.13)	0.069
Some college or above	0.65 (0.44–0.96)	0.030	0.63 (0.36–1.11)	0.111
Family income-to-poverty ratio	0.99 (0.88–1.12)	0.900	1.02 (0.85–1.22)	0.857
Other Hispanic	2.26 (1.21–4.22)	0.010	1.74 (0.85–3.56)	0.133
Mexican American	2.39 (1.37–4.16)	0.002	1.84 (1.12–3.04)	0.016
Non-Hispanic Black	1.80 (1.10–2.95)	0.019	1.51 (0.79–2.87)	0.208
Non-Hispanic Asian	4.55 (2.42–8.54)	<0.001	5.18 (2.81–9.58)	<0.001
Other/Multiracial	1.04 (0.24–4.41)	0.962	0.59 (0.14–2.58)	0.485
Arsenous acid, urine (per IQR)	0.59 (0.38–0.93)	0.024	0.48 (0.30–0.76)	0.002
Cadmium, urine (per IQR)	0.99 (0.69–1.41)	0.937	1.14 (0.70–1.84)	0.598
Tin, urine (per IQR)	1.25 (0.97–1.61)	0.085	1.37 (1.18–1.59)	<0.001
Molybdenum, urine (per IQR)	1.49 (1.05–2.12)	0.024	1.45 (1.05–2.01)	0.026
Manganese, urine (per IQR)	0.98 (0.87–1.10)	0.694	0.91 (0.78–1.06)	0.206
Selenium, blood (per IQR)	1.25 (1.03–1.52)	0.024	1.05 (0.80–1.37)	0.725
Urinary creatinine (per log unit)	0.75 (0.48–1.18)	0.213	0.91 (0.50–1.67)	0.765

Odds ratios (OR) with 95% confidence intervals (CI) from multivariable logistic regression; diabetes (yes/no) was the dependent variable. n = 1,388 with 192 events. Trace element odds ratios are expressed per interquartile-range increase on the log scale. The weighted model incorporates the urinary subsample weight (WTSA2YR), sampling strata, and primary sampling units. BMI, body mass index; CI, confidence interval; IQR, interquartile range; OR, odds ratio.

Table S2. Alternative handling of values below the limit of detection.

Urinary arsenous acid was below the limit of detection (LOD) in 32.5% of participants, comprising 31.4% of those without and 39.6% of those with diabetes. Because this proportion differed between groups, two alternative specifications were examined.

Table S2a. Arsenous acid modelled as a binary above-versus-below LOD indicator.

Variable	OR (95% CI)	p
Age (per year)	1.06 (1.04–1.07)	<0.001
Male sex	1.11 (0.77–1.60)	0.569
BMI (per kg/m^2)	1.10 (1.07–1.13)	<0.001
Active smoking	1.18 (0.73–1.91)	0.494
Hypertension	0.95 (0.66–1.36)	0.776
Self-reported kidney disease	3.59 (1.69–7.61)	<0.001
Some college or above	0.64 (0.43–0.94)	0.023
Family income-to-poverty ratio	0.99 (0.88–1.11)	0.858
Other Hispanic	2.22 (1.20–4.12)	0.011
Mexican American	2.32 (1.34–4.02)	0.003
Non-Hispanic Black	1.68 (1.03–2.74)	0.036
Non-Hispanic Asian	4.57 (2.50–8.36)	<0.001
Other/Multiracial	1.00 (0.24–4.23)	1.000
Cadmium, urine (per IQR)	0.86 (0.63–1.19)	0.362
Tin, urine (per IQR)	1.18 (0.93–1.50)	0.173
Molybdenum, urine (per IQR)	1.28 (0.93–1.75)	0.127
Manganese, urine (per IQR)	0.97 (0.87–1.09)	0.623
Selenium, blood (per IQR)	1.24 (1.03–1.51)	0.027
Arsenous acid above LOD (vs below)	0.73 (0.48–1.11)	0.136

Unweighted multivariable logistic regression; diabetes (yes/no) was the dependent variable; n = 1,388 with 192 events. Urinary arsenous acid was entered as a binary indicator of detectable versus non-detectable concentration; all other trace elements were retained as continuous log-transformed variables. BMI, body mass index; CI, confidence interval; IQR, interquartile range; OR, odds ratio.

Table S2b. Analysis restricted to participants with detectable concentrations.

Variable	OR (95% CI)	p
Age (per year)	1.06 (1.04–1.08)	<0.001
Male sex	1.09 (0.68–1.75)	0.725
BMI (per kg/m^2)	1.11 (1.07–1.15)	<0.001
Active smoking	1.02 (0.55–1.89)	0.959
Hypertension	0.86 (0.53–1.38)	0.524
Self-reported kidney disease	5.49 (1.82–16.53)	0.002
Some college or above	0.64 (0.39–1.06)	0.081
Family income-to-poverty ratio	0.93 (0.80–1.08)	0.337
Other Hispanic	1.96 (0.84–4.54)	0.118
Mexican American	2.16 (1.08–4.33)	0.031
Non-Hispanic Black	1.90 (1.00–3.63)	0.050
Non-Hispanic Asian	5.87 (2.79–12.37)	<0.001
Other/Multiracial	0.58 (0.06–5.37)	0.631
Arsenous acid, urine (per IQR)	0.71 (0.54–0.93)	0.014
Cadmium, urine (per IQR)	0.87 (0.58–1.30)	0.484
Tin, urine (per IQR)	1.20 (0.90–1.59)	0.220
Molybdenum, urine (per IQR)	1.27 (0.87–1.86)	0.221
Manganese, urine (per IQR)	0.93 (0.74–1.16)	0.495
Selenium, blood (per IQR)	1.26 (0.98–1.62)	0.069

Unweighted multivariable logistic regression restricted to participants with urinary arsenous acid above the limit of detection; n = 937 with 116 events, compared with n = 1,388 and 192 events in the full sample. Trace element odds ratios

are expressed per interquartile-range increase on the log scale. BMI, body mass index; CI, confidence interval; IQR, interquartile range; OR, odds ratio.

Table S3. Multivariable models using an HbA1c-only definition of diabetes.

Variable	OR (95% CI)	p	OR (95% CI)	p
	Unweighted		Weighted	
Age (per year)	1.06 (1.04–1.07)	<0.001	1.06 (1.04–1.08)	<0.001
Male sex	1.55 (1.01–2.39)	0.046	1.35 (0.77–2.39)	0.298
BMI (per kg/m^2)	1.10 (1.06–1.13)	<0.001	1.12 (1.07–1.17)	<0.001
Active smoking	1.18 (0.67–2.09)	0.559	1.10 (0.54–2.22)	0.798
Hypertension	1.06 (0.70–1.61)	0.789	1.15 (0.62–2.13)	0.649
Self-reported kidney disease	2.57 (1.11–5.94)	0.028	1.71 (0.66–4.38)	0.267
Some college or above	0.81 (0.51–1.28)	0.370	0.97 (0.65–1.46)	0.887
Family income-to-poverty ratio	0.93 (0.81–1.07)	0.295	0.90 (0.75–1.07)	0.244
Other Hispanic	2.58 (1.24–5.36)	0.011	2.14 (0.92–4.99)	0.078
Mexican American	2.72 (1.43–5.20)	0.002	2.23 (1.04–4.75)	0.039
Non-Hispanic Black	1.71 (0.96–3.05)	0.068	1.41 (0.64–3.09)	0.396
Non-Hispanic Asian	5.35 (2.65–10.76)	<0.001	7.20 (3.71–13.97)	<0.001
Other/Multiracial	0.48 (0.05–4.61)	0.525	0.14 (0.01–2.34)	0.171
Arsenous acid, urine (per IQR)	0.40 (0.24–0.68)	<0.001	0.39 (0.24–0.66)	<0.001
Cadmium, urine (per IQR)	1.04 (0.71–1.53)	0.822	1.15 (0.74–1.79)	0.526
Tin, urine (per IQR)	0.97 (0.73–1.29)	0.843	1.26 (0.94–1.70)	0.128
Molybdenum, urine (per IQR)	1.59 (1.09–2.32)	0.016	1.56 (1.01–2.42)	0.045
Manganese, urine (per IQR)	0.92 (0.79–1.08)	0.321	0.86 (0.69–1.08)	0.191
Selenium, blood (per IQR)	1.42 (1.13–1.78)	0.002	1.18 (0.90–1.56)	0.237

Odds ratios (OR) with 95% confidence intervals (CI) from multivariable logistic regression; diabetes (yes/no) was the dependent variable. Diabetes was defined by HbA1c ≥ 6.5% alone; n = 1,388 with 126 events, compared with 192 events under the American Diabetes Association definition used in the primary analysis. The reduced number of events limits statistical power in this model. Trace element odds ratios are expressed per interquartile-range increase on the log scale. The weighted model incorporates the urinary subsample weight (WTSA2YR), sampling strata, and primary sampling units. BMI, body mass index; CI, confidence interval; IQR, interquartile range; OR, odds ratio.

Table S4. Exploratory multivariable model including the expanded element panel.

Variable	OR (95% CI)	p	OR (95% CI)	p
	Unweighted		Weighted	
Age (per year)	1.06 (1.04–1.07)	<0.001	1.06 (1.04–1.07)	<0.001
Male sex	1.17 (0.81–1.69)	0.406	1.33 (0.79–2.25)	0.282
BMI (per kg/m^2)	1.10 (1.07–1.13)	<0.001	1.11 (1.07–1.16)	<0.001
Active smoking	1.20 (0.74–1.96)	0.459	0.97 (0.46–2.04)	0.940
Hypertension	0.91 (0.63–1.31)	0.615	1.06 (0.63–1.80)	0.818
Self-reported kidney disease	2.77 (1.28–6.01)	0.010	1.64 (0.80–3.37)	0.174
Some college or above	0.65 (0.44–0.95)	0.028	0.62 (0.36–1.10)	0.102
Family income-to-poverty ratio	0.99 (0.88–1.12)	0.915	1.03 (0.86–1.24)	0.751
Other Hispanic	2.20 (1.17–4.14)	0.015	1.63 (0.77–3.44)	0.202
Mexican American	2.08 (1.17–3.70)	0.012	1.57 (0.89–2.79)	0.123
Non-Hispanic Black	1.57 (0.95–2.61)	0.081	1.42 (0.79–2.56)	0.238
Non-Hispanic Asian	4.15 (2.21–7.83)	<0.001	4.37 (2.33–8.20)	<0.001
Other/Multiracial	0.99 (0.23–4.27)	0.987	0.58 (0.13–2.63)	0.477
Arsenous acid, urine (per IQR)	0.60 (0.37–0.94)	0.028	0.49 (0.31–0.77)	0.002
Cadmium, urine (per IQR)	0.85 (0.61–1.18)	0.327	1.05 (0.71–1.55)	0.814
Tin, urine (per IQR)	1.15 (0.90–1.47)	0.252	1.24 (1.04–1.47)	0.014
Molybdenum, urine (per IQR)	1.43 (1.01–2.04)	0.044	1.44 (0.96–2.15)	0.076
Manganese, urine (per IQR)	0.97 (0.86–1.09)	0.619	0.88 (0.77–1.02)	0.089
Selenium, blood (per IQR)	1.33 (1.09–1.62)	0.005	1.09 (0.82–1.45)	0.550
Uranium, urine (per IQR)	1.51 (1.15–1.98)	0.003	1.70 (1.31–2.22)	<0.001
Iodine, urine (per IQR)	0.98 (0.75–1.28)	0.872	0.96 (0.73–1.25)	0.749
Barium, urine (per IQR)	0.79 (0.55–1.13)	0.198	0.80 (0.45–1.42)	0.446
Strontium, urine (per IQR)	0.80 (0.55–1.16)	0.240	0.84 (0.55–1.28)	0.413

Odds ratios (OR) with 95% confidence intervals (CI) from multivariable logistic regression; diabetes (yes/no) was the dependent variable. n = 1,388 with 192 events. Trace element odds ratios are expressed per interquartile-range increase on the log scale. BMI, body mass index; CI, confidence interval; IQR, interquartile range; OR, odds ratio.

Ethical considerations

The NHANES protocol was approved by the NCHS Research Ethics Review Board (Continuation of Protocol #2011-17); this secondary analysis of publicly available, de-identified data required no further approval.

Results

Participant characteristics

The descriptive sample comprised 1,486 adults, of whom 204 (13.7%) met the criteria for diabetes; 181 reported a physician diagnosis or used glucose-lowering medication, and 23 were identified by HbA1c alone. Participants with diabetes were older (median 63 vs 44

years, p<0.001) and had a higher body mass index (31.2 vs 27.3 kg/m^2, p<0.001). Diabetes prevalence rose from 1.2% among participants aged 20–29 years to 25.2% among those aged 70 years and over. Prevalence did not differ by sex (14.1% in men vs 13.4% in women, p=0.778) but varied by race/ethnicity (p=0.043), being lowest in non-Hispanic White (10.8%) and highest in non-Hispanic Asian participants (17.4%). Participants with diabetes were more likely to have completed no more than high school education (18.7% vs 10.1%, p<0.001), to have hypertension (19.3% vs 10.6%, p<0.001) and to report kidney disease (38.3% vs 12.9%, p<0.001), and were less likely to be current smokers (10.4% vs 14.9%, p=0.036). The family income-to-poverty ratio did not differ significantly (p=0.091, q=0.183). Characteristics and element concentrations are presented in Table 1.

Table 1. Baseline demographic characteristics and trace element levels of the study population stratified by diabetes status.

n=1486	Non-diabetic	Diabetic (n=204)	p	q (FDR)
	n (%*)			
Sex				
Female	652 (86.6)	101 (13.4)	0.778	
Male	630 (85.9)	103 (14.1)		
Ethnicity/Race				
Mexican American	186 (83.8)	36 (16.2)		
Other Hispanic	111 (83.5)	22 (16.5)		
Non-Hispanic White	593 (89.2)	72 (10.8)	**0.043**	
Non-Hispanic Black	226 (83.7)	44 (16.3)		
Non-Hispanic Asian	128 (82.6)	27 (17.4)		
Other and Multi-Racial	38 (92.7)	3 (7.3)		
Smoking Status				
Non-smoker	945 (85.1)	165 (14.9)	**0.036**	
Active smoker	337 (89.6)	39 (10.4)		
Hypertension				
No	850 (89.4)	101 (10.6)	**<0.001**	
Yes	432 (80.7)	103 (19.3)		
Kidney disease				
No	1253 (87.1)	186 (12.9)		

Yes	29 (61.7)	18 (38.3)	**<0.001**	
Education				
High school or less	513 (81.3)	118 (18.7)	**<0.001**	
Some college or above	769 (89.9)	86 (10.1)		
	Median (Q1-Q3)			
Age (years)	44.00 (32.00–60.75)	63.00 (52.00–71.00)	**<0.001**	
Glycohemoglobin (%)	5.40 (5.20–5.70)	6.80 (6.20–7.93)	**<0.001**	
BMI (kg/m^2)	27.30 (23.90–31.70)	31.20 (27.15–36.35)	**<0.001**	
Family income-to-poverty ratio	2.19 (1.06–4.36)	1.93 (1.03–3.79)	0.091	
Sedentary activity (min/day)	420.00 (240.00–540.00)	480.00 (300.00–480.00)	0.333	
Sleep duration (h)	7.00 (6.00–8.00)	7.00 (6.00–8.00)	0.531	
Vitamin B12 (pmol/L)	378.95 (275.30–518.80)	376.75 (280.27–598.70)	0.266	
25-hydroxyvitamin D2 + D3 (nmol/L)	63.00 (46.05–80.88)	63.70 (44.38–82.05)	0.799	
Arsenous acid, urine (ug/L)	0.48 (0.08–0.73)	0.29 (0.08–0.64)	**<0.001**	**0.004**
Total arsenic, urine (ug/L)	6.39 (3.12–13.12)	6.75 (3.60–12.78)	0.363	0.419
Uranium, urine (ug/L)	0.00 (0.00–0.01)	0.01 (0.00–0.01)	**<0.001**	**<0.001**
Tungsten, urine (ug/L)	0.05 (0.02–0.11)	0.06 (0.03–0.11)	0.052	0.119
Thallium, urine (ug/L)	0.15 (0.08–0.24)	0.14 (0.08–0.21)	0.224	0.345
Strontium, urine (ug/L)	87.43 (48.48–146.19)	74.70 (41.57–127.65)	**0.008**	**0.026**
Tin, urine (ug/L)	0.37 (0.17–0.81)	0.69 (0.26–1.52)	**<0.001**	**<0.001**
Antimony, urine (ug/L)	0.04 (0.02–0.07)	0.04 (0.03–0.08)	**0.027**	0.068
Lead, urine (ug/L)	0.32 (0.17–0.55)	0.37 (0.22–0.56)	**0.028**	0.068
Manganese, urine (ug/L)	0.09 (0.09–0.13)	0.09 (0.09–0.14)	0.140	0.246
Molybdenum, urine (µg/L)	32.61 (17.05–61.77)	37.35 (22.99–63.33)	0.061	0.122
Cesium, urine (ug/L)	4.09 (2.40–6.39)	4.02 (2.53–6.30)	0.986	0.986
Cobalt, urine (ug/L)	0.37 (0.22–0.62)	0.38 (0.24–0.67)	0.230	0.345
Cadmium, urine (ug/L)	0.17 (0.08–0.35)	0.23 (0.14–0.45)	**<0.001**	**<0.001**
Barium, urine (ug/L)	0.96 (0.50–2.01)	0.70 (0.35–1.59)	**<0.001**	**<0.001**
Mercury, urine (ug/L)	0.26 (0.09–0.56)	0.21 (0.09–0.48)	0.186	0.310
Iodine, urine (ug/L)	115.45 (65.03–216.78)	142.70 (86.00–239.15)	**0.004**	**0.015**
Manganese, blood (nmol/L)	167.08 (134.86–209.85)	175.18 (131.22–215.99)	0.346	0.415
Selenium, blood (umol/L)	2.48 (2.31–2.66)	2.49 (2.32–2.71)	0.251	0.325
Mercury, blood (nmol/L)	3.90 (2.10–8.10)	4.20 (2.10–8.55)	0.339	0.415
Cadmium, blood (nmol/L)	2.58 (1.60–5.07)	2.76 (1.85–4.11)	0.509	0.566
Lead, blood (umol/L)	0.05 (0.03–0.07)	0.05 (0.04–0.08)	0.097	0.183

Data are median (25th–75th percentile) or n (%); percentages are calculated within rows. Continuous variables were compared using the Mann–Whitney U test and categorical variables using

the chi-square test. The Benjamini–Hochberg false discovery rate procedure was applied across the 22 trace element comparisons, and adjusted q values are shown for these variables only. Urinary element concentrations are expressed as µg/L and blood elements as µmol/L, except blood manganese, mercury, and cadmium (nmol/L). Values below the limit of detection were retained at the LOD/√2 value assigned by NHANES; for urinary arsenous acid, this applied to 32.3% of measurements. Hypertension was defined as a recorded diagnosis or measured elevated blood pressure, and kidney disease as a self-reported physician diagnosis. Analyses are unweighted. BMI, body mass index; HbA1c, glycohaemoglobin; 25(OH)D, 25-hydroxyvitamin D; LOD, limit of detection.

Trace element concentrations by diabetes status

Seven of the 22 trace elements examined differed between groups at q<0.05. Participants with diabetes had higher urinary cadmium (0.23 vs 0.17 µg/L, q<0.001), tin (0.69 vs 0.37 µg/L, q<0.001), uranium (q<0.001) and iodine (142.70 vs 115.45 µg/L, q=0.015), and lower urinary barium (0.70 vs 0.96 µg/L, q<0.001), strontium (74.70 vs 87.43 µg/L, q=0.026) and arsenous acid (0.29 vs 0.48 µg/L, q=0.004). Urinary antimony, lead, and molybdenum differed before but not after correction (q=0.068, 0.068, and 0.122). Blood selenium did not differ between groups (2.49 vs 2.48 µmol/L, p=0.251, q=0.325), nor did total urinary arsenic (p=0.363), blood cadmium (p=0.509), blood lead (p=0.097), or blood mercury (p=0.339).

Multivariable analyses

Models were fitted in 1,388 participants, of whom 192 had diabetes (Table 2). Age and body mass index were the strongest correlates in both models (Model 1: OR 1.06 per year, 95% CI 1.04–1.07; OR 1.10 per kg/m^2, 1.07–1.13; both p<0.001). Non-Hispanic Asian participants had markedly higher odds than non-Hispanic White participants (Model 1: OR 5.02, 2.73–9.26; Model 2: OR 5.36, 3.01–9.55; both p<0.001), as did Mexican American participants in the weighted model (OR 1.87, 1.13–3.09, p=0.015). The estimate for non-Hispanic Black participants was readily obtained (Model 1: OR 1.73, 1.07–2.82, p=0.027; 41 of 255 had diabetes). Self-reported kidney disease was associated with diabetes in the unweighted model only (OR 3.52, 1.65–7.51, p=0.001; weighted p=0.069). Neither sex, smoking, hypertension, nor the income-to-poverty ratio was independently associated with diabetes.

Among trace elements, urinary arsenous acid was inversely associated with diabetes in both models (Model 1: OR 0.55, 0.36–0.86, p=0.008; Model 2: OR 0.47, 0.30–0.72, p<0.001). Blood selenium was associated in the unweighted model only (OR 1.26, 1.04–1.53, p=0.020; weighted OR 1.05, 0.81–1.37, p=0.695), whereas urinary tin and molybdenum were associated only under the survey design (tin OR 1.35, 1.14–1.60, p<0.001; molybdenum OR 1.41, 1.00–2.00, p=0.049). Urinary cadmium, which differed markedly between groups in unadjusted analysis, showed no independent association in either model (Model 1: OR 0.89, 0.65–1.23, p=0.481; Model 2: OR 1.10, 0.76–1.58, p=0.614). Urinary manganese was not associated with diabetes in any specification.

Sensitivity analyses

The null result for urinary cadmium was attributable to age adjustment: omitting age from an otherwise identical model produced a strong positive association (OR 1.56, 1.17–2.07, p=0.002), which disappeared when age was included. Urinary tin was similarly age-sensitive (OR 1.31, 1.04–1.64 without age), whereas the associations for arsenous acid and selenium persisted in both specifications (Table S5).

Adjustment for log-transformed urinary creatinine modestly attenuated the arsenous acid association but did not abolish it (unweighted OR 0.59, 0.38–0.93, p=0.024; weighted OR 0.48, 0.30–0.77, p=0.002), and creatinine itself was not independently associated with diabetes (p=0.213 and p=0.765). Urinary molybdenum became significant in the unweighted model after dilution adjustment (OR 1.49, 1.05–2.12, p=0.024), and urinary cadmium remained null throughout (Table S1).

Table S5. Effect of age adjustment on trace element–diabetes associations.

Variable	OR (95% CI)	p	OR (95% CI)	p
	Age excluded		Age included	
Male sex	1.31 (0.92–1.86)	0.133	1.15 (0.80–1.66)	0.445
BMI (per kg/m^2)	1.08 (1.06–1.11)	<0.001	1.10 (1.07–1.13)	<0.001
Active smoking	0.62 (0.40–0.96)	0.033	1.19 (0.73–1.92)	0.486
Hypertension	1.34 (0.95–1.89)	0.091	0.93 (0.65–1.33)	0.679
Self-reported kidney disease	4.26 (2.09–8.65)	<0.001	3.52 (1.65–7.51)	0.001
Some college or above	0.54 (0.37–0.78)	0.001	0.64 (0.44–0.95)	0.027
Family income-to-poverty ratio	1.01 (0.90–1.13)	0.865	0.99 (0.88–1.12)	0.883
Other Hispanic	1.73 (0.95–3.18)	0.074	2.32 (1.25–4.32)	0.008
Mexican American	1.75 (1.04–2.95)	0.035	2.46 (1.42–4.28)	0.001
Non-Hispanic Black	1.32 (0.83–2.10)	0.246	1.73 (1.06–2.82)	0.027
Non-Hispanic Asian	3.95 (2.21–7.04)	<0.001	5.02 (2.73–9.26)	<0.001
Other/Multiracial	0.57 (0.15–2.17)	0.413	1.01 (0.24–4.32)	0.985
Arsenous acid, urine (per IQR)	0.39 (0.26–0.60)	<0.001	0.55 (0.35–0.86)	0.008
Cadmium, urine (per IQR)	1.56 (1.17–2.07)	0.002	0.89 (0.65–1.23)	0.481
Tin, urine (per IQR)	1.31 (1.04–1.64)	0.020	1.18 (0.93–1.50)	0.166
Molybdenum, urine (per IQR)	1.08 (0.79–1.47)	0.620	1.37 (0.99–1.89)	0.054
Manganese, urine (per IQR)	1.00 (0.89–1.12)	0.970	0.97 (0.87–1.09)	0.645
Selenium, blood (per IQR)	1.22 (1.01–1.48)	0.042	1.26 (1.04–1.53)	0.020

Odds ratios (OR) with 95% confidence intervals (CI) from unweighted multivariable logistic regression; diabetes (yes/no) was the dependent variable. n = 1,388 with 192 events. Trace element odds ratios are expressed per interquartile-range increase on the log scale. BMI, body mass index; CI, confidence interval; IQR, interquartile range; OR, odds ratio.

The inverse association for arsenous acid also persisted when the exposure was modelled as a binary detection indicator, when analysis was restricted to participants with detectable concentrations, and when diabetes was defined by HbA1c alone. An exploratory model including all elements differing at q<0.10 yielded a positive association for urinary uranium. Full results are provided in Tables S2a, S2b, S3 and S4.

Table 2. Multivariable associations between trace element concentrations and diabetes mellitus

Variable	Model 1 Unweighted		Model 2 Weighted	
	OR (95% CI)	p	OR (95% CI)	p
Age (per year)	1.058 (1.043–1.073)	<0.001	1.053 (1.038–1.069)	<0.001
Male sex	1.153 (0.801–1.659)	0.445	1.280 (0.763–2.148)	0.349
BMI (per kg/m^2)	1.102 (1.074–1.131)	<0.001	1.113 (1.068–1.160)	<0.001
Active smoking	1.187 (0.733–1.920)	0.486	0.977 (0.477–2.001)	0.950
Hypertension	0.927 (0.648–1.327)	0.679	1.109 (0.663–1.854)	0.693
Self-reported kidney disease	3.523 (1.653–7.511)	0.001	1.971 (0.947–4.098)	0.069
Some college or above	0.645 (0.438–0.950)	0.027	0.631 (0.357–1.114)	0.112
Income-to-poverty ratio	0.991 (0.880–1.116)	0.883	1.017 (0.850–1.215)	0.857
Other Hispanic	2.322 (1.247–4.323)	0.008	1.746 (0.858–3.553)	0.124
Mexican American	2.463 (1.417–4.279)	0.001	1.866 (1.128–3.085)	0.015
Non-Hispanic Black	1.734 (1.065–2.823)	0.027	1.492 (0.804–2.769)	0.205
Non-Hispanic Asian	5.024 (2.727–9.257)	<0.001	5.359 (3.007–9.551)	<0.001
Other/Multiracial	1.014 (0.238–4.321)	0.985	0.590 (0.136–2.550)	0.480
Arsenous acid, urine (per IQR)	0.552 (0.355–0.858)	0.008	0.466 (0.303–0.717)	<0.001
Cadmium, urine (per IQR)	0.891 (0.647–1.228)	0.481	1.098 (0.764–1.578)	0.614
Tin, urine (per IQR)	1.183 (0.933–1.501)	0.166	1.348 (1.137–1.597)	<0.001
Molybdenum, urine (per IQR)	1.370 (0.994–1.887)	0.054	1.414 (1.001–1.998)	0.049
Manganese, urine (per IQR)	0.973 (0.868–1.092)	0.645	0.906 (0.777–1.056)	0.207
Selenium, blood (per IQR)	1.259 (1.037–1.527)	0.020	1.054 (0.812–1.368)	0.695

Odds ratios (OR) with 95% confidence intervals (CI) from multivariable logistic regression; diabetes (yes/no) was the dependent variable. Analyses include 1,388 participants, of whom 192 had diabetes. Both models were adjusted for age, sex, body mass index, race/ethnicity, education, family income-to-poverty ratio, smoking, hypertension, and self-reported kidney disease; all covariates entered are reported. Model 1 is unweighted; Model 2 accounts for the complex sampling design using the urinary subsample weight, strata, and primary sampling units, with Taylor series linearisation and standard errors clustered on primary sampling units. Element concentrations were natural log-transformed, and odds ratios are expressed per interquartile-range increase on the log scale, obtained by multiplying the regression coefficient by log(Q3) − log(Q1). The corresponding

observed concentration ranges (Q1–Q3, with the fold-increase they represent) were: urinary arsenous acid 0.08–0.72 µg/L (9.0-fold); urinary cadmium 0.08–0.36 µg/L (4.5-fold); urinary tin 0.18–0.90 µg/L (5.0-fold); urinary molybdenum 17.4–61.9 µg/L (3.6-fold); urinary manganese 0.09–0.13 µg/L (1.4-fold); blood selenium 2.32–2.67 µmol/L (1.2-fold). Reference categories: non-Hispanic White and female. Variance inflation factors were all below 2. BMI, body mass index; CI, confidence interval; IQR, interquartile range; OR, odds ratio.

Discussion

In this cross-sectional analysis of 1,388 US adults, urinary arsenous acid was inversely associated with diabetes, blood selenium was associated in unweighted but not design-based analysis, and urinary tin and molybdenum were associated only when the survey design was applied. Urinary cadmium, which differed substantially between groups in unadjusted comparisons, showed no independent association once age was taken into account. These results indicate that associations between trace element biomarkers and diabetes in cross-sectional survey data are highly sensitive to model specification, and that the direction and significance of individual estimates can depend as much on analytical choices as on the underlying exposure.

Age confounding and the cadmium association

The most consistent finding of this analysis was a negative one. Urinary cadmium was strongly associated with diabetes when age was omitted from an otherwise identical model, and showed no association whatever when age was included. The explanation is straightforward: urinary cadmium reflects cumulative lifetime burden and rises across the adult lifespan, while diabetes prevalence in this sample increased from 1.2% among participants in their twenties to over 25% among those aged 70 and above. Any analysis that does not fully account for age will therefore tend to recover a positive cadmium–diabetes association irrespective of whether a causal relationship exists.

This may explain the heterogeneity of the published literature, in which estimates for cadmium and type 2 diabetes span a wide range, and NHANES analyses have variously

reported positive, null and inverse associations.[6,9,15] Our findings do not exclude a contribution of cadmium to metabolic dysfunction, since experimental evidence for beta-cell toxicity remains substantial,[23,24] but they suggest that cross-sectional biomarker studies are poorly suited to detecting it.

Urinary arsenous acid

Urinary arsenous acid was inversely associated with diabetes across every specification examined, a direction opposite to that predicted by experimental work and to observations in populations with high exposure through drinking water.[12,14,21] Several explanations warrant consideration, and we cannot distinguish between them. Spot urinary concentrations depend on dilution, and osmotic diuresis in diabetes dilutes urinary analytes; this is the most obvious candidate for an artefactual inverse association. However, adjustment for urinary creatinine only modestly attenuated the association; creatinine itself was not independently associated with diabetes, and urinary tin and molybdenum, subject to the same dilution pressures, were associated in the opposite direction. A third of measurements fell below the limit of detection, and although the association persisted among participants with detectable concentrations, restricted analytical range remains a source of uncertainty. Reverse causation also cannot be excluded, since dietary sources of arsenic such as rice and seafood may differ between individuals with and without diagnosed diabetes. We therefore regard this finding as requiring replication in prospective data with dietary assessment rather than as evidence of a protective effect.

Selenium, tin and molybdenum

Blood selenium was associated with diabetes in the unweighted model but not after incorporation of the survey design, and did not differ between groups after correction for multiple testing. The direction of the adjusted estimate is consistent with reports of increased diabetes risk at higher selenium status despite its antioxidant function.[3,5,25] Our models specified a log-linear term and therefore cannot evaluate non-linearity; in exploratory analyses neither a quadratic term (p=0.528) nor quartile categorisation

provided evidence of a U-shaped relationship (Table S6). Proposed mechanisms centre on overexpression of glutathione peroxidase-1, which may deplete the transient hydrogen peroxide signalling required for inhibition of protein tyrosine phosphatase 1B and thereby attenuate insulin receptor phosphorylation.[26] These mechanisms derive from experimental systems and were not evaluated here; they are potential rather than established pathways. Given the discrepancy between weighted and unweighted estimates and the narrow observed selenium range in this population (interquartile range 2.32–2.67 μmol/L, a 1.2-fold difference), the present data are compatible with, but not confirmatory of, this relationship.

Urinary tin and molybdenum were associated with diabetes under the survey design, and molybdenum additionally in the unweighted model after adjustment for dilution. A previous NHANES analysis reported a positive association for tin, which the present findings support.[16] Because these associations were not robust across all specifications, both warrant further investigation rather than being regarded as established.[17,18]

Non-Hispanic Asian participants had approximately fivefold higher odds of diabetes after full adjustment, consistent with documented ethnic differences in risk at a given body mass index.[22,27] Element exposure also varies by ethnic group, making ethnicity a potential confounder here.

Implications and limitations

The evidence base itself, rather than any individual risk factor, is what these findings most clearly address. Environmental and occupational exposure to toxic metals remains a legitimate target for population-level prevention through tobacco control, emission standards and food supply monitoring. Our results nonetheless suggest that cross-sectional biomarker associations provide a weak foundation for exposure-specific policy claims and should be reproduced under adequate confounder control and appropriate survey methodology. We did not evaluate supplement use, screening performance, or clinical utility, and these data do not support recommendations regarding routine biomonitoring or selenium supplementation.

Strengths include a large, nationally representative survey with standardised laboratory methods, simultaneous measurement of a broad element panel permitting mutual adjustment, correction for multiple comparisons, use of the appropriate subsample weight with design-based variance estimation, and systematic evaluation of model specification. The cross-sectional design nonetheless precludes causal inference and cannot distinguish exposure preceding disease from its consequences. Spot urine provides a single measurement that may not represent long-term exposure. Dietary intake, occupational history, and diabetes duration were unavailable, and residual confounding cannot be excluded. The analytic sample represents the subset of examined adults with complete data, and selective missingness may have introduced bias.

Conclusion

In US adults, urinary arsenous acid was inversely associated with diabetes, while blood selenium, urinary tin, and urinary molybdenum showed associations that varied according to model specification. The association between urinary cadmium and diabetes was fully accounted for by age. These findings represent associations rather than causation; the mechanisms discussed derive from experimental literature. The central implication is methodological: estimates of the relationship between trace element biomarkers and diabetes in survey data are sensitive to adjustment for age, to the handling of urinary dilution and values below the limit of detection, and to incorporation of the sampling design. Prospective studies are required before the contribution of environmental metal exposure to diabetes can be established.

Ethical Considerations: The National Health and Nutrition Examination Survey protocol was approved by the Research Ethics Review Board of the National Center for Health Statistics (Continuation of Protocol #2011-17, approved 26 October 2010; extended for the 2013–2014 cycle), and written informed consent was obtained from all participants at the time of the original survey. The present study is a secondary analysis of publicly available, fully de-identified data and involved no additional contact with participants; it was therefore

exempt from further institutional review board approval. All procedures were conducted in accordance with the Declaration of Helsinki.

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

References

1. Landrigan PJ, Fuller R, Acosta NJR, et al. The Lancet Commission on pollution and health. *Lancet.* 2018;391(10119):462-512 (doi:10.1016/s0140-6736(17)32345-0).
2. Pyrzynska K, Sentkowska A. Selenium Species in Diabetes Mellitus Type 2. *Biol Trace Elem Res.* 2024;202(7):2993-3004. doi:10.1007/s12011-023-03900-z.
3. Stranges S, Marshall JR, Natarajan R, et al. Effects of long-term selenium supplementation on the incidence of type 2 diabetes: a randomized trial. *Ann Intern Med.* 2007;147(4):217-23. doi:10.7326/0003-4819-147-4-200708210-00175.
4. Hansen AF, Simić A, Åsvold BO, et al. Trace elements in early phase type 2 diabetes mellitus-A population-based study. The HUNT study in Norway. *J Trace Elem Med Biol.* 2017;40:46-53. doi:10.1016/j.jtemb.2016.12.008.
5. Lee KH, Jeong D. Bimodal actions of selenium essential for antioxidant and toxic pro-oxidant activities: the selenium paradox (Review). *Mol Med Rep.* 2012;5(2):299-304. doi:10.3892/mmr.2011.651.
6. Tinkov AA, Filippini T, Ajsuvakova OP, et al. The role of cadmium in obesity and diabetes. *Sci Total Environ.* 2017;601-602:741-55. doi:10.1016/j.scitotenv.2017.05.224
7. Onat T, Demir Caltekin M, Turksoy VA, et al. The Relationship Between Heavy Metal Exposure, Trace Element Level, and Monocyte to HDL Cholesterol Ratio with Gestational Diabetes Mellitus. *Biol Trace Elem Res.* 2021;199(4):1306-15. doi:10.1007/s12011-020-02499-9
8. Iritas SB, Dinc AH, Dip A, Unal BM, Ertan B, Soylemezoglu T. Levels of toxic metals and trace elements in autopsy liver tissue samples. *Med Sci.* 2017;6(2):242-8.
9. Gong R, Pu X, Cheng Z, Ding J, Chen Z, Wang Y. The association between serum cadmium and diabetes in the general population: A cross-sectional study from NHANES (1999–2020). *Front. Nutr..* 2022;9:966500. doi:10.3389/fnut.2022.966500
10. Johnson CL, Paulose-Ram R, Ogden CL, et al. National health and nutrition examination survey: analytic guidelines, 1999-2010. *Vital Health Stat 2.* 2013(161):1-24.

11. National Center for Health Statistics. National Health and Nutrition Examination Survey: Analytic Guidelines, 2011-2014 and 2015-2016. Hyattsville, MD: National Center for Health Statistics; December 14, 2018.
12. Davey JC, Nomikos AP, Wungjiranirun M, et al. Arsenic as an endocrine disruptor: arsenic disrupts retinoic acid receptor-and thyroid hormone receptor-mediated gene regulation and thyroid hormone-mediated amphibian tail metamorphosis. *Environ Health Perspect.* 2008;116(2):165-72. doi:10.1289/ehp.10131
13. Kirkley AG, Carmean CM, Ruiz D, et al. Arsenic exposure induces glucose intolerance and alters global energy metabolism. *Am J Physiol Regul Integr Comp Physiol.* 2018;314(2):R294-r303. doi:10.1152/ajpregu.00522.2016
14. Peng Q, Harlow SD, Park SK. Urinary arsenic and insulin resistance in US adolescents. *Int J Hyg Environ Health.* 2015;218(4):407-13 (doi:10.1016/j.ijheh.2015.03.006).
15. Satarug S. Is Environmental Cadmium Exposure Causally Related to Diabetes and Obesity? *Cells.* 2023;13(1). doi:10.3390/cells13010083
16. Liu B, Sun Y, Lehmler HJ, Bao W. Association between urinary tin concentration and diabetes in nationally representative sample of US adults. *J Diabetes.* 2018;10(12):977-83. doi:10.1111/1753-0407.12798
17. Rovenko BM, Perkhulyn NV, Lushchak OV, Storey JM, Storey KB, Lushchak VI. Molybdate partly mimics insulin-promoted metabolic effects in Drosophila melanogaster. *Comp Biochem Physiol C Toxicol Pharmacol.* 2014;165:76-82. doi:10.1016/j.cbpc.2014.06.002
18. Li XT, Yu PF, Gao Y, et al. Association between Plasma Metal Levels and Diabetes Risk: a Case-control Study in China. *Biomed Environ Sci.* 2017;30(7):482-91. doi:10.3967/bes2017.064
19. Alian F, Asemi-Esfahani Z, Fazilati M, Habibollahi S, Palizban A. Evaluation of Blood Concentrations of Copper and Molybdenum in Type 2 Diabetic Patients. *J. Biol.* 2017;6:139-45. doi:10.15412/J.JBTW.01060801
20. Hildebrand J, Thakar S, Watts T-L, et al. Environmental Cadmium Exposure and Type 2 Diabetes Mellitus Risk: An Overview of Systematic Reviews. *Expo. Health.* 2022;14(3):743-62. doi:10.1007/s12403-021-00449-x

21. Srisuporn P, Navasumrit P, Ngaotepprutaram T, Chaisatra K, Hunsonti P, Ruchirawat M. Arsenic exposure alters the expression of genes related to metabolic diseases in differentiated adipocytes and in newborns and children. *Int J Hyg Environ Health.* 2023;250:114124. doi:10.1016/j.ijheh.2023.114124
22. Nagai K, Chung HF, Hayashi K, et al. The Association Between Race/Ethnicity and Risk of Type 2 Diabetes in Women Varies by BMI: A Pooled Analysis of Individual Data From 15 Cohort Studies. *Diabetes Care.* 2025. doi:10.2337/dc25-1478
23. Gungor H, Kara H. Effects of selenium, zinc, insulin and metallothionein on cadmium-induced oxidative stress and metallothionein gene expression levels in diabetic rats. *J Basic Clin Physiol Pharmacol.* 2020;31(2). doi:10.1515/jbcpp-2019-0198
24. Attia SM, Varadharajan K, Shanmugakonar M, Das SC, Al-Naemi HA. Cadmium: An Emerging Role in Adipose Tissue Dysfunction. *Expo. Health.* 2022;14(1):171-83. doi:10.1007/s12403-021-00427-3
25. Rodriguez-Hernandez Z, Bel-Aguilar J, Moreno-Franco B, et al. Differential association of selenium exposure with insulin resistance and β-cell function in middle age and older adults. *Nutr Diabetes.* 2025;15(1):5. doi:10.1038/s41387-025-00361-2
26. Zhao J, Zou H, Huo Y, Wei X, Li Y. Emerging roles of selenium on metabolism and type 2 diabetes. *Front Nutr.* 2022;9:1027629. doi:10.3389/fnut.2022.1027629
27. Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and Trends in Diabetes Among Adults in the United States, 1988-2012. *Jama.* 2015;314(10):1021-9. doi:10.1001/jama.2015.10029

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):406-422 // 10.5505/amj.2026.40035

PERCEIVED STRESS AND PREPAREDNESS IN PEDIATRIC PREHOSPITAL CARE AMONG EMERGENCY MEDICAL SERVICES PERSONNEL: A CROSS-SECTIONAL STUDY

Ümit Alper Özcan[1], Ramiz Yazıcı[2], Hüseyin Mutlu[3], Medine Akkan Öz[4], Uğur Şakar[5], Yunus Yatmaz[6], Özgür Ömer Yıldız[7]

[1]Department of Pediatrics, Bilkent City Hospital, Ankara, Türkiye
[2]Department of Emergency Medicine, University of Health Sciences, Kanuni Sultan Süleyman Education and Research Hospital, İstanbul, Türkiye
[3]Department of Emergency Medicine, Aksaray University Medical School, Aksaray, Türkiye
[4]Department of Emergency Medicine, Gülhane Training and Research Hospital, Ankara, Türkiye
[5]Department of Emergency Medicine, Dışkapı Yıldırım Beyazıt Training and Research Hospital, University of Health Sciences, Ankara, Türkiye
[6]Department of Emergency Medicine, Yenimahalle Training and Research Hospital, Ankara, Türkiye
[7]Department of Thoracic Surgery, Atatürk Sanatoryum Training and Research Hospital, Ankara, Türkiye

Correspondence:
Ramiz Yazıcı (e-mail: dr.ramiz.yazici@gmail.com)

Submitted: 28.06.2026 // Accepted: 14.08.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: Pediatric emergencies are relatively uncommon in prehospital care, but they may cause considerable stress among healthcare providers because of limited exposure, high perceived clinical risk, and age-specific management requirements. This study aimed to evaluate perceived stress during pediatric patient care among prehospital healthcare providers and to examine its relationship with perceived knowledge, clinical self-confidence, pediatric case exposure, and perceived team support.

Materials and Methods: This cross-sectional questionnaire-based study included prehospital healthcare providers working in Ankara Provincial Emergency Medical Services. The online questionnaire was distributed to 312 eligible providers, and 187 participants were analyzed, yielding a response rate of 59.9%. Data were collected using an anonymous questionnaire covering sociodemographic and professional characteristics and five Likert-type subscales. Internal consistency was assessed using Cronbach's alpha. Since subscale scores were not normally distributed, Spearman correlation analysis was used to examine relationships between variables.

Results: The mean perceived stress score was 3.19±1.10, indicating a moderate-to-high level of stress. Overall, 48.1% of participants had high perceived stress scores. Perceived stress was negatively correlated with perceived knowledge, clinical self-confidence, and pediatric case exposure. Team support had the highest mean score among the subscales but was not directly correlated with perceived stress. More than half of the participants reported exposure to only 0–5 pediatric cases during the previous year.

Conclusion: Prehospital healthcare providers experience considerable stress when caring for pediatric patients. Lower perceived knowledge, reduced clinical self-confidence, and limited pediatric case exposure were associated with higher perceived stress. Targeted pediatric training and repeated simulation-based practice may improve preparedness in prehospital pediatric care.

Keywords: Emergency medical services, pediatrics, psychological stress, self-confidence, health personnel, prehospital care.

Introduction

Prehospital healthcare is a critical area of healthcare that is carried out in unpredictable conditions with limited resources and under severe time constraints.[1,2] Healthcare workers providing care under these circumstances face various traumatic and stressful situations during their careers. Although most of these stressors are related to adult patient populations, current literature consistently shows that pediatric patients cause unparalleled and disproportionately high levels of stress for prehospital healthcare providers.[1,3] In the US, it has been reported that anxiety is the second most common cause threatening patient safety during interventions involving pediatric patients, and that this anxiety is caused by cognitive overload, difficulty in decision-making, and communication problems.[1] Reports also show that even during simulated pediatric cardiac arrest scenarios, paramedics experience a 41.3% increase in immediate anxiety levels.[2] There are many related underlying mechanisms for this intense stress caused by pediatric cases in the prehospital environment. The leading mechanism is the rarity of exposure to pediatric cases. Epidemiological data consistently show that pediatric patients account for a rather small proportion of all emergency healthcare calls.[4,5] According to the US national database, pediatric calls comprise 7.4% of all emergency healthcare calls.[4] Also in the US, retrospective analysis of 371,746 emergency healthcare calls revealed that pediatric patients represent only 5.9% of the total cohort[5]. This low exposure rate feeds the vicious cycle known as "paradox of rarity", where clinical conditions requiring the highest levels of skills provide very little opportunity for practice.

As a natural consequence of this vicious cycle, gaps in knowledge and low clinical self-confidence are widely reported in prehospital healthcare providers (PHPs). A national needs assessment for PHPs has highlighted that pediatric airway management and provider anxiety experienced during pediatric interventions are defined as priority training requirements and that there are critical gaps in knowledge due to lack of exposure.[3] A qualitative study conducted with 42 PHPs revealed that participants are aware of their own lack of knowledge and that continuing education programs are insufficient in terms of

pediatric content.[6] In a similar study, PHPs clearly report a negative impact on pediatric airway management caused by lack of rare exposure and insufficient pediatric training and remark on the necessity of recurrent training programs to prevent skill and knowledge erosion.[7] A meta-analysis reported that low self-confidence and perception of self-sufficiency are associated with low rates of initiating treatment in pediatric patients, and that this can have adverse effects on quality of care.[8] This theory suggests that apart from knowledge, clinical self-confidence may be an independent determinant of stress during pediatric interventions.

When all this literature is assessed as a whole, it is clear that perceived stress in prehospital pediatric care is multidimensional and cannot be explained by a single factor. The relationship between stress and potential protective factors such as knowledge level, clinical self-confidence, exposure to pediatric cases, and team support must be assessed through a holistic approach. However, in current literature, there are no studies assessing all four of these variables concurrently, particularly none that review PHPs and pediatric cases in the Ankara Provincial Emergency Healthcare Services. This study aims to bridge this gap by assessing the perceived stress levels of PHPs during pediatric interventions and examining the relationship between this stress and level of knowledge, clinical self-confidence, exposure to pediatric cases, and perceived team support.

Materials and Methods

Study design and participants

This study is a single-center, cross-sectional, descriptive, and analytical questionnaire study designed to evaluate the perceived stress levels of the prehospital healthcare providers (PHPs) working in the Ankara Provincial Emergency Healthcare Services (EHS) and the relationship between this stress and level of knowledge, clinical self-confidence, exposure to pediatric cases, and perceived team support. The study was approved by the Ankara Provincial Health Directorate Ethics Committee (Decision No: 2026-01-5, Date: 28/01/2026). The study was conducted in accordance with the principles of the Helsinki

Declaration. Informed consent was acquired from the PHPs actively working in the Ankara Provincial EHS between March 2026 and April 2026, who agreed to participate in the study, through the information sheet at the beginning of the questionnaire, prior to participation. PHPs actively working in the Ankara Provincial Emergency Healthcare Services who were 18 years of age or older and volunteered to participate in the study were included. Participants who did not complete the study, who wanted to withdraw during the data collection phase, or left more than 50% of the questionnaire incomplete were excluded from the study.

The online questionnaire was distributed to 312 eligible prehospital healthcare providers working in Ankara Provincial Emergency Medical Services during the study period. A total of 187 participants completed the questionnaire and met the inclusion criteria. No participant was excluded from the final study sample because of incomplete responses or withdrawal. Therefore, the final analytical sample consisted of 187 participants, and the overall response rate was 59.9% (187/312). One participant entered an invalid age value; therefore, age-related analyses were performed using data from 186 participants, while all other analyses were conducted with the full sample of 187 participants.

Data collection tools

Data were collected using a structured questionnaire developed by the researchers based on a review of the literature on pediatric prehospital care, provider stress, clinical confidence, pediatric case exposure, and team-based care.[1,3-8] Items addressing perceived stress were informed by reports of heightened provider anxiety during pediatric calls[1,2]; items on perceived knowledge and clinical self-confidence by studies documenting pediatric knowledge gaps and low provider confidence [3,6,8]; items on pediatric case exposure by data on the rarity of pediatric calls and the resulting skill decay [4,5,7]; and items on perceived team support by qualitative reports of communication and teamwork difficulties during pediatric prehospital calls.[1,6] Since there was no single validated instrument specifically addressing perceived stress and preparedness-related factors during pediatric prehospital interventions in the local emergency medical services setting, an item pool was created to

cover five conceptual domains: perceived stress, perceived knowledge, clinical self-confidence, pediatric case exposure and experience, and perceived team support.

The questionnaire consisted of two parts. The first part included questions on sociodemographic and professional characteristics: age, gender, professional title, total duration of professional experience, duration of prehospital work experience, and the number of pediatric cases managed in the previous year. Professional title was categorized as paramedic, emergency medical technician (EMT), physician, nurse, or other. Total professional experience and prehospital work experience were categorized as 0-2 years, 3-5 years, 6-10 years, and ≥11 years. The number of pediatric cases managed in the previous year was categorized as 0-5, 6-10, 11-20, and >20.

The second part included 25 Likert-type items grouped into five subscales, with five items in each subscale: perceived stress, perceived knowledge, clinical self-confidence, pediatric case exposure and experience, and perceived team support. Each item was scored on a 5-point Likert scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). Higher scores indicated a higher level of the relevant construct. Sample items were as follows: “Treating pediatric patients causes more stress than treating adult patients” for perceived stress; “I have sufficient theoretical knowledge about pediatric emergencies” for perceived knowledge; “I can safely perform the initial assessment of a pediatric patient” for clinical self-confidence; “I feel that my stress level decreases as I gain more experience with pediatric cases” for pediatric case exposure and experience; and “I believe that cooperation within the team facilitates the intervention process in pediatric cases” for perceived team support.

One reverse-coded item was included in the clinical self-confidence subscale, and one reverse-coded item was included in the pediatric case exposure and experience subscale. Reverse-coded items were recoded as 6 minus the original score before calculating subscale scores. Each subscale score was calculated as the mean of the relevant items. Subscale scores ranged from 1.00 to 5.00.

Before data collection, the draft questionnaire was reviewed by three experts with experience in emergency medicine, pediatric emergency care, and prehospital emergency medical services. The experts evaluated the items in terms of relevance, clarity, comprehensibility, and domain appropriateness. Based on their feedback, minor linguistic revisions were made to improve clarity, while the five-domain structure of the questionnaire was preserved.

A pilot test was conducted with 30 prehospital healthcare providers who were not included in the final analysis. The pilot test aimed to assess item clarity, approximate completion time, and the technical functionality of the online form. No major structural changes were required after the pilot test; however, minor wording revisions were made to improve readability.

Internal consistency reliability was assessed using Cronbach's alpha coefficient for each subscale. No exploratory or confirmatory factor analysis was performed; therefore, the subscale structure reflects the conceptual domains defined by the research team rather than an empirically verified factor structure.

The complete item list of the questionnaire is provided as a supplementary file. The questionnaire was administered anonymously via Google Forms and did not collect personally identifiable information. The aim of the study was explained to the participants at the beginning of the form. Participants were informed that their responses would be used only for research purposes and that no individual performance assessment would be conducted. Electronic informed consent was obtained before participation. Participants were able to leave any question unanswered and withdraw from the study at any time.

Statistical analysis

All statistical analyses were performed using IBM SPSS Statistics (IBM Corp., Armonk, NY, USA). For descriptive statistics, continuous variables were calculated as mean, standard deviation, median, minimum, and maximum values; categorical variables as frequency and percentage distribution. The internal consistency reliability of subscales was assessed using

the Cronbach alpha coefficient. The normality of the data was examined using the Shapiro-Wilk test; based on the results obtained, the subscale scores did not follow a normal distribution, so non-parametric statistical methods were chosen. Mann-Whitney U test was used for the comparison of two independent groups, and the Kruskal-Wallis H test was used for comparing more than two independent groups. For Kruskal-Wallis results found to be significant, multiple comparisons were performed using a Bonferroni-corrected Mann-Whitney U test.

Multiple linear regression analysis was performed to identify independent factors associated with perceived stress during pediatric patient care. The perceived stress subscale score was entered as the dependent variable. Perceived knowledge, clinical self-confidence, pediatric case exposure, and perceived team support scores were included as independent variables. All predictors were entered into the model simultaneously using the enter method. Regression results were reported as standardized beta coefficients (β) and p-values. Multicollinearity was assessed using tolerance and variance inflation factor (VIF) values. To assess the relationship between variables, Spearman's rank correlation analysis was performed. In all analyses, $p < 0.05$ was considered to be statistically significant.

Results

Of the 312 eligible prehospital healthcare providers invited to participate, 187 completed the questionnaire and were included in the final analysis, yielding a response rate of 59.9%. No participant was excluded from the final study sample due to incomplete responses or withdrawal. Age data were valid for 186 participants because one participant entered an invalid age value. The mean age of the participants was 34.24 ± 5.78 years, and 59.9% (n=112) were male. More than half the participants were paramedics (50.8%, n=95), and 42.8% (n=80) were EMTs. In terms of professional experience, most (72.7%, n=136) had 11 years or more of professional experience. When the number of pediatric cases attended to in the past year was assessed, 58.8% (n=110) of participants had seen only 0-5 pediatric

cases. The sociodemographic and professional characteristics of the participants are shown in Table 1.

Table 1. Sociodemographic and professional characteristics of the participants (n=187)

Variable	Category	n	%	Mean ± SD
Age (year)	-	186*	-	34.24 ± 5.78
Gender	Male	112	59.9	-
	Female	75	40.1	-
Professional title	Paramedic	95	50.8	-
	EMT	80	42.8	-
	Physician	7	3.7	-
	Nurse	4	2.1	-
	Other	1	0.5	-
Total professional experience	0-2 years	3	1.6	-
	3-5 years	21	11.2	-
	6-10 years	27	14.4	-
	≥11 years	136	72.7	-
Years working in prehospital	0-2 years	13	7.0	-
	3-5 years	31	16.6	-
	6-10 years	39	20.9	-
	≥11 years	104	55.6	-
Number of pediatric cases attended in the past year	0-5	110	58.8	-
	6-10	17	9.1	-
	11-20	20	10.7	-
	>20	40	21.4	-

** The age of 1 participant was entered incorrectly*
Data is presented as mean ± SD or n (%)
SD: Standard deviation; n: number; EMT: Emergency medical technician

Perceived level of stress was calculated as 3.19 ± 1.10; this value corresponds to a medium-high level of stress on the 5-point Likert scale. While 48.1% (n=90) of participants reported high stress levels (> 3.50), 33.7% (n=63) reported medium (2.01-3.50) and 18.2% (n=34) reported low levels (≤ 2.00). An item-based analysis showed that the "Treating pediatric patients causes more stress than treating adult patients" item received the highest stress score (3.53 ± 1.22), and the "I feel that my stress levels affect my performance" item received the lowest (2.94 ± 1.23). The mean score of perceived level of knowledge was 3.64 ± 0.80, and the highest self-confidence score (3.88 ± 0.86) was obtained from the "I can safely perform initial assessment of a pediatric patient" item. The descriptive analysis, internal consistency reliability coefficient (Cronbach's alpha), and results of the normality test relating to the five sub-scales used in the study are shown in Table 2.

Table 2. Descriptive statistics, reliability and normality test results of the sub-scales

Sub-scale	n	Mean ± SD	Med	Min	Max	α	S-W p
Perceived Stress	187	3.19 ± 1.10	3.20	1.00	5.00	0.931	<0.001
Perceived Knowledge	187	3.64 ± 0.80	3.60	1.00	5.00	0.913	<0.001
Clinical Self-confidence	187	3.67 ± 0.68	3.60	1.60	5.00	0.778	0.002
Exposure to Cases	187	3.45 ± 0.60	3.40	1.60	5.00	0.704	0.013
Team Support	187	3.88 ± 0.85	4.00	1.00	5.00	0.945	<0.001

n: number; α: Cronbach's alpha, internal consistency coefficient; S-W: Shapiro-Wilk normality test; Med: Median; SD: Standard deviation

Correlation analysis revealed a negative correlation between perceived stress levels and perceived knowledge levels ($p < 0.001$), and between perceived stress and clinical self-confidence ($p < 0.001$). A similar negative correlation was observed between perceived stress levels and exposure to pediatric cases ($p=0.001$). The assessment of the relationship between independent variables revealed a positive correlation between perceived knowledge levels and clinical self-confidence, between perceived knowledge levels and exposure to pediatric cases, and between clinical self-confidence and exposure to pediatric

cases (p< 0.001). Results of the correlation analysis and multiple linear regression results are shown in Table 3.

Table 3. Spearman correlation matrix and multiple linear regression results

Variable	1	2	3	4	5	β (Regression)
Perceived Stress	1.000					-
Perceived Knowledge	-0.249***	1.000				-0.095
Clinical Self-confidence	-0.263***	0.644***	1.000			-0.101
Exposure to cases	-0.237**	0.548***	0.570***	1.000		-0.219*
Team Support	-0.007	0.350***	0.340***	0.360***	1.000	0.206**

*Correlation coefficients are Spearman rho values. The regression column shows standardized beta coefficients obtained from multiple linear regression analysis. *p<0.05; **p<0.01; ***p<0.001.*

Discussion

Providing care during critical incidents is the most common known cause of stress, anxiety, and even post-traumatic stress disorder (PTSD) among hospital workers.[9] It has been reported that the prevalence of stress and anxiety among prehospital healthcare providers (PHPs) is 10% worldwide.[9-10] In the less-controlled prehospital environment where outpatient care is provided to acutely ill patients, it can be expected that this stress will similarly increase when caring for pediatric patients.[11] In our study, we also found that PHPs perceive medium-high levels of stress (mean 3.19 ± 1.10) when providing care for pediatric patients, and that almost half the participants (48.1%) were in the high stress group. At the item level, the highest-scoring statement was the one expressing that pediatric patients cause more stress than adult patients. Since adult encounters were not assessed in this study, this reflects the participants' own comparative perception rather than a measured difference

between the two patient groups; nevertheless, it indicates that pediatric cases are perceived as a distinct source of stress by PHPs.

The large-scale study by Guise et al., conducted in 44 states in the US with 753 PHPs, revealed that provider anxiety is the second most common factor affecting pediatric patient safety and that this anxiety can cause cognitive overload, difficulty in decision-making, and communication problems.[1] Similarly, a multi-centered randomized controlled study conducted by Lacour et al. with 150 high-level paramedics from 14 different EHS' showed a 41,3% increase in the anxiety levels of the control group during simulated pediatric cardiac arrest scenarios.[2] McGahern et al., studying the effects of outcome letters delivered to 470 paramedics in Canada, reported that pediatric cases cause long-term emotional distress for providers and outcome notifications relieve this distress.[12] Therefore, the high stress levels found in our study are not unique to our country and reflect the universal challenges of pediatric care.

In our study, we identified a negative correlation between level of pediatric knowledge and stress. This finding is consistent with other evidence indicating that gaps in knowledge increase anxiety. Hansen et al. conducted the national Delphi study with 737 emergency medical professionals and found that pediatric airway management and anxiety experienced while attending to children are priority training needs and that there are critical gaps in knowledge due to lack of exposure.[3] In their focus group study with 42 emergency healthcare professionals from urban, suburban, and rural areas, Brown et al. revealed that participants were aware of their own lack of knowledge and knew that current continuing training programs were insufficient in terms of pediatric content.[6] Lyng et al. clearly expressed that pediatric management skills of PHPs were adversely affected by rare clinical exposure and insufficient pediatric training and that repeat training programs were necessary to prevent loss of skill and knowledge.[7] After conducting a simulation curriculum on 187 PHPs through 5 academic centers in Massachusetts, Farrell et al. reported a 9.8% increase in knowledge levels and a 93% increase in the self-confidence of participants.[13] The common indication of

these studies is that pediatric knowledge is both a variable that can be directly influenced by training and also, that it has a defining role in reducing perceived stress levels.

Clinical self-confidence showed the strongest inverse association with perceived stress among the preparedness-related variables in our bivariate analysis (rho = -0.263). This is in line with the scoping review by Fowler et al., which reported that pediatric cases frequently evoke anxiety and low confidence among paramedics across different countries, independently of the health system in which they work.[8] It is also notable that in the simulation curriculum reported by Farrell et al., self-confidence increased by 93% while knowledge increased by only 9.8%,[13] suggesting that self-confidence may respond more rapidly to structured practice than knowledge does. In systems such as ours, where more than half of the providers encountered five or fewer pediatric cases in a year, simulation-based repetition may therefore be the more accessible route to reducing perceived stress.

Our study determined that 58.8% of participants have seen only 0-5 pediatric cases in the past year. This rate is consistent with international epidemiological data. Using NEMSIS data, Diggs et al. revealed that pediatric calls comprised only 7.4% of all emergency health calls in the US nationwide.[4] In their retrospective study reviewing 371,746 emergency health calls, Rampogal et al. determined that pediatric patients accounted for only 5.9% of the total cohort and that documentation of vital signs was significantly more incomplete in children compared to adults.[5] In the national assessment involving 8,166 emergency healthcare centers, Hewes et al. reported that 80% of the institutions responded to less than 100 cases annually.[14] These low exposure rates reflect a vicious cycle known, in the literature, as the "paradox of rarity", where conditions requiring the highest skill levels provide the least opportunity for practice. The negative correlation between exposure to pediatric cases and perceived stress established in our study materializes the clinical effects of this paradox.

Maloney et al. reported that the peer-support program, implemented over 20 months, resulted in no statistically significant difference in the burnout sub-scale scores of PHP's pre- or post-program periods[15]. Studying the effects of insecure attachments on the stress of ambulance workers, Halpern et al. found that coping mechanisms and social support do not

play a mediating role in this relationship and that individual psychological characteristics are more decisive than team support in acute stress responses.[16] We believe that perceived team support having the highest mean score (3.88 ± 0.85) among all sub-scales in our study is not because team support is experienced, but because acute stress experienced during pediatric interventions is caused by internal reasons such as perceived individual competence and lack of experience.

Our finding that the highest-scoring item concerned pediatric patients being more stressful than adult patients is supported by international data based on within-provider comparisons. In a German cross-sectional survey of 1,005 EMS providers reporting on emergencies involving children with life-limiting conditions, 60.4% reported feeling overwhelmed during pediatric calls, significantly more often than during adult calls, and only 26.9% felt confident about future pediatric calls compared with 75.6% for adult calls.[17] The parallel with our results is notable: in both settings the deficit is concentrated in pediatric-specific confidence rather than in general professional competence, supporting the interpretation that perceived stress in this context arises from case-specific unfamiliarity rather than from general occupational burden.

Although perceived team support was not significantly correlated with perceived stress in the bivariate analysis, it showed a positive association in the multivariable regression model after adjustment for perceived knowledge, clinical self-confidence, and pediatric case exposure. This finding should be interpreted cautiously. One possible explanation is a suppression effect caused by intercorrelations among the preparedness-related variables. Another possibility is that personnel who experience higher stress during pediatric cases may become more aware of the need for team support. Therefore, this result should not be interpreted as evidence that team support increases stress, but rather as an indication that the relationship between team dynamics and perceived stress may be more complex than a direct linear association.

Two features of our data support this interpretation. First, perceived team support received the highest mean score of all subscales (3.88 ± 0.85), with a median of 4.00, indicating that

most participants perceived their teams as supportive; a protective factor that is already perceived favorably by the majority offers limited room to differentiate high- and low-stress providers. Second, team support correlated positively and significantly with perceived knowledge, clinical self-confidence, and pediatric case exposure (rho = 0.340-0.360), yet showed essentially no bivariate association with perceived stress (rho = -0.007). This pattern suggests that team support operates on the preparedness-related domain rather than directly on the affective stress response, which is consistent with Halpern et al., who reported that individual psychological characteristics were more decisive than social support in acute stress reactions among ambulance workers,[16] and with Maloney et al., who found no significant change in burnout subscale scores after a 20-month peer support program.[15] Taken together, these findings indicate that team support may be necessary but not sufficient for reducing pediatric-specific stress, and that interventions targeting individual competence are likely to be required alongside it.

Considering the negative correlation of both perceived knowledge and clinical self-confidence with perceived stress, our findings suggest that targeted training has the potential to reduce stress during pediatric interventions. Lacour et al. showed that the mobile application for drug doses reduced stress by 27% during simulated pediatric resuscitation and that this effect was independent of professional experience[2]. Farrell et al. reported that after state-wide implementation of a simulated curriculum, knowledge and self-confidence both increased simultaneously.[13]

Our study has limitations. Firstly, due to the cross-sectional design, causality between variables could not be established. Secondly, data collection using self-reporting scales carries the risk of social desirability bias. Thirdly, the study was conducted in a single EMS system and included only personnel of the Ankara Provincial Ambulance Service; since staffing, pediatric call volume, and in-service training differ between provinces and national EMS systems, our findings reflect one regional service rather than prehospital providers in general. In addition, the questionnaire was newly developed for this study and has not undergone a formal validation process; although expert review, pilot testing, and acceptable

internal consistency support its usability, its construct validity was not established, and the subscale scores should therefore be regarded as preliminary indicators rather than validated measures. Furthermore, perceived stress was assessed only in relation to pediatric encounters, with no parallel measurement for adult encounters; any comparison between the two patient groups therefore rests on participants' self-reported perception rather than on measured data. Finally, the use of convenience sampling further limits the generalizability of the findings. Multicenter studies are needed before these associations can be considered generalizable. Despite all these limitations, our study is one of the first studies to assess the stress experienced by prehospital healthcare providers during pediatric interventions in Türkiye from a multi-dimensional perspective.

Conclusion

In this cross-sectional study of 187 PHPs working in the Ankara Provincial EHS, we found that PHPs perceived medium-high levels of stress (mean 3.19 ± 1.10) during pediatric interventions, and that almost half the participants (48.1%) were in the high stress group. We think that stress levels of PHPs increase when pediatric knowledge levels are low and that increasing knowledge through training would be effective in reducing the perception of stress. For future studies, assessing pre- and post- training stress levels of PHPs during pediatric interventions and planning multi-centered studies involving PHPs from different provinces would increase the generalizability of the findings.

Ethical Considerations: Ethical approval for the study was obtained from the Ankara Provincial Health Directorate Ethics Committee (Decision No: 2026-01-5, Date: January 28, 2026). The study was conducted in accordance with the Declaration of Helsinki. Informed consent was obtained electronically from all participants before participation.

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

References

1. Guise JM, Hansen M, O'Brien K, et al. Emergency medical services responders' perceptions of the effect of stress and anxiety on patient safety in the out-of-hospital emergency care of children: a qualitative study. BMJ Open. 2017;7(2):e014057. doi: 10.1136/bmjopen-2016-014057
2. Lacour M, Bloudeau L, Combescure C, et al. Impact of a mobile app on paramedics' perceived and physiologic stress response during simulated prehospital pediatric cardiopulmonary resuscitation: study nested within a multicenter randomized controlled trial. JMIR Mhealth Uhealth. 2021;9(10):e31748. doi: 10.2196/31748
3. Hansen M, Meckler G, Dickinson C, et al. Children's Safety Initiative: a national assessment of pediatric educational needs among emergency medical services providers. Prehosp Emerg Care. 2015;19(2):287-91. doi: 10.3109/10903127.2014.959223
4. Diggs LA, Sheth-Chandra M, De Leo G. Epidemiology of pediatric prehospital basic life support care in the United States. Prehosp Emerg Care. 2016;20(2):230-8. doi: 10.3109/10903127.2015.1076099
5. Ramgopal S, Elmer J, Escajeda J, Martin-Gill C. Differences in prehospital patient assessments for pediatric versus adult patients. J Pediatr. 2018;199:200-5.e6. doi: 10.1016/j.jpeds.2018.03.069
6. Brown SA, Hayden TC, Randell KA, Rappaport L, Stevenson MD, Kim IK. Improving pediatric education for emergency medical services providers: a qualitative study. Prehosp Disaster Med. 2017;32(1):20-6. doi: 10.1017/S1049023X16001230
7. Lyng J, Harris M, Mandt M, et al. Prehospital pediatric respiratory distress and airway management training and education: an NAEMSP position statement and resource document. Prehosp Emerg Care. 2022;26(sup1):102-10. doi: 10.1080/10903127.2021.1992551
8. Fowler J, Beovich B, Williams B. Improving paramedic confidence with paediatric patients: a scoping review. Australas J Paramedicine. 2018;15(1):1-12. doi: 10.33151/ajp.15.1.559

9. De Boer J,Lok A, Van't Verlaat EV, et al. Work-related critical incidents in hospital-based health care providers and the risk of post-traumatic stress symptoms, anxiety, and depression: a meta-analysis. Soc Sci Med 2011;73:316–26. doi:10.1016/j.socscimed.2011.05.009
10. Berger W, Coutinho ES, Figueira I, et al. Rescuers at risk: a systematic review and meta-regression analysis of the worldwide current prevalence and correlates of PTSD in rescue workers. Soc Psychiatry Psychiatr Epidemiol. 2012;47(6):1001-11. doi:10.1007/s00127-011-0408-2
11. LeBlanc VR, MacDonald RD, McArthur B, King K, Lepine T. Paramedic performance in calculating drug dosages following stressful scenarios in a human patient simulator. Prehosp Emerg Care. 2005;9(4):439-44. doi:10.1080/10903120500255255
12. McGahern C, Cantor Z, De Mendonca B, et al. Closing the loop: the value of outcome letters for prehospital pediatric care. Pediatr Emerg Care. 2024;40(4):261-4. doi: 10.1097/PEC.0000000000002978
13. Farrell C, Dorney K, Mathews B, et al. A statewide collaboration to deliver and evaluate a pediatric critical care simulation curriculum for emergency medical services. Front Pediatr. 2022;10:903950. doi: 10.3389/fped.2022.903950
14. Hewes HA, Ely M, Richards R, Shah MI, et al. Ready for children: assessing pediatric care coordination and psychomotor skills evaluation in the prehospital setting. Prehosp Emerg Care. 2019;23(4):510-8. doi: 10.1080/10903127.2018.1542472
15. Maloney LM, Hoffman J, Peralta E, et al. Informal peer support and intentional acts of kindness may attenuate the impact of work-related stressors on compassion satisfaction, secondary traumatic stress, and burnout of emergency medical services clinicians. Air Med J. 2024;43(4):333-9. doi: 10.1016/j.amj.2024.03.005
16. Halpern J, Maunder RG, Schwartz B, Gurevich M. Attachment insecurity, responses to critical incident distress, and current emotional symptoms in ambulance workers. Stress Health. 2012;28(1):51-60. doi: 10.1002/smi.1401
17. Hauch H, El Mohaui N, Vaillant V, et al. Prehospital emergency medicine for children receiving palliative home care in Germany: a cross-sectional, exploratory study of EMS providers. *Front Pediatr.* 2023;11:1104655. doi:10.3389/fped.2023.1104655

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):423-444 // 10.5505/amj.2026.43788

A RETROSPECTIVE DESCRIPTIVE REVIEW AND TITLE-BASED THEMATIC CONTENT ANALYSIS OF FAMILY MEDICINE SPECIALTY THESES IN TÜRKİYE: 2015–2024

Muhammed Fatih Baran[1], Selma Pekgör[1]

[1]University of Health Sciences, Konya City Hospital, Family Medicine Department, Konya, Türkiye

Correspondence:
Muhammed Fatih Baran (e-mail: mbaran355@gmail.com)

Submitted: 08.04.2026 // Accepted: 11.08.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: This study aimed to evaluate family medicine specialty theses in Türkiye (2015–2024) in terms of institutional contribution, author sex distribution, and topic trends.

Materials and Methods: The Council of Higher Education National Thesis Center (YÖKTEZ) was searched using filters "Family Medicine" (department), "Medical Specialty Thesis" (type), and "Permitted" (access) for 01 Jan 2015–31 Dec 2024. After excluding non-eligible/duplicate/non-accessible records, 4,715 theses were included. Year, institution type (University of Health Sciences (UHS) vs other universities), author sex (inferred from names where possible), and topic categories were assigned using a predefined title-based thematic framework. Chi-square tests were applied; $p<0.05$ was considered significant.

Results: Overall, 64.2% of theses were authored by females (n=3,024) and 35.8% by males (n=1,691). UHS accounted for 45.8% (n=2,160) and other universities for 54.2% (n=2,555); sex distribution did not differ by institution type (p=0.075). The annual number of theses increased from 115 in 2015 to 767 in 2024 and reached its highest level in 2023 with 839 theses. The UHS share exceeded 50% after 2022, rising from 16.5% in 2015 to 53.0% in 2022, 53.3% in 2023, and 50.2% in 2024. The most common topic was primary care practice (25.9%, n=1,223), followed by attitudes/perceptions/approaches (9.3%, n=440) and infectious diseases (8.2%, n=387). Primary care practice remained the leading topic throughout the study period, although its annual proportion varied significantly from 27.8% in 2015 to 18.9% in 2019 and 29.3% in 2024 ($p<0.001$). Significant year-by-year changes were observed for primary care practice ($p<0.001$), infectious diseases ($p<0.001$), and other clinical topics ($p<0.001$). Infectious diseases showed a pandemic-period increase, rising from 4.6% in 2020 to 13.7% in 2021 and 13.5% in 2022, before decreasing to 3.8% in 2024. Knowledge/awareness topics increased from 0.9% in 2015 to 5.4% in 2023 and 3.8% in 2024 (p=0.041). Other clinical topics increased from 0% in 2015–2016 to 5.9% in 2024 ($p<0.001$). Linear regression showed a significant increase in annual thesis production (β=84.92, 95% CI=66.07–103.76, $R^2=0.931$, $p<0.001$). Infectious disease-related theses were more frequent in the pandemic/post-pandemic period than in the pre-pandemic period (OR=1.56, 95% CI=1.20–2.02, $p<0.001$).

Conclusion: Between 2015 and 2024, family medicine thesis production increased, female authorship remained predominant, UHS contributed increasingly, and topic patterns demonstrated statistically significant, albeit proportionally modest, shifts—particularly during the pandemic period. These findings demonstrate both a quantitative expansion in thesis production and significant temporal changes in institutional and thematic patterns in family medicine specialty theses in Türkiye.

Keywords: Family medicine, family medicine specialization education, family medicine theses, thematic content analysis, descriptive review

Introduction

Family medicine is a medical discipline that prioritizes continuity, comprehensiveness, and easy accessibility of primary care services and follows individuals throughout their lifetime without discrimination by age, sex, or disease status.[1,2,3]

In Türkiye, medical specialty residents are required to write a thesis in their specialty field to complete their training.[4] During the 2010s, there was a marked increase in the number of registered theses reported to the Council of Higher Education (YÖK) by Ministry of Health–affiliated training and research hospitals. Indeed, a study examining family medicine theses between 2000 and 2020 identified 1,628 theses and reported a substantial increase in annual thesis numbers after 2014.[5]

Similarly, Mengüllüoğlu and Ünlüoğlu, who evaluated family medicine specialty theses between 2005 and 2015, reported that Istanbul-based training and research hospitals and a limited number of universities drove production; however, topics were largely repetitive and survey-based, and publication rates were limited. These findings suggest that family medicine theses can be used as a mirror of primary care practice in the country, while also highlighting the need to increase research diversity and methodological depth.[6]

A similar pattern is observed in other bibliometric thesis/article studies directly or indirectly related to family medicine: In medical specialty theses on health literacy, family medicine and public health departments were predominant; most studies were cross-sectional; and topics rapidly became more specific after certain years.[7]

However, a shared limitation of existing studies is that they either stop in the early 2000s or fail to capture institutional expansion after 2016, when the University of Health Sciences (UHS) began reporting a large number of family medicine specialty theses to the YÖK system.[1]

In particular, the increase in both the total number of theses after 2015 and the growing share of UHS indicate the need for updated analyses in this field. Moreover, changing interests over time in a rapidly evolving world may have led to differences by sex and topic. Therefore, this study aimed to conduct a retrospective descriptive review and title-based thematic content analysis of family medicine theses produced in recent years to describe institutional contribution, the sex distribution of authors, and trends in thesis topics.

Materials and Methods

In this study, we conducted a retrospective descriptive thesis review of theses and title-based thematic content analysis of medical specialty theses in family medicine completed between 1 January 2015 and 31 December 2024 and registered with metadata in the Council of Higher Education National Thesis Center (YÖK National Thesis Center; YÖKTEZ) in Türkiye. The online search interface of the YÖK National Thesis Center was used as the data source. Searches were performed via the National Thesis Center homepage under the Search module using the following filters: Department/Discipline: "Family Medicine", Thesis Type: "Medical Specialty", Permission Status: "Permitted", and then filtering by year; retrieved records were merged.

Author sex was determined based on the first names recorded in the YÖKTEZ metadata. Names that are conventionally used as female or male names in Türkiye were classified accordingly. For names with potential ambiguity, spelling variants, or limited sex-specific use, additional verification was performed using publicly available name-use information and, when available, contextual information within the thesis record. After this verification process, all authors were classified as female or male; therefore, no record was excluded from the analysis because of undetermined sex.

Topic classification was performed using a predefined thematic coding framework developed by the research team before the final analysis. A total of 11 primary categories were established. The initial categories were created after reviewing family medicine

thesis titles, common primary care research areas, and previously published bibliometric studies on medical specialty theses. Each thesis was assigned to one primary topic category according to the main focus reflected in its title. When a title included more than one possible topic, the dominant research focus was prioritized. For example, theses primarily addressing diabetes, obesity, or thyroid diseases were classified under “metabolic/endocrine disorders,” whereas studies focusing on vaccination, COVID-19, or other communicable diseases were classified under “infectious diseases.” Survey-based studies primarily evaluating knowledge, attitudes, perceptions, awareness, or approaches were classified according to their dominant construct. Theses focusing on specific clinical branches other than the main targeted categories (e.g., general surgery, dermatology, neurology) were classified as 'other clinical topics', whereas non-clinical, purely administrative, or highly heterogeneous subjects were grouped under 'other topics'.

To improve standardization, the coding rules and category definitions were discussed before classification. A second researcher independently reviewed a randomly selected subset comprising approximately 10% of the total theses (n=471) to assess inter-rater reliability. Disagreements between reviewers were resolved through discussion and consensus. Inter-rater agreement was evaluated using Cohen’s kappa coefficient. The inter-rater agreement was excellent, with a Cohen’s kappa value of 0.86.

Theses explicitly indicating the family medicine department/clinic were included directly. Theses registered under a different institute or department but addressing family medicine/primary care service delivery in the title or abstract were excluded due to department mismatch. Theses without access permission, duplicate records, and theses determined to be unrelated to family medicine were excluded. A total of 6175 theses were initially identified. After excluding 1460 records (1362 due to being unrelated to family medicine, 24 duplicates, and 74 lacking access permissions), a final total of 4,715 theses were included in the study.

Ethics approval

Ethics approval was not required because this study was based exclusively on publicly available thesis metadata obtained from the YÖKTEZ database. No human participants were recruited, no intervention was performed, and no personal, sensitive, or identifiable private data were collected. The analyzed information consisted only of publicly accessible bibliometric variables such as thesis year, institution, author name as recorded in the database, and thesis title/topic. Therefore, the study did not involve human subject research requiring ethics committee approval.

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics for Windows, Version 27.0. In statistical analyses, descriptive data were presented as numbers (n) and percentages (%). The Chi-square test was used to compare thesis counts and sex distribution by year. Sex distribution by institution type was also evaluated using the Chi-square test.

To increase analytical depth, additional temporal analyses were performed. Linear regression analysis was used to evaluate the trend in the annual number of theses between 2015 and 2024. For pre–post pandemic comparisons, the study period was divided into two intervals: 2015–2019 as the pre-pandemic period and 2020–2024 as the pandemic/post-pandemic period. Binary logistic regression analyses were performed to assess whether the pandemic/post-pandemic period was associated with the likelihood of major topic categories compared with the pre-pandemic period. Results were reported as odds ratios (ORs), 95% confidence intervals (CIs), and p-values. Institution type was also compared between the two periods using the Chi-square test, and ORs with 95% CIs were calculated.

Results

A total of 4,715 medical specialty theses completed in the field of family medicine between 2015 and 2024 were included in the study. Overall, 64.2% (n=3,024) were authored by female authors and 35.8% (n=1,691) by male authors (Table 1). The proportion of female authors was higher in each year. Of the theses, 45.8% (n=2,160) were conducted within UHS and 54.2% (n=2,555) at other universities. The annual number of theses increased from 115 in 2015 to 767 in 2024 and reached its highest level in 2023 with 839 theses. As illustrated in Figure 1, despite minor fluctuations, there has been a steady and marked upward trend in overall thesis production throughout the study period.

Table 1. Distribution of Total Theses by Year, Sex and Institution Type

	Sex distribution by year		**Institution distribution by**		
Year	**Male**	**Female**	**Other**	**UHS**	**Total**
	n (%)	**n (%)**	**n (%)**	**n (%)**	**n (%)**
2015	54 (46.9)	61 (53.1)	96 (83.5)	19 (16.5)	115 (2.4)
2016	70 (35.1)	129 (64.9)	170 (85.4)	29 (14.6)	199 (4.2)
2017	93 (36.7)	160 (63.3)	145 (57.3)	108 (42.7)	253 (5.4)
2018	87 (31.0)	193 (69.0)	169 (60.4)	111 (39.6)	280 (5.9)
2019	131 (33.5)	260 (66.5)	228 (58.3)	163 (41.7)	391 (8.3)
2020	163 (37.6)	270 (62.4)	241 (55.7)	192 (44.3)	433 (9.2)
2021	271 (37.0)	461 (63.0)	400 (54.6)	332 (45.4)	732 (15.5)
2022	274 (38.8)	432 (61.2)	332 (47.0)	374 (53.0)	706 (15.0)
2023	287 (34.2)	552 (65.8)	392 (46.7)	447 (53.3)	839 (17.8)
2024	261 (34.0)	506 (66.0)	382 (49.8)	385 (50.2)	767 (16.3)
Total	1691 (35.8)	3024 (64.2)	2555 (54.2)	2160 (45.8)	4715 (100)

UHS: University of Health Sciences. Row percentages were used, excluding the total column.

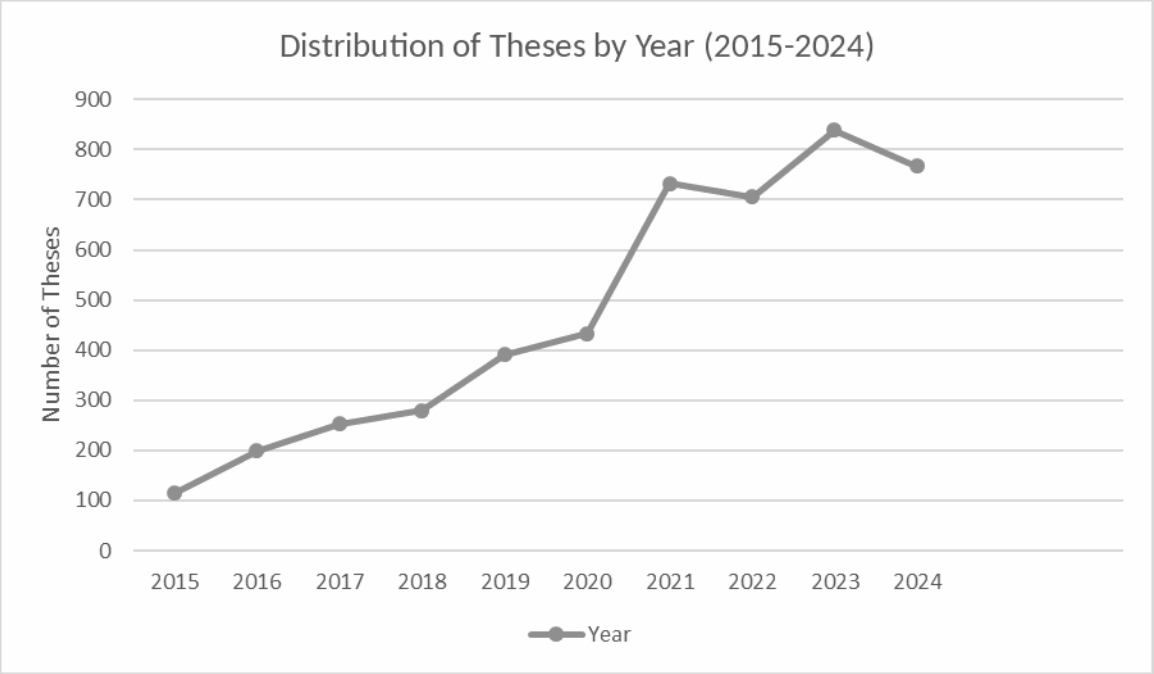


Figure 1. Annual Number of Family Medicine Theses in Türkiye (2015–2024)

In the distribution of total thesis counts by university, UHS ranked first with 2,160 theses. This was followed by Trakya University (n=145), Ondokuz Mayıs University (n=128), Dokuz Eylül University (n=128), and İzmir Kâtip Çelebi University (n=124) (Table 2).

Table 2. Top 10 institutions by number and percentages of family medicine theses

Institution	Number of theses	Percentage (%)
UHS	2160	45.8
Trakya University	145	3.1
Ondokuz Mayıs University	128	2.7
Dokuz Eylül University	128	2.7
İzmir Kâtip Çelebi University	124	2.6
İstanbul Medeniyet University	93	2.0
Aydın Adnan Menderes University	88	1.9
Atatürk University	83	1.8
Kocaeli University	80	1.7
Çukurova University	79	1.7

UHS: University of Health Sciences

When thesis counts were examined by institution over time, other universities predominated in the early years of the study period, whereas the share of UHS exceeded 50% after 2022 (Table 1, Figure 2).

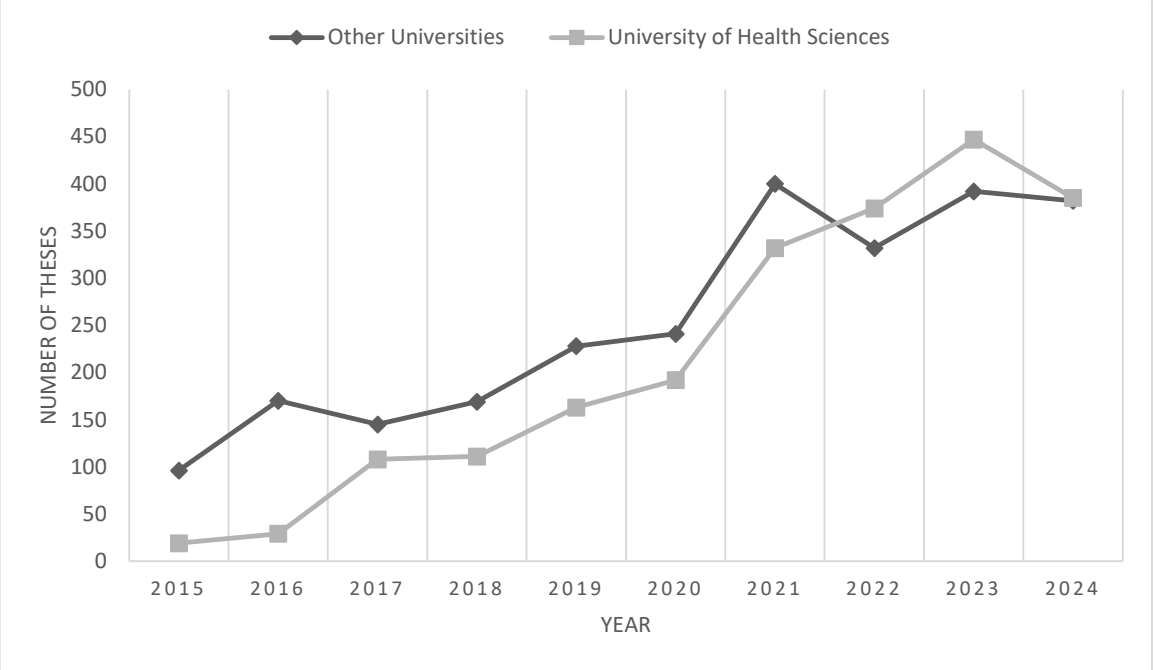


Figure 2. Distribution of Thesis Counts by Institution Type Across Years

In the distribution by topic, the most frequently studied area was primary care practice (25.9%; n=1,223). This was followed by attitudes, perceptions, and approaches (9.3%; n=440), infectious diseases (8.2%; n=387), and metabolic/endocrine disorders (6.7%; n=314). The proportion of theses categorized as other clinical topics was 5.0% (n=235), and other topics were 22.8% (n=1078) (Table 3).

Table 3. Most frequent thesis topic categories

Topic	n	%
Primary care practice	1223	25.9
Attitudes, perceptions, and approaches	440	9.3
Infectious diseases	387	8.2
Metabolic/endocrine (Diabetes/obesity/thyroid)	314	6.7
Women's/children's and reproductive health	248	5.3
Mental health (depression/anxiety, etc.)	243	5.2
Other clinical topics	235	5.0
Nutrition and physical activity	199	4.2
Knowledge level and awareness	182	3.9
Educational and counseling interventions	166	3.5
Other topics	1078	22.8
Total	**4715**	**100**

Column percentages were used.

When the distribution of the ten most frequent thesis topics was assessed by year, primary care practice showed an overall increasing trend despite a decline in 2018 and 2019 ($p<0.001$, Cramer's V=0,083). Although attitudes, perceptions, and approaches demonstrated an increasing trend over the years, this change was not statistically significant ($p=0.763$). Infectious diseases increased markedly between 2021 and 2023, and the year-by-year distribution was statistically significant ($p<0.001$, Cramer's V=0,140). The year-by-year distributions of metabolic/endocrine disorders, women's/children's and reproductive health, mental health, nutrition and physical activity, and educational and counseling interventions were not statistically significant ($p>0.05$). These grouped as other clinical topics increased noticeably year by year after

2019, and this increase was statistically significant (p<0.001, Cramer's V=0,096). Knowledge level and awareness also showed a significant increase over the years (p=0.041, Cramer's V=0,061) (Figure 3).

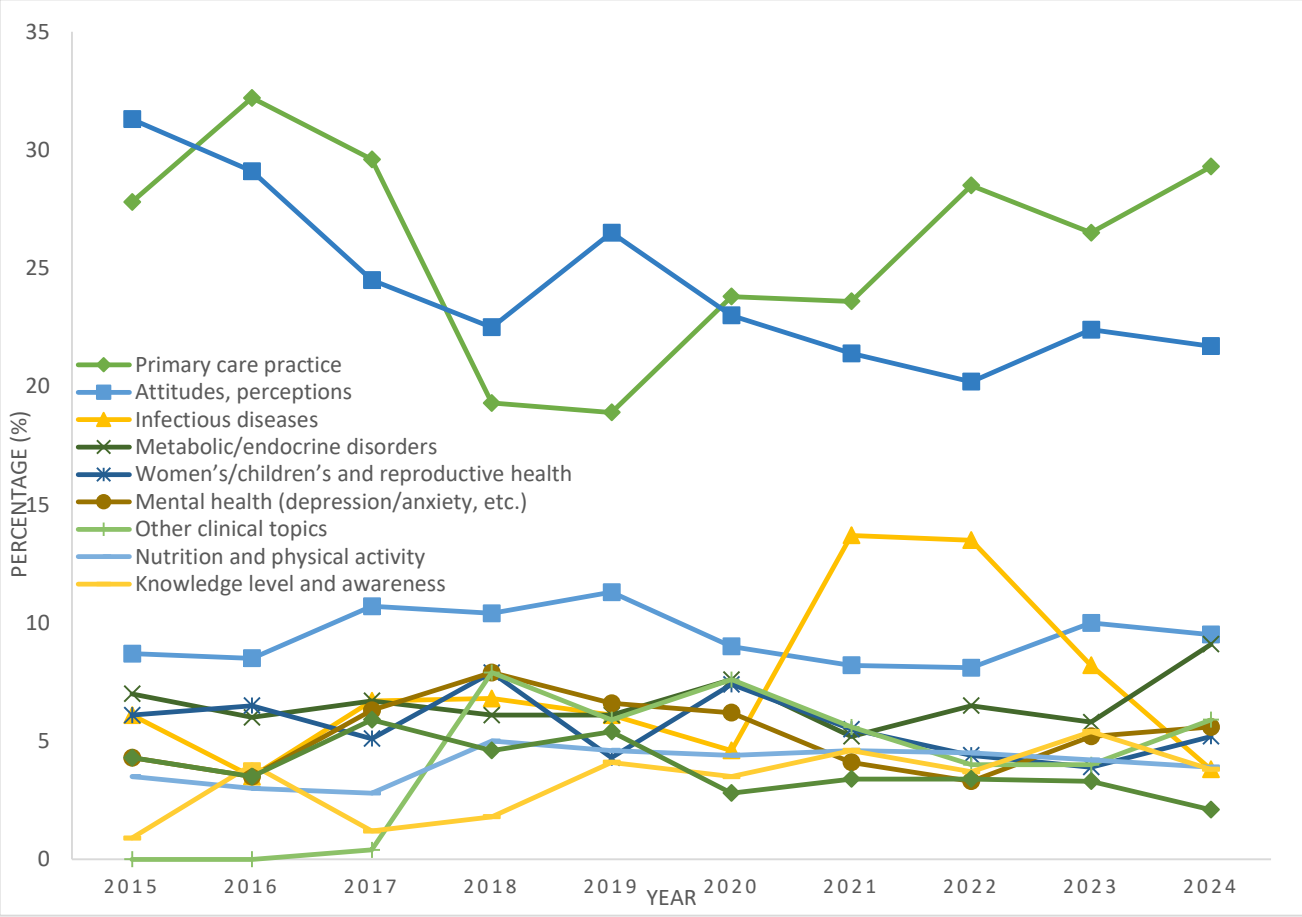
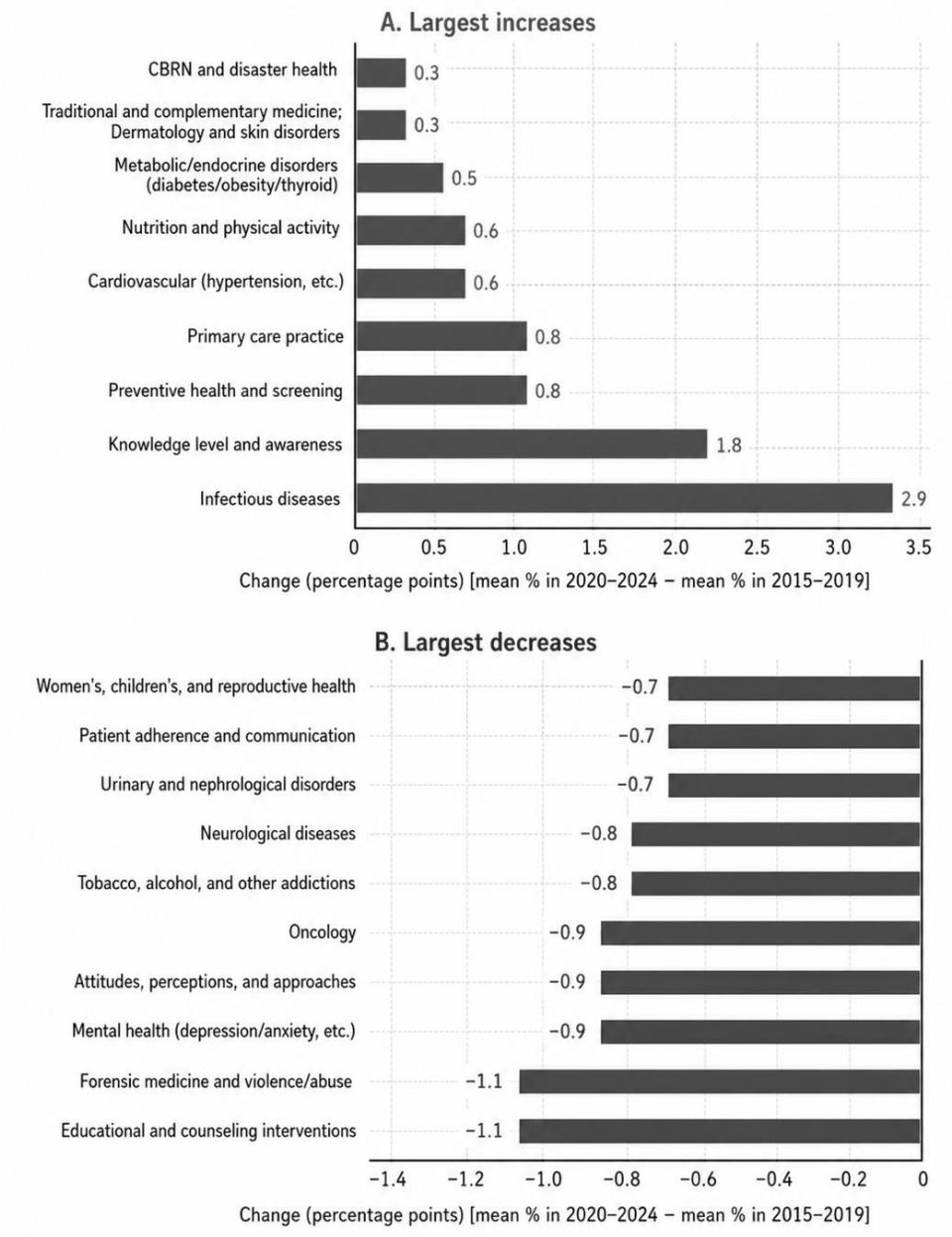


Figure 3. Year-by-Year Distribution of the 10 Most Frequent Thesis Topics

Among the topics showing the greatest increases were infectious diseases and knowledge level and awareness; among those showing the greatest decreases were educational and counseling interventions and forensic medicine and violence/abuse (Figure 4).



**CBRN: Chemical, Biological, Radiological and Nuclear*

Figure 4. Thesis Topics with the Largest Proportional Increases and Decreases (2015-2019 vs 2020-2024)

Binary logistic regression analyses showed that infectious disease-related theses were significantly more likely to be observed in the pandemic/post-pandemic period than in the pre-pandemic period (OR=1.56, 95% CI=1.20–2.02, p<0.001). Similarly, the likelihood of theses categorized as other clinical topics (OR=1.49, 95% CI=1.07–2.07, p=0.017) and knowledge level and awareness (OR=1.63, 95% CI=1.11–2.40, p=0.011) was significantly higher in the pandemic/post-pandemic period. In contrast, educational and counseling interventions were significantly less likely to be observed after 2020 (OR=0.60, 95% CI=0.44–0.83, p=0.002). No statistically significant pre–post pandemic difference was found for primary care practice, attitudes/perceptions/approaches, metabolic/endocrine disorders, women's/children's and reproductive health, mental health, or nutrition and physical activity. The distribution of institution type differed significantly between the pre-pandemic and pandemic/post-pandemic periods. The proportion of UHS-affiliated theses increased from 34.7% (n=430) before the pandemic to 49.8% (n=1,730) in the pandemic/post-pandemic period. UHS-affiliated theses were significantly more likely to be observed in the pandemic/post-pandemic period than in the pre-pandemic period (OR=1.86, 95% CI=1.63–2.13, p<0.001) (Table 4).

Table 4. Crude logistic regression analysis of major thesis topics and institutional affiliation according to the pandemic period

Outcome/Period	Crude OR	95% CI	p-value
Infectious diseases			
Pre-pandemic (2015-2019)	Ref.	-	-
Pandemic/Post-pandemic (2020-2024)	1.56	1.20–2.02	<0.001
Other clinical topics			
Pre-pandemic (2015-2019)	Ref.	-	-
Pandemic/Post-pandemic (2020-2024)	1.49	1.07–2.07	0.017
Knowledge/awareness			
Pre-pandemic (2015-2019)	Ref.	-	-
Pandemic/Post-pandemic (2020-2024)	1.63	1.11–2.40	0.011
Educational/counseling interventions			
Pre-pandemic (2015-2019)	Ref.	-	-
Pandemic/Post-pandemic (2020-2024)	0.60	0.44–0.83	0.002
UHS-affiliated theses			
Pre-pandemic (2015-2019)	Ref.	-	-
Pandemic/Post-pandemic (2020-2024)	1.86	1.63–2.13	<0.001

*: All models represent crude (unadjusted) binary logistic regression analyses

Discussion

This study showed a marked increase in the number of family medicine specialty theses between 2015 and 2024. The increase became more evident after 2020, alongside changes in institutional structure. The proportion of female authors remained consistently high throughout the study period. These findings represent a strengthened continuation of trends reported in previous national bibliometric studies examining the

2005–2015 and 2000–2020 periods; however, the total number of theses and the relative contribution of UHS in our series were substantially higher than in those studies.[5,6] In addition, the distribution of topics was analyzed in detail in the present study: increases were observed in infectious diseases and knowledge level and awareness, whereas a period-specific decline was noted in primary care practice. The increase in thesis production after 2015 may be related not only to growing academic interest in family medicine, but also to Türkiye's specialty training structure, where completing and defending a thesis is a mandatory legal requirement for graduating from medical residency. Consequently, the expansion of residency quotas in family medicine has directly translated into a surge in thesis output.

When distributions were examined by year, the increase between 2015 and 2019 was mainly driven by university-based theses, whereas from 2020 onward, the number of UHS-affiliated theses was comparable to, and in some years higher than, that of other universities. This inflection point is consistent with administrative arrangements that increased the role of clinics/training and research hospitals in family medicine specialty training in Türkiye, and it may also be explained by the entry of institutions that began reporting theses in bulk to YÖK after 2016. The rise in UHS output may additionally be related to the affiliation of training and research hospitals with UHS in subsequent years. While the 2000–2020 analysis noted this shift only briefly for the most recent years, our 10-year series shows that it has become a dominant pattern. A similar institutional concentration has been shown in bibliometric studies of public health theses in Türkiye.[8] Furthermore, it is crucial to recognize that UHS operates as a national umbrella organization encompassing dozens of training and research hospitals, whereas other entries represent single-center universities. This aggregation inherently inflates UHS's total output compared to individual universities, a structural distinction that must be considered when interpreting institutional supervision structures and output volumes. In this respect, although sex distribution did not differ by institution type, the shift in the institutional center of production did not diminish the predominance of female authors; rather, as volume increased, this predominance persisted. This finding aligns with

international studies reporting a steady rise in the number of female physicians and female authorship in primary care/general practice over the past decade, including settings where women have outnumbered men.[9] This enduring female predominance may be attributed to the increasing proportion of female students enrolling in medical schools in Türkiye, alongside the specific appeal of family medicine, which often offers a more favorable work-life balance compared to surgical specialties.

One of the most striking findings of our study is the asymmetry in topic distribution. The high proportion of theses on attitudes, perceptions, and approaches suggests that family medicine research may preferentially focus on areas that are highly accessible, low-cost, and readily studied through surveys. Mengüllüoğlu and Ünlüoğlu reported that a substantial portion of theses in the 2005–2015 period were survey-based and focused on attitudes/knowledge.[6] Our findings indicate that this orientation persisted after 2015 and did not show a meaningful proportional decline.[5] This pattern is also consistent with bibliometric reviews in other medical specialties in Türkiye, which have attributed topic choices to easy access to field data, limited statistical requirements, and feasibility within a short time frame.[10] The predominance of primary care practice, attitudes/perceptions, and knowledge/awareness topics may be explained by the feasibility of these subjects within residency training. Family medicine residents often conduct their theses while continuing clinical duties; therefore, cross-sectional, questionnaire-based, and routinely accessible clinical topics may be preferred. Although several topic categories showed statistically significant year-by-year differences, the corresponding Cramer's V values were small. This indicates that, while temporal changes were statistically detectable in this large dataset, the magnitude of these changes was generally limited. Therefore, both statistical significance and effect size should be considered when interpreting the observed topic trends.

Conversely, the significant year-by-year variation in primary care practice and its decline, particularly in 2018–2021 suggests that topic selection in family medicine theses is shaped not only by clinical preferences but also by contemporary needs in field practice.

It has been reported in many countries that the COVID-19 pandemic beginning in 2020 narrowed the topic spectrum of primary care studies and redirected research toward specific headings such as infection, immunization, re-organization of service delivery, and remote follow-up.[11,12] In our study, the sharp increase in infectious disease theses in 2021 and the statistically significant year-by-year difference likely represent the local reflection of this global trend. Moreover, international literature has shown that redesign of primary care services, disruptions in chronic disease follow-up, and changes in immunization programs during the pandemic shaped primary care research.[13,14] The increase in infectious disease topics in our dataset is consistent with this literature. Although infectious diseases increased, the primary driver appears to have been the classification of many pandemic-period theses as infectious disease studies in the context of COVID-19. The increase in infectious disease-related theses during the pandemic/post-pandemic period probably reflects the central role of family medicine in COVID-19 follow-up, vaccination counseling, risk communication, and continuity of primary care services.

The absence of a statistically significant year-by-year difference in mental health topics can be interpreted in two ways. First, although international publications on mental health increased substantially during the pandemic period, this increase was not reflected in our thesis dataset to the same extent.[15,16] This may be because family medicine thesis topics in mental health shifted internally toward pandemic-related mental health issues. Second, because the duration and content of mental health rotations in family medicine training vary across institutions, fluctuations in thesis topic selection are expected. In contrast, the lack of significant variation over time in chronic disease–related topics (metabolic/endocrine disorders), nutrition and physical activity, and women's/children's and reproductive health suggests that these core domains retained a stable place in family medicine training, generating a consistent volume of theses each year. This stability is compatible with the concept that patient-centered care in primary care has increasingly integrated chronic disease management over 2010–2024.[16]

Our findings also offer an indirect explanation for the frequently reported problem of low conversion rates of theses into journal publications. International studies examining the publication of medical specialty theses suggest that theses that are survey-based, cross-sectional, and service-oriented may face greater difficulty being accepted in peer-reviewed journals, whereas theses with methodological diversity and mixed quantitative–qualitative designs are more likely to be published.[10,17] Although our study did not directly assess the methodological quality or the actual thesis-to-publication conversion rates, the thematic distribution suggests a high reliance on descriptive, single-center surveys. Such designs typically face greater challenges in being accepted by high-impact peer-reviewed journals compared to mixed-method or multi-center analytical studies. While such studies can describe everyday problems of primary care effectively, they are often single-center, limited in sample size, and strongly influenced by local service conditions.

These findings also have academic and educational implications for family medicine residency training in Türkiye. The concentration of theses in feasible, cross-sectional, and survey-based topics suggests that residents may need stronger methodological support during thesis planning, including training in research design, biostatistics, data management, and scientific writing. Residency programs may benefit from structured thesis monitoring systems, multicenter research networks, and mentorship models that encourage clinically relevant but methodologically stronger studies. Improving thesis quality may also increase the likelihood of thesis-to-publication conversion, thereby transforming mandatory residency theses from a graduation requirement into a more visible contribution to family medicine literature and primary care policy.

A key strength of this study is that, unlike previous Turkish bibliometric thesis and article examples, it simultaneously demonstrates the post-2015 increase in thesis volume, the UHS contribution, and the pandemic-related topic shift.

This study has several limitations. First, author sex was inferred from names recorded in the YÖKTEZ metadata rather than self-reported information; therefore, name-based

classification may be subject to misclassification and does not allow assessment of gender identity. Second, topic classification was based mainly on thesis titles and a predefined thematic framework. Although coding rules were standardized, title-based classification may not fully reflect the detailed content, methodology, or secondary aims of each thesis, especially for studies addressing more than one topic. Third, the study was limited to accessible records in the YÖK National Thesis Center database. Restricted-access theses, incomplete metadata, duplicate entries, delayed registrations, or inconsistencies in department and institution names may have affected the bibliometric counts. Finally, thesis-to-article conversion and citation performance were beyond the scope of this study. Therefore, the findings should be interpreted as reflecting accessible YÖKTEZ-based thesis records rather than all family medicine residency research activity in Türkiye. Additionally, because separate logistic regression models were utilized to explore temporal associations across multiple topics, the potential for Type I error due to multiple comparisons should be acknowledged. Furthermore, because of the thematic classification approach used in the study, some theses may potentially overlap across closely related categories. We acknowledge that this possibility may have resulted in a degree of subjectivity or misclassification.

In conclusion, family medicine specialty thesis production in Türkiye increased substantially between 2015 and 2024, with a growing contribution of UHS and topic shifts partly influenced by the pandemic. These findings indicate that residency theses reflect changes in training capacity, institutional structure, and primary care priorities. Strengthening methodological support, thesis supervision, biostatistics training, and scientific writing mentorship may improve thesis quality and increase thesis-to-publication conversion. At the policy level, national monitoring of thesis topics may help reduce repetition, encourage multicenter practice-based research, and align thesis planning with primary care needs. Future studies should evaluate thesis quality, publication conversion rates, citation impact, and methodological characteristics.

Ethical Considerations: Ethics approval was not required because this study was based exclusively on publicly available thesis metadata obtained from the YÖKTEZ database. No human participants were recruited, no intervention was performed, and no personal, sensitive, or identifiable private data were collected. The analyzed information consisted only of publicly accessible bibliometric variables such as thesis year, institution, author name as recorded in the database, and thesis title/topic. Therefore, the study did not involve human subject research requiring ethics committee approval.

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

References

1. Uğur F, Dağcıoğlu BF. Review of specialty theses assessing family physicians' knowledge levels. Turkish Journal of Family Medicine & Primary Care. 2024;18(1):101-8.
2. Starfield B. Global health, equity, and primary care. The Journal of the American Board of Family Medicine. 2007;20(6):511–3.
3. Marušic A, Marušic M. Teaching students how to read and write science: a mandatory course on scientific research and communication in medicine. Academic Medicine. 2003;78(12):1235–9.
4. Regulation on Specialty Training in Medicine and Dentistry (Türkiye). Resmî Gazete; 3 September 2022. Sayı:31942 (Original Turkish title: Tıpta ve Diş Hekimliğinde Uzmanlık Eğitimi Yönetmeliği.)
5. Karakullukçu A, Ardıç C. Bibliometric analysis of medical speciality dissertation studies in family medicine departments and clinics between 2000–2020. Ankara Medical Journal. 2023;23(3).
6. Mengüllüoğlu NÖ, Ünlüoğlu İ. Evaluation of family medicine specialty theses between 2005–2015. Ankara Medical Journal. 2017;17(4). (Original Turkish title: 2005-2015 Yılları Arasındaki Aile Hekimliği Uzmanlık Tezlerinin Değerlendirilmesi.)
7. Celik NM, Kuçlu T. Bibliometric analysis of medical specialty theses on health literacy in Türkiye. Turkish Health Literacy Journal. 2022;3(2):38–95.
8. Öztürk ENY, Yücel M, Uyar M. Bibliometric analysis of theses made in public health and registered in the National Thesis Center of the Council of Higher Education in Türkiye. Hipokrat Tıp Dergisi. 2022;2(3):15–22.
9. Giannos P, Katsikas Triantafyllidis K, Paraskevaidi M, Kyrgiou M, Kechagias KS. Female dynamics in authorship of scientific publications in the Public Library of Science: a 10-year bibliometric analysis of biomedical research. European Journal of Investigation in Health, Psychology and Education. 2023;13(2):228–37.
10. Rador MC, Çiçek Ç. From specialization theses to peer-reviewed articles: the pathway in plastic surgery. Turkish Journal of Plastic Surgery. 2025;33(4):181–7.
11. Michael GC, Aliyu I. Impact of COVID-19 pandemic on primary care research in a resource-limited setting: a commentary. The Pan African Medical Journal. 2020;35(Suppl 2):15.
12. Khalil-Khan A, Khan MA. The impact of COVID-19 on primary care: a scoping review. Cureus. 2023;15(1).
13. Damar M, Pinto AD, Erenay FS, Aydin O. Impact of COVID-19 on primary health care research trends and suggestions for better service approaches via blockchain-based applications. Blockchain in Healthcare Today. 2025;8(1).

14. Michalaki F, Triantafillopoulou KM, Pagkozidis I, Tirodimos I, Dardavesis T, Tsimtsiou Z. The impact of COVID-19 pandemic on Primary Health Care through 'health providers' eyes': a qualitative study of focus groups and individual interviews in Greece. European Journal of General Practice. 2024;30(1):2382218.
15. Deml MJ, Minnema J, Dubois J, Senn O, Streit S, Rachamin Y, et al. The impact of the COVID-19 pandemic on the continuity of care for at-risk patients in Swiss primary care settings: a mixed-methods study. Social Science & Medicine. 2022;298:114858.
16. Tong L-K, Tam H-L, Mao A-M, editors. A bibliometric review of person-centered care research 2010–2024. Healthcare. 2025;MDPI.
17. Jakab M, Kittl E, Kiesslich T. How many authors are (too) many? A retrospective, descriptive analysis of authorship in biomedical publications. Scientometrics. 2024;129(3):1299–328.

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):445-466 // 10.5505/amj.2026.53911

DEVELOPMENT AND VALIDATION OF THE DIABETES MELLITUS KNOWLEDGE SCALE: A VALIDITY AND RELIABILITY STUDY

Gülsüm Öztürk Emiral[1], Abide Aksungur[2], Hüsna Sarıca Çevik[3], Hamit Harun Bağcı[4]

[1]Department of Public Health, Ankara Yıldırım Beyazıt University, Ankara, Türkiye
[2]Haymana Vocational School, Ankara University, Ankara, Türkiye
[3]Department of Family Medicine, Ankara University Medical School, Ankara, Türkiye
[4]General Directorate of Public Health, Republic of Türkiye Ministry of Health, Ankara, Türkiye

Correspondence:
Gülsüm Öztürk Emiral (e-mail: dr.gulsum.ozturk@gmail.com)

Submitted: 19.02.2026 // Accepted: 27.07.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: Diabetes Mellitus (DM) is a common non-communicable disease globally. Since children spend a significant amount of time at school, collaboration with school staff is essential for managing both emergencies and long-term complications related to DM. Therefore, assessing the knowledge level of school staff about DM is of great importance. This study aimed to develop a Diabetes Mellitus Knowledge Scale (DM-KS) and to conduct its validity and reliability analysis among teachers.

Materials and Methods: This methodological study was carried out with 379 teachers working under the Altındağ District National Education Directorate during the 2022–2023 academic year. The DM-KS consists of 23 items. Validity was assessed through content validity, item analysis, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and known-groups comparison.

Results: The DM-KS comprises two subscales: "DM General Information and Symptoms" (11 items) and "DM Treatment and Complication Management" (12 items). Item difficulty indices ranged between 0.37 and 0.84, and discrimination coefficients ranged between 0.40 and 0.91. According to EFA, four items were excluded due to overlapping factors, and the factor loadings of the remaining items ranged from 0.39 to 0.76. The explained variances of subscales 1 and 2 were 38.2% and 41.4%, respectively. Cronbach's alpha values were 0.84-0.87. Participants who had received training scored significantly higher in both subscales.

Conclusions: The DM-KS demonstrated acceptable reliability and preliminary evidence of validity for assessing diabetes-related knowledge among teachers. Further validation in independent and diverse samples is warranted.

Keywords: Type 1 diabetes mellitus; diabetes knowledge; diabetes knowledge scale; school health; psychometrics; validity and reliability.

Introduction

Type 1 diabetes is one of the most prevalent long-term diseases of childhood globally. It has major health consequences for individuals and society.[1,2] The IDF Diabetes Atlas (2021) estimates that DM affects 10.5% of adults aged 20–79 years (536.6 million people) and was associated with approximately 6.7 million deaths worldwide in 2021.[3] According to World Health Organization (WHO) data, there are approximately 9 million individuals diagnosed with Type 1 DM, the majority of whom are in high-income countries.[4] In Türkiye, type 1 DM represents a major pediatric burden: approximately 1,700–2,000 children are newly diagnosed each year, and about 30,000 children under 18 live with the condition, most of them school-age.[5,6] Teachers play a critical role in the daily management and safety of children with type 1 diabetes mellitus (T1 DM) in school settings. Inadequate knowledge among school personnel may lead to delayed recognition of hypoglycemia or hyperglycemia, inappropriate emergency responses, and increased risk of acute complications, which can negatively affect both health outcomes and academic performance of students with diabetes. International organizations emphasize that improving diabetes awareness and preparedness in schools is essential for ensuring safe educational environments and preventing diabetes-related emergencies among children. Therefore, assessing teachers' knowledge levels is an important step in planning effective educational interventions and school health policies.[4-6] At school, children need prompt recognition of hypoglycemia and ketoacidosis, meal and activity planning, assistance with glucose monitoring and treatment, and coordinated support from teachers, staff, and families. Therefore, this study aimed to develop and validate a reliable and valid scale to assess teachers' knowledge regarding type 1 diabetes mellitus, including recognition and management of hypoglycemia and hyperglycemia and emergency management in school settings, and to introduce the Diabetes Mellitus Knowledge Scale (DM-KS).

Materials and Methods

This methodological study was conducted with primary, middle, and high school teachers serving in Altındağ District National Education Directorate during the 2022-2023 academic year. This study was carried out within the scope of the project titled: "Evaluation of the Knowledge Level of School Teachers about Diabetes Mellitus", and SBÜ Dr. Abdurrahman Yurtaslan Ankara Onkoloji Research and Hospital Ethics Committee approval (2022-10/172) was received to conduct this study.

Development of Diabetes Mellitus Knowledge Scale

An extensive literature review was conducted to develop the DM-KS. A widely accepted and used DM-KS among teachers could not be found. Therefore, it was decided to develop the DM-KS. While developing the DM-KS, a question pool consisting of 40 items was prepared. In line with expert opinions, the questions were reduced to 27 items and grouped under two subheadings: “Diabetes Mellitus General Information and Symptoms (Subscale 1)” and “Diabetes Mellitus Treatment and Complication Management (Subscale 2)”.

In both subscales of DM-KS, correct answers were evaluated as "2", “I don't know” answers as "1", and incorrect-blank answers were evaluated as "0" points. No correction formula was applied for chance-based success.

Whether DM-KS measures the desired structure was evaluated with content validity. For content validity, the opinions of experts, including public health specialists, family medicine specialists, linguists, internal medicine specialists, and endocrine specialists, were consulted. The questions were examined by experts, and each question was evaluated as 'necessary-adequate', 'necessary but insufficient', and 'unnecessary'. The question root and options were changed. Then, expert opinions were transformed into statistically interpretable quantifiers. The content validity index (CVI) was calculated for each question. In both subgroups, the CVI was above 0.80 (0.87-0.90).[7]

Study group, procedure, and questionnaire form

Our study was conducted with teachers who serve under the Altındağ District National Education Directorate. After obtaining the necessary permissions and informing the school officials, the schools were visited during working hours. The subject and purpose of the study were explained. Those who agreed to participate in the study were asked to complete the questionnaire.

The questionnaire form consists of 2 parts. The first part includes the sociodemographic characteristics of the teachers, and the second part comprises the DM-KS items (27 items).

According to the records of the Ministry of National Education, 4,418 teachers were employed in the Altındağ district at the time of the study. The sample size recommended for developing a new assessment tool is generally 5–10 times the number of scale items.8 Since the initial version of the DM-KS consisted of 27 items, the minimum required sample size was calculated as 270 participants. To ensure representation across educational levels, schools were randomly selected from among primary, secondary, and high schools in the district. Reserve schools were also identified in advance to compensate for potential non-response. Participants were recruited from the selected schools, and all eligible teachers who were present during the data collection period and consented to participate were included in the study. Participation was voluntary, and no incentives were offered. The study was completed with 379 teachers, exceeding the recommended sample size for scale development studies.

Item Analysis

Item difficulty and difficulty index should be calculated when developing assessment tools, such as those measuring ability and success that require performance. The item difficulty index is used to distinguish how difficult or easy an item is, and the item discrimination coefficient is used to distinguish between those who know and those who do not. In our study, item difficulty (Pj) and discrimination coefficients (rjx) were

calculated with the formula P= (H+L/N)*100 and d= (H-L/N/2) formulas (ABBID). Another method used to determine item validity is to compare the item scores of the lower and upper 27% groups. It is based on the assumption that people in the upper 27% group possess the characteristic that the assessment tool aims to measure positively, while those in the lower 27% group exhibit a negative effect. This assumption is also expected to emerge among the items. The scores of the lower- and upper-27 % groups from the evaluation tool are analysed using a significance test to determine the difference between the two averages in independent samples. It is desirable that there is a statistical difference between the two averages. The statistical difference between the lower and upper 27% groups reveals evidence of the validity and distinctiveness of the items in the assessment tool.[7]

Construct Validity

The Kaiser-Meyer-Olkin (KMO) value was examined to decide whether the sample size was suitable for factor analysis. The KMO value ranges from 0 to 1, and it is required to be close to 1 for factor analysis to be effective. In the Exploratory Factor Analysis (EFA) to be performed, a KMO coefficient of 0.60 and above is considered sufficient, but it is desired to be above 0.80. While performing EFA, not all items were analysed at the same time, and it was checked whether they were collected under the determined dimensions.[8]

One method used to determine the construct validity of the developed assessment tool is ""known groups comparison". The intergroup differentiation method assumes that the groups are different in terms of the area to be measured. In our study, teachers who received DM training were expected to achieve higher scores on DM-KS and were compared with those who did not receive training using the Mann-Whitney U test.[7]

One frequently used method in determining construct validity is Confirmatory Factor Analysis (CFA). CFA is a method that allows checking the factor structures revealed in EFA. While the relationships between the items in the evaluation tool are examined with EFA, the relationships between the factors are evaluated with CFA. CFA is a modelling

method used to evaluate the latent structure of a measurement tool and to test the fit between observed variables and latent variables. In this study, the factor structure of the DM-KS was examined using CFA with the LISREL 8.51 software program.

Internal Consistency

The Cronbach's alpha coefficient was calculated to test the reliability of DM-KS and to determine internal consistency.

Scoring DM-KS - Statistical Analysis

Each correct answer given to the DM-KS items was scored as "2", incorrectly marked as "0", and those marked as "I don't know" were scored as "1". As the score obtained from each domain increased, the level of knowledge increased.

Data were evaluated using IBM SPSS Statistics for Windows, Version 20.0 (Armonk, NY: IBM Corp.). Sociodemographic characteristics were presented with descriptive statistical data (mean, frequency, standard deviation, etc.). The conformity of distribution was evaluated with the Shapiro-Wilk test. Mann-Whitney U, Kruskal-Wallis test, and Spearman correlation analysis were used for statistical analysis.

Ethical Considerations

This study was carried out within the scope of the project titled: 'Evaluation of the Knowledge Level of School Teachers about Diabetes Mellitus', and University of Health Sciences Dr. Abdurrahman Yurtaslan Ankara Oncology Training & Research Hospital Ethics Committee approval (2022-10/172) was received to conduct this study.

Results

DM-KS Validity-Reliability Analysis

The DM-KS item difficulty indexes ranged from 0.37 to 0.84, and the discrimination coefficients ranged from 0.40 to 0.91. According to the item analysis results, no items were extracted from the developed assessment tool.

As a result of the EFA analysis performed to test the construct validity, four items were excluded from the assessment tool because they had a factor load of more than 0.40 in each of the two subscales (overlapping factor) ("item 10: The main effect of insulin hormone is to raise blood sugar.", "item 11: Insulin therapy is in the form of oral tablets.", "item 17: In case of low blood sugar, if the child is conscious, he/she can be sent home.", “Item 21: It is not appropriate for a student with controlled Diabetes Mellitus to attend classes that require exercise, such as physical education.”). The item analysis and EFA results of DM-KS are given in Table 1.

Table 1. Item analysis results and factor loads of DM-KS.

DM-KS items	Pj	rjx	Factor Load
Diabetes Mellitus General Information and Symptoms			
KMO:0.870			
Bartlett's Test of Sphericity:1117.482; 0.000			
Explained Variance:38.20			
1- Diabetes mellitus is a disease with high blood sugar that occurs when the pancreas is unable to produce enough insulin or the body is unable to use insulin effectively.	0.78	0.51	0.61
2- The blood sugar level of an individual without diabetes should not exceed 100 mg/dL in fasting and 140 mg/dL in satiety (2 hours after starting the meal).	0.47	0.70	0.59
3- The two most common types of diabetes mellitus are Type 1 and Type 2 diabetes.	0.53	0.83	0.69
4- The disease, characterised by an increase in blood sugar due to absolute insulin deficiency, is called “Type 1 diabetes mellitus”.	0.37	0.78	0.63
5- Type 1 is the most common type of diabetes mellitus in childhood.	0.40	0.78	0.64
6- Diabetes mellitus is one of the leading chronic diseases worldwide.	0.73	0.53	0.56
7- Findings such as constant thirst/dry mouth, drinking more water than usual, and frequent and copious urination (polyuria) are among the most common findings of diabetes mellitus.	0.84	0.42	0.60
13- A blood sugar level lower than 70 mg/dL is considered hypoglycaemia (low blood sugar).	0.48	0.91	0.70
14- Symptoms such as hand and foot tremors, pale colour, and cold sweating may indicate low blood sugar.	0.75	0.55	0.60
18- Symptoms such as a feeling of thirst, acetone odour in the mouth, and sleepiness may indicate high blood sugar.	0.55	0.80	0.65
22- Diabetes is a lifelong disease.	0.63	0.55	0.50
Diabetes Mellitus Treatment and Complication Management			
KMO:0.893			
Bartlett's Test of Sphericity:1549.145; ≤0.001			
Explained Variance:41.42			

8- Insulin therapy is used in the treatment of type 1 diabetes mellitus.	0.46	0.70	0.51
9- The pancreas is the main organ in which the insulin hormone is secreted.	0.63	0.75	0.65
12- Insulins are drugs subject to the cold chain.	0.39	0.50	0.39
15- In case of low blood sugar, sugar cubes/fruit juice should be given if the child is conscious.	0.65	0.72	0.65
16- In case of low blood sugar, emergency health services should be notified if the child is unconscious.	0.84	0.40	0.56
19- Exercise programs are part of the treatment of diabetes mellitus.	0.60	0.73	0.63
20- In situations that cause anxiety, such as exams, the blood sugar balance may deteriorate.	0.70	0.66	0.68
23- Exercise programs are effective in lowering blood sugar.	0.60	0.80	0.71
24- Failure to control blood sugar creates long-term health problems.	0.83	0.52	0.68
25- If blood sugar is not kept under control in diabetes mellitus, eye damage (blindness, etc.) may occur.	0.65	0.80	0.74
26- If blood sugar is not kept under control in diabetes mellitus, severe kidney disease may develop that may require dialysis and/or kidney transplantation.	0.70	0.70	0.76
27- If blood sugar is not kept under control in diabetes mellitus, diseases such as heart attack and stroke, which may result in death, may develop due to damage to the heart and brain vessels.	0.53	0.81	0.69

Total Explained Variance: 40%, Pj: item difficulty index; rjx: item discrimination coefficients.

In the reliability analysis, the Cronbach's alpha coefficient for subscale-1 was 0.84, and the item-total correlation coefficient ranged between 0.394 and 0.604; the Cronbach's alpha coefficient for subscale-2 was 0.87, and the item-total correlation coefficient ranged between 0.330 and 0.674. Cronbach's alpha coefficient did not increase when any item was removed in both subscales. The reliability analyses of the study are given in Table 2.

Table 2. DM-KS Reliability Analysis Results

DM-KS items	Corrected item-total correlation	Cronbach's alpha if item deleted
Diabetes Mellitus General Information and Symptoms		
***Cronbach's alpha coefficient*: 0.836**		
1- Diabetes mellitus is a disease with high blood sugar that occurs when the pancreas is unable to produce enough insulin or the body is unable to use insulin effectively.	0.507	0.823
2- The blood sugar level of an individual without diabetes should not exceed 100 mg/dL in fasting and 140 mg/dL in satiety (2 hours after starting the meal).	0.481	0.825
3- The two most common types of diabetes mellitus are Type 1 and Type 2 diabetes.	0.592	0.815
4- The disease, characterised by an increase in blood sugar due to absolute insulin deficiency, is called "Type 1 diabetes mellitus".	0.529	0.820
5- Type 1 is the most common type of diabetes mellitus in childhood.	0.540	0.819
6- Diabetes mellitus is one of the leading chronic diseases worldwide.	0.452	0.827
7- Findings such as constant thirst/dry mouth, drinking more water than usual, and frequent and copious urination (polyuria) are among the most common findings of diabetes mellitus.	0.496	0.824
13- A blood sugar level lower than 70 mg/dL is considered hypoglycaemia (low blood sugar).	0.604	0.813
14- Symptoms such as hand and foot tremors, pale colour, and cold sweating may indicate low blood sugar.	0.491	0.824
18- Symptoms such as a feeling of thirst, acetone odour in the mouth, and sleepiness may indicate high blood sugar.	0.541	0.819
22- Diabetes is a lifelong disease.	0.394	0.832
Diabetes Mellitus Treatment and Complication Management		
***Cronbach's alpha coefficient*: 0.865**		
8- Insulin therapy is used in the treatment of type 1 diabetes mellitus.	0.434	0.863
9- The pancreas is the main organ in which the insulin hormone is secreted.	0.555	0.854
12- Insulins are drugs subject to the cold chain.	0.330	0.865

15- In case of low blood sugar, sugar cubes/fruit juice should be given if the child is conscious.	0.566	0.854
16- In case of low blood sugar, emergency health services should be notified if the child is unconscious.	0.470	0.860
19- Exercise programs are part of the treatment of diabetes mellitus.	0.542	0.855
20- In situations that cause anxiety, such as exams, the blood sugar balance may deteriorate.	0.586	0.852
23- Exercise programs are effective in lowering blood sugar.	0.628	0.849
24- Failure to control blood sugar creates long-term health problems.	0.574	0.854
25- If blood sugar is not kept under control in diabetes mellitus, eye damage (blindness, etc.) may occur.	0.642	0.848
26- If blood sugar is not kept under control in diabetes mellitus, severe kidney disease may develop that may require dialysis and/or kidney transplantation.	0.674	0.847
27- If blood sugar is not kept under control in diabetes mellitus, diseases such as heart attack and stroke, which may result in death, may develop due to damage to the heart and brain vessels.	0.599	0.851
Total scale Cronbach's alpha coefficient:0.891		

As another analysis method for DM-KS, the scores of the upper and lower groups (27%) were compared, and it was determined that the scores of the upper group were higher in all subscales of DM-KS (for each subscale, $p<0.05$). The comparison of the item scores of the lower and upper 27% groups is given in Table 3.

Table 3. Comparison of the item scores of the lower and upper 27% groups.

DM-KS		Mean (SD)	Median (Q1-Q3)	Statistical analysis z; p
Subdomain-1	Group of upper 27%	20.6 (1.5)	20.5 (20.0-22.0)	12.398; <0.001
	Group of lower 27%	12.4 (2.5)	12.0 (11.0-14.0)	
Subdomain-2	Group of upper 27%	22.4 (1.7)	23.0 (22.0-24.0)	12.284; <0.001
	Group of lower 27%	13.2 (3.9)	14.0 (12.0-16.0)	

Data are presented as mean (SD) and median (interquartile range). Mann–Whitney U test was used.

For intergroup differentiation, the DM-KS scores of individuals who had received DM training were compared with those who had not received training. As predicted, those who received training beforehand had higher scores in both subscales. Comparisons between groups are given in Table 4.

Table 4. Comparisons between groups.

DM-KS		Mean (SD)	Median (Q1-Q3)	Statistical analysis z; p
Subdomain 1	Those who have not received training on DM	16.5 (3.7)	17.0 (14.0-19.5)	2.521; 0.012
	Those who received training on DM	17.7 (3.6)	19.0 (16.0-20.0)	
Subdomain 2	Those who have not received training on DM	18.2 (4.6)	19.0 (16.0-22.0)	3.189; 0.001
	Those who received training on DM	20.1 (3.7)	21.0 (19.0-23.0)	

Mann–Whitney U test was used. *Q1–Q3: interquartile range (25th–75th percentiles).*

In the CFA conducted for DM-KS, although several goodness-of-fit indices did not meet commonly recommended thresholds, the t-values and factor loadings in the measurement model were within acceptable limits. The measurement model (T values) obtained from the CFA is presented in Figure 1, and the fit indices for this model are listed in Table 5.

Table 5. Fit indices obtained from CFA.

Fit indices	**Fit values of the measurement model obtained for the DM-KS**
χ^2 /df	6.50
RMSEA	0.12
SRMR	0.08

CFI	0.89
GFI	0.74
AGFI	0.69

df: degrees of freedom; RMSEA: root mean square error of approximation; SRMR: standardized root mean square residual; CFI: comparative fit index; GFI: goodness of fit index; AGFI: adjusted goodness of fit index.

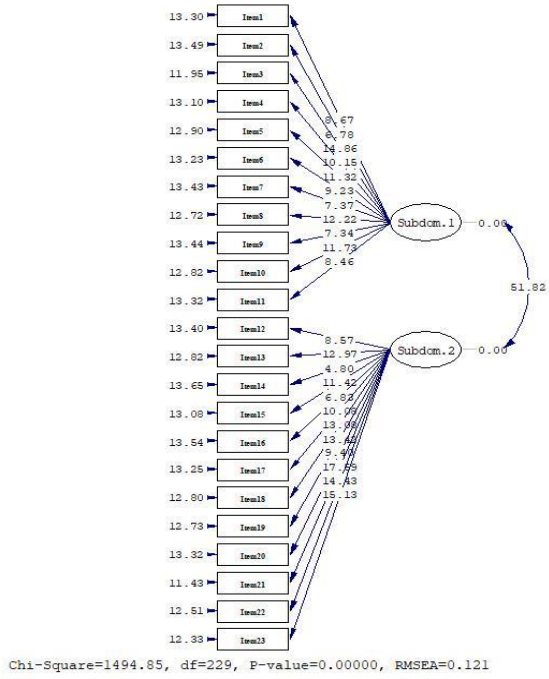


Figure 1. The measurement model obtained from the CFA

Characteristics of Participants

281 (74.1%) of the participants in the study were female. Their ages ranged from 22 to 62 years, with a mean (SD) of 39.7 (9.3) years. DM-KS scores were not significantly associated with most sociodemographic characteristics ($p > 0.05$). However, teachers with a chronic disease demonstrated significantly higher DM-KS scores compared with those without a chronic disease ($p = 0.008$) (Table 6).

Table 6. Comparison of sociodemographic characteristics of participants and DM-KS.

		n (%)	DM-KS	Statistical Analysis
			Median (Q1-Q3)	Kw;z: p
Age group	<35	130 (34.3)	16.0 (12.0-20.0)	5.229; 0.073
	35-44	126 (33.2)	17.0 (10.0-22.0)	
	≥45	123 (32.5)	19.0 (13.0-22.0)	
Sex	Female	281 (74.1)	17.0 (13.0-22.0)	1.677; 0.094
	Male	98 (25.9)	16.0 (10.0-21.0)	
School type	Primary school	97 (25.6)	17.0 (13.0-21.0)	2.263; 0.323
	Middle school	126 (33.2)	16.5 (11.0-21.0)	
	High school	156 (41.2)	17.5 (11.5-22.0)	
Working position	Director/Manager	27 (7.1)	20.0 (12.0-23.0)	1.114; 0.265
	Teacher	352 (92.9)	17.0 (12.0-21.0)	
Perceived income level	Poor	45 (11.9)	16.0 (10.0-21.0)	2.767; 0.251
	Middle	250 (66.0)	17.0 (12.0-21.0)	
	Good	84 (22.2)	18.0 (13.5-22.0)	
Having a child	Yes	271 (71.5)	17.0 (11.0-21.0)	0.532; 0.594
	No	108 (28.5)	17.0 (13.0-22.0)	
Having a chronic disease	Yes	77 (20.3)	19.0 (14.0-23.0)	***2.637; 0.008***
	No	302 (79.7)	16.5 (11.0-21.0)	

Kruskal Wallis/ *Mann–Whitney U test was used.* Q1–Q3: interquartile range (25th–75th percentiles).

Discussion

DM has become a major public health problem in Türkiye. According to the IDF, approximately 9 million individuals were living with DM in Türkiye in 2021, and an estimated 30,000 children under 18 years have type 1 DM.[5,6] Because children spend a substantial proportion of their time at school, early recognition of DM symptoms, awareness of emergency signs in children with DM, and prevention of acute and long-term complications are critically dependent on school personnel.[9] Consequently, teachers must possess adequate DM knowledge, and our study therefore aimed to develop a Diabetes Mellitus Knowledge Scale (DM-KS) and to conduct its validity and reliability analysis among teachers.

In developing a new assessment tool, items should not be excessively easy or difficult and should discriminate between knowledgeable and non-knowledgeable individuals. For this purpose, an item discrimination coefficient of ≥0.20 is generally recommended.[10] In our study, DM-KS item difficulty indices ranged from 0.37 to 0.84, and discrimination coefficients from 0.40 to 0.91, indicating satisfactory psychometric performance. Item discrimination was further confirmed by comparing the lower and upper 27% groups, and all items of the DM-KS were discriminative. In the Diabetes Knowledge Questionnaire validation study by Hsieh et al., item difficulties ranged from 0.285 to 0.993 and discriminations from 0.199 to 0.578, while Garcia et al. reported difficulty indices between 0.14 and 0.96 and discrimination values between 0.27 and 0.37.[11,12] Overall, the item characteristics of DM-KS were comparable or superior to those reported in previous instruments.

Exploratory factor analysis (EFA) was used to examine the construct validity of DM-KS. Bartlett's test of sphericity indicated that the item correlation matrix was suitable for factor analysis, and the Kaiser–Meyer–Olkin measure showed adequate sampling. In line with recommendations in the social sciences, where an explained variance of 40–60% is considered acceptable,[7] the two-factor structure of DM-KS accounted for 40% of the total variance and was considered satisfactory.

Factor loadings reflect the weights of items on their respective latent constructs and are typically expected to be ≥0.30.[7] In DM-KS, all item loadings exceeded 0.30, indicating that each item contributed adequately to its factor. Consistent with our approach, Hsieh et al. removed items with factor loadings <0.30 in their Taiwanese validation study, and Yavuz and Erol excluded items with loadings <0.20, with the lowest retained loading being 0.30.[12,13]

Construct validity was also supported by known-groups (intergroup) differentiation. As hypothesised, teachers who had previously received DM training achieved significantly higher DM-KS scores than those without such training, demonstrating the scale's ability to discriminate between groups with differing knowledge levels. Similarly, Bukhsh et al. reported higher diabetes knowledge scores among individuals with good glycaemic control compared with those with poor control, and Al-Qazaz et al. found parallel associations between knowledge and self-management indicators.[14,15] These convergent findings further support the construct validity of DM-KS.

Internal consistency is a key component of reliability. A Cronbach's alpha coefficient of ≥0.70 is generally regarded as acceptable.[7] In our study, Cronbach's alpha values for the overall DM-KS and for both subscales exceeded 0.80, indicating high internal consistency. Previous diabetes knowledge scales developed or adapted in different languages have reported alpha values of ≥0.70, and our results therefore fall within, or above, the range reported in the literature.[11,16] Differences in scoring formats, sample characteristics, and scale length across studies may partly account for variability in alpha values, yet the overall pattern consistently supports adequate reliability.

Confirmatory factor analysis (CFA) was conducted to test the two-factor structure identified in EFA. CFA evaluates the fit between observed variables and latent constructs by multiple fit indices; however, there is no universal consensus on optimal cut-off values, and indices are influenced by factors such as sample size, model complexity, and data distribution[17]. In our analysis, all item t-values exceeded 1.96 and all item–factor correlations were statistically significant, supporting the specified factor structure.

Although some goodness-of-fit indices did not meet commonly recommended thresholds, the findings provided preliminary support for the proposed two-factor structure when considered together with the factor loadings, theoretical coherence, and other psychometric findings. It is well recognized that model fit indices may be affected by sample characteristics, model complexity, and item distributions, and strict adherence to cut-off values alone may lead to incorrect conclusions regarding model adequacy. Previous methodological literature emphasizes that acceptable model fit should be interpreted holistically by considering theoretical justification, factor loadings, and reliability estimates rather than relying solely on numerical thresholds. Therefore, while the findings support the proposed factor structure, the factorial validity evidence should be interpreted with caution and considered preliminary. Furthermore, because both EFA and CFA were conducted using the same sample, future studies involving independent samples are needed to provide additional evidence regarding the stability and generalizability of the factor structure.[18-20]

After finalising DM-KS, we examined its ability to distinguish subgroups defined by relevant teacher characteristics. Teachers with chronic diseases had higher DM-KS scores than those without chronic conditions. Comparable observations have been reported in studies from rural China, where adults with chronic diseases or a positive family history showed more extensive knowledge of chronic conditions, and among older individuals with a family history of chronic disease.[21,22] These findings likely reflect increased health-seeking behaviour and more frequent contact with healthcare professionals, which in turn may enhance disease-related knowledge.

Strengths and Limitations

This study contributes to the existing literature by providing a psychometrically tested instrument developed and evaluated among teachers to assess knowledge regarding type 1 diabetes mellitus in school settings. Although several diabetes knowledge scales exist in the literature, most of them were developed for patients or healthcare professionals and may not adequately capture the educational and practical needs of school personnel. The

present scale addresses this gap by focusing on teachers, who play a critical role in recognizing symptoms, preventing complications, and responding to diabetes-related emergencies during school hours. In addition, the inclusion of both exploratory and confirmatory factor analyses provides preliminary evidence regarding the structural validity of the instrument. Therefore, *DM-KS* may serve as a useful tool for assessing educational needs, planning training programs, and evaluating interventions aimed at improving diabetes management and safety in school environments.

This study has several limitations that should be considered when interpreting the findings. First, the study was conducted in a single district, and participants were recruited using a convenience sampling approach, although schools were randomly selected, which may limit the generalizability of the results to other regions or populations. Second, data were collected using self-report questionnaires, which may introduce response bias, including social desirability bias, and may not fully reflect participants' actual knowledge levels. Third, although both exploratory and confirmatory factor analysis were performed, the CFA was conducted on the same sample rather than an independent validation sample, which may limit the strength of the structural validation. In addition, some CFA model fit indices did not fully meet the recommended thresholds, and therefore the factorial validity findings should be interpreted with caution. Additionally, test–retest reliability was not assessed due to time constraints, and therefore the temporal stability of the scale could not be evaluated. Future studies including larger and more diverse samples, independent validation groups, and longitudinal reliability assessments are recommended to further strengthen the psychometric properties of the scale.

In conclusion, the Diabetes Mellitus Knowledge Scale (DM-KS) demonstrated acceptable reliability and preliminary evidence of validity for assessing teachers' knowledge regarding type 1 diabetes mellitus and its management in school settings. The findings suggest that the scale has promising psychometric properties and may serve as a practical tool for identifying knowledge gaps among teachers and guiding educational

interventions. Further validation in independent samples is warranted to confirm the stability and generalizability of its factor structure.

The DM-KS may also be useful for evaluating the effectiveness of training programs and supporting school health policies aimed at enhancing preparedness for diabetes care in educational environments. Future research involving larger and more diverse populations, independent validation samples, and longitudinal studies assessing temporal stability would provide additional evidence regarding the applicability and psychometric performance of the scale across different settings.

Ethical Considerations: This study was carried out within the scope of the project titled: "Evaluation of the Knowledge Level of School Teachers about Diabetes Mellitus", and SBÜ Dr. Abdurrahman Yurtaslan Ankara Onkoloji Research and Hospital Ethics Committee approval (2022-10/172) was received to conduct this study.

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

References

1. Alberti KGMM, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus. Diabet Med. 1998;15(7):539-553. doi:10.1002/(SICI)1096-9136(199807)15:7<539::AID-DIA668>3.0.CO;2-S.
2. Turkish Society of Endocrinology and Metabolism. Diabetes mellitus ve komplikasyonlarının tanı, tedavi ve izlem kılavuzu. 15th ed. Ankara: Bayt Yayınları; 2022.
3. International Diabetes Federation. IDF Diabetes Atlas 2021 [Internet]. Brussels: International Diabetes Federation; 2021 [cited 2025 Aug 7]. Available from: https://diabetesatlas.org/atlas/tenth-edition/
4. World Health Organization. Diabetes [Internet]. Geneva: World Health Organization; 2023 [cited 2025 Aug 7]. Available from: https://www.who.int/news-room/fact-sheets/detail/diabetes
5. Roglic G. WHO global report on diabetes: A summary. Int J Noncommun Dis. 2016;1(1):3-8. doi:10.4103/2468-8827.184853
6. Hatun S. Okulda diyabet programı eğitim platformu [Internet]. 2022 [cited 2025 Aug 7]. Available from: https://okuldadiyabet.meb.gov.tr/okulda-diyabet-programi-hakkinda
7. Alpar R. Spor, sağlık ve eğitim bilimlerinden örneklerle uygulamalı istatistik ve geçerlik güvenirlik. 7th ed. Ankara: Detay Yayıncılık; 2022.
8. Tavşancıl E. Tutumların ölçülmesi ve SPSS ile veri analizi. 6th ed. Ankara: Nobel Akademik Yayıncılık; 2010.
9. Republic of Turkey Ministry of Health, Public Health Institution of Turkey. Türkiye Diyabet Programı 2015-2020. Ankara: Kuban Yayıncılık; 2014.
10. Baykul Y. Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması. 7th ed. Ankara: Pegem Akademi; 2024.
11. Garcia AA, Villagomez ET, Brown SA, Kouzekanani K, Hanis CL. The Starr County diabetes education study: Development of the Spanish-language diabetes

knowledge questionnaire. Diabetes Care. 2001;24(1):16-21. doi:10.2337/diacare.24.1.16

12. Hsieh MH, Chen YC, Ho CH, Lin CY. Validation of diabetes knowledge questionnaire (DKQ) in the Taiwanese population: Concurrent validity with diabetes-specific quality of life questionnaire module. Diabetes Metab Syndr Obes. 2022;15:2391-2403. doi:10.2147/DMSO.S369552
13. Yavuz K, Erol S. Developing a diabetes knowledge scale for adults and its psychometric properties. Clin Exp Health Sci. 2022;12(3):711-718. doi:10.33808/clinexphealthsci.1014888
14. Bukhsh A, Lee SWH, Pusparajah P, Khan AH, Khan TM. Psychometric properties of the Urdu version of diabetes knowledge questionnaire. Front Public Health. 2017;5:139. doi:10.3389/fpubh.2017.00139
15. Al-Qazaz HK, Hassali MA, Shafie AA, Sulaiman S, Sundram S. The 14-item Michigan diabetes knowledge test: Translation and validation study of the Malaysian version. Pract Diabetes Int. 2010;27(6):238-241. doi:10.1002/pdi.1495
16. Cacic M, Kruljac I, Mirosevic G, Vrkljan M. Validation of diabetes knowledge questionnaire in Croatian with assessment of diabetes knowledge and quality of life in patients with type 2 diabetes mellitus. Clin Diabetol. 2020;9(6):387-393. doi:10.5603/DK.2020.0054.
17. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. J Educ Res. 2006;99(6):323-338. doi:10.3200/JOER.99.6.323-338
18. Brown TA. Confirmatory factor analysis for applied research. 2nd ed. New York: Guilford Press; 2015.
19. Brown TA, Moore M. Confirmatory factor analysis. In: Hoyle RH, editor. Handbook of structural equation modeling. New York: Guilford Press; 2012. p. 361-379.
20. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociol Methods Res. 1992;21(2):230-258. doi:10.1177/0049124192021002005
21. Tian M, Chen Y, Zhao R, Chen L, Chen X, Feng D, et al. Chronic disease knowledge and its determinants among chronically ill adults in rural areas of Shanxi Province

in China: A cross-sectional study. BMC Public Health. 2011;11:948. doi:10.1186/1471-2458-11-948.

22. Song Y, Ma W, Yi X, Wang S, Sun X, Tian J, et al. Chronic diseases knowledge and related factors among the elderly in Jinan, China. PLoS One. 2013;8(6). doi:10.1371/journal.pone.0068599

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):467-481 // 10.5505/amj.2026.60094

EVALUATION OF SOCIAL MEDIA ADDICTION AND OBESITY AWARENESS IN CHILDREN AGED 10–14

Sultan İlay Tunçez[1], İrfan Şencan[2]

[1]The Central District Health Directorate, Düzce, Türkiye

[2]Department of Family Medicine, Sincan Training and Research Hospital, Ankara, Türkiye

Correspondence:

Sultan İlay Tunçez (e-mail: ilaytncz@gmail.com)

Submitted: 17.03.2026 // Accepted: 27.07.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine

Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: This study aimed to evaluate levels of social media addiction and obesity awareness among children aged 10–14, examine their relationship, and identify factors affecting both conditions within a primary care context.

Materials and Methods: An analytical cross-sectional study was conducted with 228 children aged 10–14 years who attended central secondary schools in a district in Türkiye. Those who agreed to participate in the survey were asked to complete a Personal Information Form to record their socio-demographic characteristics, as well as the Social Media Addiction Scale for Adolescents(SMAS-A) and the Obesity Awareness Scale(OAS). Approval for the study was obtained from the Ankara City Hospital Ethics Committee (approval number: E1-22-2566).

Results: Among the participants, 23.2% were overweight, 10.1% were obese, and the mean daily social media use was 128.88 minutes. Social media addiction levels were low (18.20±7.54), while obesity awareness was high (56.39±7.76). A weak but statistically significant negative correlation was found between social media addiction and obesity awareness. In multivariable linear regression analysis, social media addiction scores remained independently associated with obesity awareness. While no significant association was found between BMI and social media addiction, a weak negative association was observed between BMI and obesity awareness. Regular exercise was significantly associated with both scale scores. A positive significant correlation was detected between daily social media use and addiction scores, whereas a negative significant correlation was found with obesity awareness.

Conclusion: School-based and primary care-oriented initiatives that promote physical activity and healthy digital habits may contribute to improving adolescent health, increasing obesity awareness, and preventing problematic social media use.

Keywords: Social media addiction, obesity, health knowledge, practice, adolescent, primary health care.

Introduction

In Turkey, 91.3% of children aged 6–15 use the internet, and 66.1% of those use it for social media.[1] Social media encompasses online environments where users can communicate, exchange information, and actively participate in content creation and sharing.[2] While these platforms facilitate the exchange of ideas and information, studies indicate that adolescents may be more likely than adults to exhibit problematic social media use, which can have negative effects on social relationships, physical and mental health, and academic performance.[2-3]

Although problematic social media use has attracted increasing scientific attention, it has not yet been recognized as a distinct diagnostic entity in the Statistical Manual of Mental Disorders-V (DSM-V). Behavioral addictions related to internet use are currently limited to gaming disorder, resulting in relatively limited investigation of other forms of problematic online behavior.[2,4] Problematic engagement with digital media has been linked to several lifestyle-related behaviors, including decreased physical activity, prolonged sedentary time, unhealthy eating patterns, and sleep-related problems.[5-6] Therefore, social media addiction may be associated with obesity.

Early adolescence, defined as the age range of 10–14 years, is a critical developmental period characterized by rapid biological, cognitive, and psychosocial changes.[7] During this stage, individuals begin to gain autonomy, peer relationships become more prominent, and long-term behavioral patterns start to emerge. In addition, immature impulse control may increase susceptibility to risk-taking behaviors. Problematic social media use during this period may be attributed to decreased parental control and authority, an increased need for social interaction, access to information and entertainment, and using social media as a coping mechanism for emerging emotional difficulties.[2,6-7] These characteristics make early adolescence a critical period for preventive health interventions.

Health-related risks such as obesity and problematic digital media use may be associated with lifestyle behaviors established during this developmental period.[2,7] Unhealthy dietary

patterns, including the consumption of high-calorie fast food, a sedentary lifestyle, and prolonged screen time, contribute to energy imbalance and are well-established risk factors for obesity. The rising prevalence of childhood obesity is a major global public health concern as it increases the burden of chronic non-communicable diseases.[8-10]

The relationship between social media addiction and obesity is thought to be mediated by behavioral and lifestyle changes associated with addictive use patterns. While previous studies have primarily focused on the relationship between physical inactivity and obesity, adolescents with problematic internet use have also been reported to exhibit behaviors such as unhealthy eating, reduced physical activity, and impaired social functioning, all of which may increase the risk of obesity.[11] Nevertheless, elucidating the relationship between social media addiction and obesity awareness from a primary care perspective remains limited. The focus of previous research has been on obesity outcomes or physical inactivity, with limited attention to obesity awareness as a behavioral and cognitive dimension. Nevertheless, obesity awareness may influence how individuals perceive health risks and adopt healthy lifestyle behaviors.

This study aims to evaluate levels of social media addiction and obesity awareness among children aged 10–14, examine the relationship between the two, and identify factors affecting both conditions within a primary care context.

Materials and Methods

This analytical cross-sectional study was conducted between June 6 and June 13, 2022, with students aged 10–14 attending central secondary schools in the Konya district of Yulak. The study population consisted of 634 students enrolled in these schools. Using the sample size formula for known populations, the sample size was calculated as 240, with a 5% margin of error and a 95% confidence interval. However, due to school absenteeism related to seasonal agricultural work, as well as limitations in voluntary participation based on informed consent, the study was ultimately completed with 228 students.

The inclusion criteria for the study were:

I.being aged 10–14 years

II.having sufficient cognitive ability to answer the questionnaires

III.being enrolled in a secondary school in Konya.

Students who met these criteria and who voluntarily agreed to participate were included in the study; those who wished to withdraw during the data collection process were excluded.

Ethical approval for the study was obtained from the Clinical Research Ethics Committee of Ankara City Hospital (approval no: E1-22-2566, date: 1 June, 2022), and institutional permission was obtained from the provincial directorate of national education (no: E83688308-605.99-46555979, date: 28 March, 2022). All participants and their parents/legal guardians were informed about the study and provided written informed consent. Data were collected by the researchers through face-to-face interviews using a personal information form and two scales: the SMAS-A and the OAS. The Personal Information Form included questions on age, gender, height (in m), weight (in kg), number of siblings, ownership of a personal room and digital devices, physical activity habits, and duration of daily social media use (minutes/day). The researcher took height and weight measurements of the participants using calibrated devices and calculated BMI by dividing weight (kg) by the square of height (m^2).

The SMAS-A, developed by Ozgenel et al., is based on the DSM-V Internet Gaming Disorder criteria and consists of nine items on a five-point Likert scale. Scores range from 9 to 45, with higher scores indicating a higher level of social media addiction. The Cronbach's alpha reliability coefficient of the scale was reported as 0.904.[12]

The OAS was originally developed by Allen in 2011 and subsequently adapted into Turkish by Kafkas and Ozen, who examined its psychometric properties. Participants respond using a four-point Likert-type response format. Scores may vary between 20 and 80, with larger

values representing greater awareness of obesity-related issues. The Turkish version demonstrated good internal consistency (Cronbach's α = 0.87).[13]

A statistical analysis of the data was performed using SPSS software (version 22.0). The normality of the data distribution was assessed using the Kolmogorov–Smirnov test. Categorical variables were compared using the chi-squared test or Fisher's exact test. Group comparisons of continuous variables were performed using an independent samples t-test or ANOVA for normally distributed data and a Mann–Whitney U test or Kruskal–Wallis test for non-normally distributed data. Relationships between continuous variables were examined using Pearson or Spearman's correlation analysis, depending on the distribution characteristics of the data. Multiple linear regression analysis was performed to identify factors associated with obesity awareness. A p-value of less than 0.05 was considered statistically significant.

Results

Of the 228 students who participated in the study, 52.6% (n=120) were female, and 47.4% (n=108) were male. The mean age of the participants was 12.13±1.15 years. According to BMI classification, 52.2% of participants were of normal weight, 23.2% were overweight, 14.5% were underweight, and 10.1% were obese. The mean time spent on social media each day was 128.88±115.70 minutes. The demographic characteristics and anthropometric measurements of the participants are presented in Table 1, while their personal digital device ownership and lifestyle characteristics are presented in Table 2.

Participants achieved an average total score of 18.20±7.54 (range 9–43) on the SMAS-A, and an average total score of 56.39±7.76 (range 23–77) on the OAS (Table 3).

The total SMAS-A and OAS scores were not normally distributed. Spearman's correlation analysis revealed a statistically significant but weak negative correlation between the two scores ($r=-0.237$; $p<0.001$) (Table 4).

Table 1. Demographic Characteristics and Body Measurements

Variable	Mean	Median	SD	Min	Max
Age (years)	12.13	12.00	1.15	10.00	14.00
Number of	2.99	3.00	1.33	1.00	10.00
Height (m)	1.51	1.52	0.09	1.17	1.75
Weight (kg)	46.07	45.00	12.06	22.00	96.00
BMI (kg/m^2)	19.67	19.02	3.66	11.72	31.83
Social Media	128.88	90.00	115.70	0.00	720.00

Table 2. Personal Digital Device Ownership and Lifestyle Characteristics

Variable n=228 (%)	Yes n (%)	No n (%)
Phone	88 (38.6)	140 (61.4)
Tablet-PC	131 (57.5)	97 (42.5)
Own Room	138 (60.5)	90 (39.5)
Exercise	108 (47.4)	120 (52.6)

Table 3. Distribution of Participants' SMAS-A and OAS Scores

	Mean	Median	SD	Min	Max
SMAS-A Total Score	18.20	17.50	7.542	9	43
OAS Total Score	56.39	57.00	7.761	23	77

Table 4. Correlation Between SMAS-A Total Scores and OAS Total Scores

Spearman's rho correlation		OAS Total Score
SMAS-A Total Score	Correlation Coefficient (r)	-.237*
	p	<0.001*
	n	228

When OAS scores were compared according to whether participants had their own room, those with their own room had significantly higher mean OAS scores (57.19±8.01) than those who did not (p=0.010). No statistically significant difference was found between the same groups in terms of SMAS-A scores (p=0.758) (Table 5).

Students who engaged in at least 30 minutes of physical activity, on at least three days per week, had significantly lower mean SMAS-A scores (16.76±5.41) than those who did not (p=0.011), as well as significantly higher mean OAS scores (58.45±6.68) (p<0.001) (Table 5).

Table 5. Comparison of SMAS-A and OAS Total Scores According to Exercise Status and Personal Room Ownership

Variable	Group	n	Mean±SD	p
SMAS-A	Exercise (+)	108	16.76±5.41	0.011*
	Exercise (-)	120	20.11±8.76	
OAS	Exercise (+)	108	58.45±6.68	<0.001*
	Exercise (-)	120	54.54±8.21	
	Own Room (+)	138	57.19±8.01	0.010*
	Own Room (-)	90	55.18±7.24	

*Mann–Whitney U test was used. p<0.05 was considered statistically significant.

No statistically significant relationship was found between participants' BMI and their total SMAS-A scores (p=0.245). However, a weak negative correlation was found between BMI and OAS scores (r=-0.141, p=0.033) (Table 6).

A moderate positive correlation was found between daily social media usage time and SMAS-A scores (r=0.544; p<0.001), while a weak negative correlation was found between daily social media usage time and OAS scores (r =-0.141; p=0.034) (Table 6).

Table 6. Correlation of SMAS-A and OAS Total Scores with BMI and Social Media Duration

N=228	SMAS-A Total Score		OAS Total Score	
	Correlation Coefficient (r)	p	Correlation Coefficient (r)	P
BMI	0.077	0.245	-0.141*	0.033*
Social Media Duration		p<0.001*	-0.141*	0.034*

*p<0.05 indicates statistical significance

A multivariable linear regression model was constructed to identify independent predictors of obesity awareness scores. The model was statistically significant (F=6.503, p<0.001) and explained 17.1% of the variance in OAS scores (R^2=0.171). Female sex, having a personal room, engaging in regular exercise, and lower SMAS-A scores were independently associated with higher OAS scores. Age, BMI, and daily social media use were not significant predictors. Regression analysis findings are presented in Table 7.

Table 7. Multivariate Linear Regression Analysis of Factors Associated with OAS Scores

Variable	**B**	**Beta**	**p**	**%95 Confidence Interval**	
				Lower Bound	Upper Bound
Age	0,048	0,428	0,910	-0,795	0,892
Gender (Reference: Man)	2,107	0,136	0,041*	0,090	4,124
BMI	-0,262	-0,124	0,060	-0,535	0,011
Social Media Duration	0,007	0,111	0,146	-0,003	0,017
Own Room (Reference: No)	2,247	0,142	0,023*	0,316	4,179
Exercise (Reference: No)	3,670	0,237	<0,001*	1,665	5,675
SMAS-A	-0,256	-0,249	0,001*	-0,409	-0,102
Constant=60,588 R=0,414 R^2=0,171 Durbin-Watson=1,839 F=6,503 p<0,001					

Discussion

As social media use becomes more widespread, it is increasingly important to understand the effect of digital behaviors on health during early adolescence. Unlike previous studies, which have primarily focused on BMI, physical inactivity, or obesity-related outcomes, this study examined the relationship between social media addiction and obesity awareness, a cognitive and attitudinal health indicator. This approach may contribute to a better understanding of the potential links between social media use and health-related cognitive processes. Furthermore, evaluating this relationship from a primary healthcare perspective, focusing on early adolescence, may contribute to an area that has received limited attention in the literature.

In this study, a statistically significant but weak, negative correlation was identified between social media addiction and obesity awareness. Furthermore, multivariable linear regression analysis demonstrated that social media addiction remained independently associated with obesity awareness after adjustment for age, sex, BMI, regular exercise, daily social media use, and ownership of a private room. This finding suggests that the observed association may not be explained solely by factors such as physical activity, sex, or duration of social media use. The absence of a robust correlation observed in this study may be attributable to several factors, including the modest sample size, the cross-sectional design, and the unique characteristics inherent to early adolescence. These findings may indicate that levels of social media addiction observed during early adolescence are associated with disparities in obesity awareness. Although the underlying mechanisms remain unclear, problematic social media use may be associated with differences in health information processing, health risk perception, or engagement with healthy lifestyle messages.

Research directly examining the relationship between social media addiction and obesity awareness remains sparse. Most previous research investigating digital media use has focused on obesity, BMI, physical inactivity, or lifestyle-related outcomes and has reported associations between internet or social media use and obesity-related lifestyle factors.[5,11,14.]

These findings offer indirect support for the association observed in this study. The construct of obesity awareness is multifaceted, reflecting how individuals perceive health risks and their attitudes towards healthy lifestyle behaviors. Consequently, the evaluation of obesity awareness during early adolescence may facilitate a more comprehensive understanding of current weight status, as well as identify cognitive and behavioral tendencies that could influence future health behaviors.

Mean total SMAS-A scores indicate that participants' social media addiction levels were generally closer to the low end of the scale, which is consistent with previous studies.[15] However, other studies have found that problematic social media use increases with age.[16] This may indicate that problematic social media use remains relatively limited during early adolescence.

The mean total OAS score indicated a generally good level of obesity awareness. Our findings are consistent with previous studies that have reported high levels of obesity awareness among adolescents.[17-18.] This finding may be related to the influence of health education in schools, health information disseminated through media channels, and increasing public awareness regarding obesity. However, it should be noted that high levels of awareness do not necessarily result in the adoption of healthy behaviors. Consequently, the consideration of awareness and behavior as discrete yet interconnected constructs may facilitate the development of more efficacious interventions.

Regular physical activity was associated with both lower social media addiction scores and higher obesity awareness levels. Similar findings have been reported in high school and university students, although one study found no significant relationship between physical activity and obesity awareness, possibly due to differences in age distribution and low activity prevalence. [19-21] Overall, existing evidence supports the possibility that physical activity may be associated with more controlled social media use and higher obesity awareness.

Additionally, students with their own room had significantly higher obesity awareness scores, though no significant difference was observed in social media addiction scores. This finding may be related to socioeconomic conditions and access to health-related information. However, the lack of an association between having a private room and social media addiction may suggest that the physical environment alone may not determine this behavior.

The negative correlation between BMI and obesity awareness scores suggests that individuals with a higher BMI may be less aware of their obesity. This finding highlights the importance of interventions that enhance awareness for at-risk individuals. While some studies have found no significant correlation between BMI and obesity awareness, others have reported significant differences across BMI categories.[18,21-22] In contrast, although previous studies have reported an association between internet addiction and BMI, no significant relationship was observed between BMI and social media addiction scores in the present study. [23] This discrepancy may be related to the fact that our study focused specifically on individuals in early adolescence (10-14 years).

Finally, consistent with previous studies, daily social media use was correlated with both SMAS-A and OAS scores, suggesting that the duration of digital media use may be associated with these variables.[24-25] However, in the multivariable analysis, daily social media use was not identified as an independent predictor of obesity awareness. This finding may indicate that the observed association is influenced by other individual and environmental factors.

From a primary healthcare perspective, these findings support the importance of assessing social media usage habits during routine adolescent check-ups and periodic follow-up visits, alongside providing lifestyle counselling on physical activity and nutrition. In particular, conducting brief assessments regarding social media use, physical activity levels, and obesity awareness during routine health checkups can help identify at-risk groups at an early stage, as well as raise awareness of health risks among children who have not yet developed weight issues and promote healthy lifestyle behaviors. For adolescents identified as having low obesity awareness or problematic social media use, brief counseling interventions and family-based educational approaches can be implemented.

In conclusion, our findings suggest a potential association between social media addiction and obesity awareness. This indicates that when evaluating adolescent health, not only behavioral outcomes should be considered but also health awareness and perceptions that may influence future health behaviors. Interventions aimed at promoting regular physical activity, encouraging healthy digital habits, and improving obesity awareness may contribute to the promotion of adolescent health. In this context, early adolescence represents an important developmental period that warrants attention within preventive health strategies.

This study has several limitations. Data collection at the end of the academic year may have affected participation rates. The inability to recruit the planned number of participants may have made it more difficult to identify associations of small magnitude. The single-center design may limit generalizability of the findings. Additionally, while device ownership was assessed, patterns of use were not examined in detail.

Ethical Considerations: The study was approved by the Clinical Research Ethics Committee of Ankara City Hospital with approval number: E1-22-2566

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

This research was presented as an oral presentation at the 18th Family Medicine Autumn School in the Turkish Republic of Northern Cyprus on September 26, 2024.

References

1. TUIK. Hanehalkı Bilişim Teknolojileri (BT) Kullanım araştırması, 2024 [Internet] https://data.tuik.gov.tr/Bulten/Index?p=Cocuklarda-Bilisim-Teknolojileri-Kullanim-Arastirmasi2024-53638 (Accessed:23.03.2026).
2. Ciacchini R, Orrù G, Cucurnia E, Sabbatini S, Scafuto F, Lazzarelli A, et al. Social media in adolescents: A retrospective correlational study on addiction. *Children*. 2023;10(2):278.
3. Lin S, Mastrokoukou S, Longobardi C. Social relationships and social media addiction among adolescents: Variable-centered and person-centered approaches. *Comput Human Behav*. 2023;147:107869
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-V*. 5th ed. Washington, DC: American Psychiatric Association; 2013.
5. Muslu M, Gokcay GF. Teknoloji bağımlısı çocuklarda obeziteye neden olan risk faktörleri. *Gümüşhane Univ Saglik Bilim Derg*. 2019;8(2):72-9.
6. Andreassen CS. Online social network site addiction: A comprehensive review. *Curr Addict Rep*. 2015;2(2):175-84.
7. Blum RW, Astone NM, Decker MR, Mouli VC. A conceptual framework for early adolescence: a platform for research. *Int J Adolesc Med Health*. 2014;26(3):321-31.
8. Kansra AR, Lakkunarajah S, Jay MS. Childhood and adolescent obesity: A review. *Front Pediatr*. 2021;8:581461.
9. World Health Organization. Obesity [Internet]. https://www.who.int/health-topics/obesity#tab=tab (Accessed: 23.03.2026).
10. Cali AM, Caprio S. Obesity in children and adolescents. *J Clin Endocrinol Metab*. 2008;93(11 Suppl 1):S31-6.
11. Li M, Deng Y, Ren Y, Guo S, He X. Obesity status of middle school students in Xiangtan and its relationship with internet addiction. *Obesity.* 2014;22(2):482-7.
12. Ozgenel M, Canpolat O, Eksi H. Ergenler için Sosyal Medya Bağımlılığı Ölçeği (ESMBÖ): Geçerlik ve güvenirlik çalışması. *Addicta(Turk J Addict).* 2019;6(3):629-54.

13. Kafkas ME, Ozen G. Obezite Farkındalık Ölçeği'nin (OFÖ) Türkçeye uyarlanması: Bir geçerlik ve güvenirlik çalışması. *Inonu Univ Beden Egit Spor Bilim Derg.* 2014;1(2):1-15.
14. Bayrak T, Cihan B. Yeni medyada bağımlılık sonucu gelişen dijital obezite olgusu üzerine bir inceleme: Netflix Türkiye örneği. *E J New Media*. 2021;5(1):1-15.
15. Buzzi C, Tucci M, Ciprandi R, Brambilla I, Caimmi S, Ciprandi G, et al. The psycho-social effects of COVID-19 on Italian adolescents' attitudes and behaviors. *Ital J Pediatr.* 2020;46(1):69.
16. Güney M, Taştepe T. Ergenlerde sosyal medya kullanımı ve sosyal medya bağımlılığı. *Ankara Saglik Bilim Derg.* 2020;9(2):183-90.
17. Ozmeral M. *Obesity awareness in adolescents attending high school and their parents*. Ph.D. Thesis. University of Health Sciences, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Departmant of Family Medicine;İstanbul; 2020.33-59
18. Yildirim M, Kartal BSA. Lise öğrencilerinin obezite farkındalık düzeylerinin incelenmesi. *J Soc Humanit Adm Sci.* 2022;8(51):183-92.
19. Can B, Hazar Z, Koç S Spor yapan ve yapmayan lise öğrencilerinin sosyal medya bağımlılık düzeylerinin bazı değişkenlere göre incelenmesi (karma model araştırması). *Gaziantep Univ Spor Bilim Derg.*2021;6(1):15-39.
20. Alemdag S. Üniversite öğrencilerinin fiziksel aktiviteye katılım ve sosyal medya ilişkisi. *Akad Sos Arast Derg.* 2022;10(128):85-96.
21. Ozkan I, Adibelli D, Kirak M, Sahin S. Üniversite öğrencilerinin obezite farkındalıkları ile beden kitle indeksleri arasındaki ilişki. *Acibadem Univ Saglik Bilim Derg*. 2020;(1):120-6.
22. Yesiltepe A, Karadag G, Gurkan KP. The effect of obesity awareness on blood pressure and body mass index in primary school students. *Gevher Nesibe Journal of Medical and Health Sciences*. 2022;7(17):111-8.
23. Canan F, Yildirim O, Ustunel TY, Sinani G, Kaleli AH, Gunes C, et al. The relationship between internet addiction and body mass index in Turkish adolescents. *Cyberpsychol Behav Soc Netw*. 2014;17(1):40-5.

24. Fullerton S, Taylor AW, Dal Grande E, Berry N. Measuring physical inactivity: Do current measures provide an accurate view of "sedentary" video game time? *J Obes.* 2014;2014:287013.
25. Inal Ö. Eating disorders, emotional eating, internet and social media addiction in adult individuals during the Covid-19 period. *Akd Med J*. 2022;8(3):291-7

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):482-499 // 10.5505/amj.2026.61257

THE ROLE OF THE C-REACTIVE PROTEIN-TO-ALBUMIN RATIO AND HEMATOLOGICAL INFLAMMATORY MARKERS IN PREDICTING THE SEVERITY OF OBSTRUCTIVE SLEEP APNEA SYNDROME

Memet Aslanyavrusu[1], Ali Sönmez[1], Mehmet Fatih Göl[1]

[1]Department of Neurology, Kayseri City Hospital, Kayseri, Türkiye

Correspondence:

Memet Aslanyavrusu (e-mail: aslanyavrusu@yahoo.com)

Submitted: 30.01.2026 // Accepted: 02.07.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine

Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: It is well established that systemic inflammatory responses induced by intermittent hypoxia are associated with disease severity in obstructive sleep apnea syndrome (OSAS). This study aimed to investigate the association between the C-reactive protein–to–albumin ratio (CAR), hematological inflammatory markers, and disease severity in OSAS, and to assess their diagnostic performance for predicting moderate-to-severe OSAS.

Materials and Methods: We included 230 individuals aged 18–65 years who underwent full-night polysomnography in our sleep laboratory between June 2025 and December 2025. Participants were classified as normal, mild, moderate, or severe OSAS according to the apnea–hypopnea index (AHI). Serum C-reactive protein (CRP) and albumin levels were measured in all participants, and CAR, neutrophil-to-lymphocyte ratio (NLR), and systemic immune-inflammation index (SII) were calculated. Independent predictors of moderate-to-severe OSAS were assessed using multivariable logistic regression analyses. The discriminative performance of inflammatory markers was evaluated using receiver operating characteristic (ROC) curve analysis.

Results: As OSAS severity increased, CRP, CAR, NLR, and SII levels showed a significant upward trend, whereas serum albumin levels decreased (all $p < 0.001$). In multivariable analyses, CAR, NLR, and SII were independently associated with the presence of moderate-to-severe OSAS. ROC analysis demonstrated that CAR had the highest diagnostic accuracy for distinguishing moderate-to-severe OSAS.

Conclusion: Increasing OSAS severity is associated with a marked elevation in systemic inflammatory burden. CAR may serve as a supportive inflammatory marker for identifying patients with moderate-to-severe OSAS.

Keywords: Sleep apnea, obstructive, C-reactive protein, albumins, inflammation

Introduction

Obstructive sleep apnea syndrome (OSAS) is characterized by recurrent collapse or instability of the upper airway during sleep, leading to intermittent hypoxemia and repeated arousals that disrupt normal sleep architecture and systemic physiological regulation.[1] Epidemiological studies report that the prevalence of OSAS among adults aged 30–60 years is approximately 24% in men and 9% in women. In addition, the landmark study by Benjafield et al. demonstrated that nearly one billion individuals worldwide are affected by OSAS.[2] Obesity, typically assessed by body mass index (BMI), represents the most important risk factor for OSAS. In addition, advancing age, male sex, alcohol consumption, cigarette smoking, and upper airway anatomical features also contribute to the development of the disease.[3]

Intermittent hypoxia, increased sympathetic activity, and sleep fragmentation—hallmark features of OSAS pathophysiology—trigger systemic inflammatory responses that contribute to the development of cardiovascular, metabolic, and neurodegenerative complications. Experimental and clinical evidence suggests that inflammation is not merely a consequence of OSAS but may also exacerbate upper airway collapsibility, thereby perpetuating disease severity and facilitating excessive daytime sleepiness.[4] In this context, elevated levels of inflammatory mediators such as C-reactive protein (CRP), tumor necrosis factor-α, interleukin-6, and interleukin-8 have been consistently reported in patients with OSAS.[5]

CRP is a liver-derived positive acute-phase reactant that rapidly increases in response to inflammation or infection. In contrast, albumin is a negative acute-phase reactant with antioxidant and anti-inflammatory properties. The C-reactive protein–to–albumin ratio (CAR) has therefore emerged as a composite index that may more comprehensively reflect the severity of systemic inflammation and has been suggested to be associated with adverse clinical outcomes. Previous studies have demonstrated that CAR is linked to prognosis in various clinical conditions, including stroke, malignancy, and sepsis.[6-7]

On the other hand, neutrophil activation accompanied by suppression of lymphocyte responses represents a key component of systemic inflammatory processes. The neutrophil-to-lymphocyte ratio (NLR), which integrates these two opposing cellular responses, has been widely used as a prognostic marker in numerous clinical settings, particularly in malignancies, chronic inflammatory diseases, and cardiovascular disorders.[8-10] However, data regarding the clinical and prognostic value of NLR in patients with OSAS remain limited. More recently, the systemic immune-inflammation index (SII), calculated as platelet × neutrophil/lymphocyte, has been proposed as a potentially more sensitive indicator among peripheral blood inflammatory markers.[10] Nevertheless, the clinical significance of SII and its relationship with disease severity in OSAS have not yet been fully elucidated.

Given that a substantial proportion of patients with undiagnosed OSAS initially present to primary care settings, simple and accessible inflammatory markers may support early identification, facilitate risk stratification, and guide appropriate referral for polysomnography. In this context, this study explores whether CAR and other hematological inflammatory indices, including NLR and SII, are associated with increasing clinical severity in patients with OSAS.

Materials and Methods

This retrospective study included newly diagnosed OSAS patients aged 18–65 years who had not received any previous OSAS-specific treatment, including CPAP therapy, and were evaluated in our sleep laboratory between June 2025 and December 2025. Patients diagnosed with mild, moderate, or severe OSAS based on full-night polysomnography (PSG), as well as individuals with normal PSG findings who met the inclusion criteria, were included in the analysis. OSAS severity was classified according to the apnea–hypopnea index (AHI). Participants with an AHI <5 were classified as normal, those with 5 ≤ AHI <15 as mild OSAS, those with 15 ≤ AHI <30 as moderate OSAS, and those with an AHI ≥ 30 as severe OSAS .[11] Ethical approval was obtained from the Kayseri City Hospital Ethics Committee (Approval

No. 628, November 4, 2025). Given that data were analyzed retrospectively, individual informed consent was not obtained.

Demographic data, including age, sex, and body mass index (BMI), were recorded for all participants. Blood samples were obtained on the same day as polysomnography evaluation, and serum albumin and CRP levels were measured to calculate CAR, NLR, and SII values. Individuals with a history of malignancy, inflammatory or rheumatologic diseases, chronic infections, acute infection, chronic obstructive pulmonary disease, chronic liver disease, or recent antibiotic or anti-inflammatory drug use were excluded from the study.

Statistical Analysis

All statistical procedures were conducted using SPSS software (version 22.0). Sample size estimation was performed using G*Power software (version 3.1). Based on a one-way ANOVA model (fixed effects, omnibus test), an alpha level of 0.05, a statistical power of 80%, and a medium effect size (Cohen's f = 0.25), the minimum required sample size was estimated to be approximately 180 participants. This approach was selected during the study planning stage to estimate the sample size required for comparisons across the four OSAS severity groups. Because some variables were not normally distributed, non-parametric methods were subsequently used for group comparisons. The final study population consisted of 230 participants, exceeding the minimum required sample size and ensuring adequate statistical power for the planned analyses.

Prior to comparative analyses, the distribution pattern of continuous variables was evaluated using the Kolmogorov–Smirnov test to determine the appropriate analytical approach. Variables were summarized in accordance with their distributional characteristics, whereas categorical data were presented as frequencies and percentages.

Group differences across OSAS severity levels were examined using non-parametric methods when distributional assumptions were not met and chi-square testing for categorical variables. Statistical significance was defined as $p < 0.05$.

To identify independent determinants of clinically relevant disease (AHI ≥15 events/hour), multivariable logistic regression models were constructed. Adjustment was made for age, sex, and BMI in all models. Given the potential overlap between inflammatory indices, CAR, NLR, and SII were evaluated separately. Associations are expressed as odds ratios with 95% confidence intervals. Model fit was assessed using calibration testing.

Discriminative performance was explored using ROC curve methodology, with overall accuracy quantified by the area under the curve. Cut-off values were selected based on optimal combined sensitivity and specificity.

Results

A total of 230 participants were included in the study and classified according to OSAS severity as normal (n = 49), mild (n = 27), moderate (n = 39), or severe OSAS (n = 115). There was no statistically significant difference in age among the groups (45.8 ± 12.4, 48.0 ± 11.1, 46.6 ± 10.6, and 49.7 ± 10.0 years, respectively; p = 0.246). Male sex predominated across all groups, and the proportion of male participants increased significantly with increasing OSAS severity (normal: 61.2%, mild OSAS: 66.7%, moderate OSAS: 71.8%, severe OSAS: 77.4%; p <0.001).

Significant differences were observed among the groups in terms of BMI, WBC, neutrophil count, CRP, albumin, NLR, CAR, and SII (all p <0.001). BMI, CRP, CAR, NLR, and SII values showed a progressive increase from the normal group to the severe OSAS group, whereas albumin levels decreased significantly with increasing OSAS severity. In contrast, no significant differences were observed in lymphocyte or platelet counts among the groups (p = 0.583 and p = 0.539, respectively) (Table 1).

Table 1. Distribution and comparison of inflammatory and hematological parameters according to OSAS severity

Variable	Normal	Mild	Moderate	Severe	H (χ^2)	p-value
Age	45.8 ± 12.4	48.0 ± 11.1	46.6 ± 10.6	49.7 ± 10.0	3.44	0.246
BMI (kg/m^2)	27.8 (25.4–32.4)	28.1 (27.0–30.1)	30.1 (27.8–36.0)	34.4 (30.1–38.1)	41.93	<0.001
WBC (×10^3/μL)	7.6 (6.4–9.3)	8.1 (6.9–9.7)	9.4 (7.9–10.4)	9.6 (8.1–11.5)	20.63	<0.001
Neutrophil (×10^3/μL)	4.4 (3.5–5.2)	4.5 (3.7–5.9)	5.5 (4.1–6.8)	5.9 (4.6–7.3)	28.37	<0.001
CRP (mg/L)	1.2 (0.7–1.8)	2.6 (1.5–4.6)	2.7 (1.9–7.7)	3.6 (2.3–6.8)	74.30	<0.001
Albumin (g/L)	46 (44–47)	45 (44.5–46)	45 (43.5–45.5)	44 (43–45)	23.60	<0.001
NLR	1.4 (1.2–1.7)	1.5 (1.3–1.9)	1.7 (1.3–2.1)	1.9 (1.6–2.4)	34.08	<0.001
CAR	0.02 (0.01–0.04)	0.05 (0.03–0.11)	0.06 (0.04–0.18)	0.08 (0.05–0.16)	76.87	<0.001
SII	384 (289–493)	410 (321–546)	435 (338–561)	508 (388–695)	21.33	<0.001
Lymphocyte (×10^3/μL)	2.9 (2.4–3.3)	2.7 (2.5–3.5)	3.1 (2.3–3.9)	2.8 (2.3–3.4)	1.95	0.583
Platelet (×10^3/μL)	257 (223–287)	280 (243–317)	262 (237–287)	267 (229–320)	2.17	0.539

Kruskal–Wallis test statistic; BMI, body mass index; CAR, C-reactive protein-to-albumin ratio; CRP, C-reactive protein; NLR, neutrophil-to-lymphocyte ratio; OSAS, obstructive sleep apnea syndrome; SII, systemic immune-inflammation index; WBC, white blood cell count. Data are presented according to their distribution characteristics. Non-normally distributed variables are expressed as median (interquartile range), whereas normally distributed variables are presented as mean ± standard deviation. Group comparisons were performed using the Kruskal–Wallis test for continuous variables and the chi-square test for categorical variables.

Multivariable logistic regression analyses were performed to identify independent predictors of moderate-to-severe OSAS. To avoid multicollinearity, CAR, NLR, and SII were

evaluated in separate regression models adjusted for age, sex, and BMI. In Model 1, CAR emerged as a strong and independent predictor of moderate-to-severe OSAS (OR = 1.158, 95% CI: 1.068–1.257; p <0.001). In Model 2, increased NLR values were independently associated with OSAS severity (OR = 5.379, 95% CI: 2.442–11.848; p <0.001). In Model 3, SII was also independently associated with the presence of moderate-to-severe OSAS (OR = 1.004, 95% CI: 1.002–1.006; p <0.001). Across all models, older age, male sex, and higher BMI remained significant predictors of moderate-to-severe OSAS (Table 2). Hosmer–Lemeshow goodness-of-fit testing demonstrated acceptable calibration for all multivariable logistic regression models (Model 1: p = 0.18, Model 2: p = 0.41, and Model 3: p = 0.42).

Table 2. Multivariable Logistic Regression Models for Predicting Moderate-to-Severe OSAS

Variable	Model 1: CAR		Model 2: NLR		Model 3: SII	
	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p
Age (years)	1.07 (1.04–1.11)	<0.001	1.08 (1.04–1.11)	<0.001	1.08 (1.04–1.12)	<0.001
Male sex	5.44 (2.29–12.92)	<0.001	4.90 (2.08–11.51)	<0.001	6.04 (2.56–14.25)	<0.001
BMI (kg/m^2)	1.17 (1.08–1.26)	<0.001	1.22 (1.13–1.31)	<0.001	1.20 (1.12–1.29)	<0.001
CAR (per 0.01 increase)	1.16 (1.07–1.26)	<0.001	—	—	—	—
NLR	—	—	5.38 (2.44–11.85)	<0.001	—	—
SII	—	—	—	—	1.004 (1.002–1.006)	<0.001

BMI, body mass index; CAR, C-reactive protein-to-albumin ratio; CI, confidence interval; NLR, neutrophil-to-lymphocyte ratio; OR, odds ratio; OSAS, obstructive sleep apnea syndrome; SII, systemic immune-inflammation index. Independent predictors of moderate-to-severe OSAS (AHI ≥ 15 events/hour) were evaluated using

multivariable logistic regression analysis. To minimize potential collinearity among inflammatory indices, CAR, NLR, and SII were analyzed in separate adjusted models.

In ROC curve analyses, CAR demonstrated the highest diagnostic performance for discriminating moderate-to-severe OSAS, with an AUC of 0.811. CRP showed good discriminative ability (AUC = 0.807), whereas NLR demonstrated acceptable discriminative performance (AUC = 0.718). In contrast, neutrophil count (AUC = 0.613), WBC count (AUC = 0.651), and SII (AUC = 0.668) exhibited more limited discriminative capacity. Because serum albumin levels were inversely associated with OSAS severity, the reciprocal albumin value (1/albumin) was used in ROC analysis, revealing a moderate discriminative ability for albumin (AUC = 0.663) (Table 3, Figure 1).

Table 3. ROC Analyses of Inflammatory Markers for Predicting Moderate-to-Severe OSAS

Marker	AUC (95% CI)	p-value	Optimal Cut-off	Sensitivity (%)	Specificity (%)
CAR	0.811 (0.747–0.871)	<0.001	≥ 0.045	80.5	67.1
CRP (mg/L)	0.807 (0.752–0.862)	<0.001	≥ 2.0	80.2	65.3
NLR	0.718 (0.649–0.781)	<0.001	≥ 1.45	82.7	51.5
Neutrophil(× 10^3/μL)	0.613 (0.546–0.680)	<0.001	≥ 3.4	75.6	42.5
WBC (×10^3/μL)	0.651 (0.611–0.751)	<0.001	≥ 7.8	77.4	40.3
SII	0.668 (0.595–0.736)	<0.001	≥ 340	81	54.2
Albumin (g/L)	0.663 (0.584-0.744)	<0.001	≤ 42.5	79.2	64.6

AUC, area under the curve; CI, confidence interval; CAR, C-reactive protein-to-albumin ratio; CRP, C-reactive protein; NLR, neutrophil-to-lymphocyte ratio; OSAS, obstructive sleep apnea syndrome; SII, systemic immune-inflammation index; WBC, white blood cell count. The discriminatory capacity of each marker was evaluated

using receiver operating characteristic (ROC) curve analysis. Diagnostic performance was quantified by the AUC, and optimal cut-off values were determined using the Youden index.

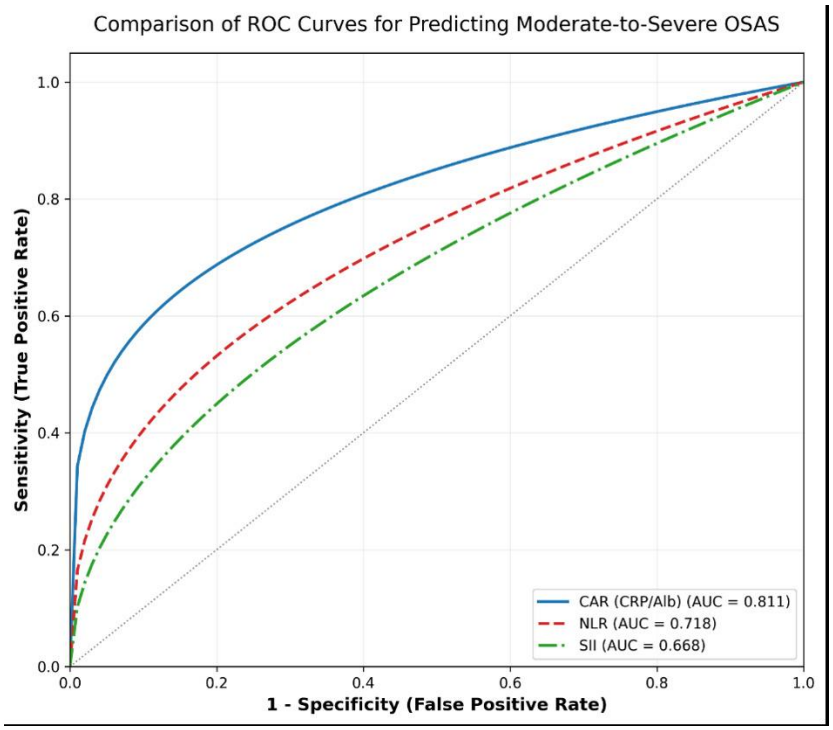


Figure 1. Comparison of ROC curves for systemic inflammatory markers in predicting moderate-to-severe OSAS.

Among the evaluated inflammatory markers, CAR yielded the highest AUC (0.811), indicating superior discriminatory performance for identifying moderate-to-severe OSAS. NLR (AUC = 0.718) and SII (AUC = 0.668) demonstrated comparatively lower performance. The diagonal reference line represents no discrimination (AUC = 0.5). All evaluated indices were significantly associated with clinically significant OSAS (AHI ≥ 15).

Discussion

This study comprehensively demonstrates the association between OSAS severity and systemic inflammatory responses using both humoral (CAR) and cellular (NLR, SII) biomarkers. Our findings indicate that systemic inflammatory burden increases progressively with advancing OSAS severity and that CAR, NLR, and SII are independently

associated with the presence of moderate-to-severe OSAS. Notably, the superior diagnostic accuracy of CAR compared with the other markers (AUC: 0.811) supports the notion that this ratio reflects the inflammatory processes underlying OSAS in a more integrative manner.

Intermittent hypoxia, a hallmark pathophysiological feature of OSAS, induces a process resembling recurrent ischemia–reperfusion injury at the tissue level, leading to enhanced generation of reactive oxygen species.[12] This increased oxidative stress burden has been shown to shift the balance between two major transcriptional pathways in neutrophils—hypoxia-inducible factor-1α (HIF-1α) and nuclear factor kappa B (NF-κB)—toward a pro-inflammatory response. Unlike chronic sustained hypoxia, OSAS-specific intermittent hypoxia appears to induce relatively insufficient HIF-1α–mediated adaptive responses while selectively activating the NF-κB pathway. Enhanced NF-κB activation, in turn, promotes the expression of pro-inflammatory cytokines such as TNF-α, IL-6, and IL-8, thereby amplifying the systemic inflammatory response.[13] At the cellular level, intermittent hypoxia has been associated with delayed neutrophil apoptosis through preservation of antiapoptotic mechanisms (Bax/Mcl-1 balance), resulting in prolonged neutrophil survival, whereas lymphocyte responses—particularly at the level of CD4+ T cells—appear to be suppressed, leading to immune dysfunction.[14-15] The elevated NLR and SII values observed in our study may therefore reflect this cellular imbalance and the hypoxia-driven inflammatory cascade as manifested in peripheral blood parameters. Moreover, oxidative stress–induced consumption of antioxidant proteins such as albumin provides a plausible biochemical explanation for the observed increase in CAR.

Several large-scale studies have reported an association between elevated CRP levels and severe OSAS.[16] Lundetræ et al. demonstrated that severe OSAS was an independent predictor of CRP levels exceeding 3 mg/L.[17] However, CRP levels may also be influenced by comorbid conditions such as diabetes and chronic obstructive pulmonary disease, as well as by obesity, age, and sex, which limits its ability to reflect inflammatory burden and disease severity when used in isolation.[18] Consequently, interest has increasingly shifted toward composite biomarkers that allow a more comprehensive assessment of the acute-phase response. In

this context, the CAR, which integrates a positive acute-phase reactant that increases during inflammation with a negative acute-phase protein that decreases under inflammatory and oxidative stress conditions, has emerged as a more balanced and reliable index of systemic inflammation.[19] Recent studies have demonstrated associations between CAR and disease severity or prognosis in cardiovascular, infectious, and inflammatory conditions.[20] Specifically for OSAS, Hızlı et al. reported that CAR may be useful in discriminating moderate and severe disease.[21] Consistent with these findings, our multivariable and ROC analyses indicate that CAR outperforms other inflammatory markers in predicting the presence of moderate-to-severe OSAS.

The NLR represents a composite marker that simultaneously reflects two opposing cellular components of the inflammatory response. While neutrophils indicate ongoing nonspecific inflammation and innate immune activation, lymphocytes reflect adaptive and regulatory immune responses. Accordingly, NLR is considered a holistic inflammatory biomarker that captures both cellular inflammation and immune balance.[22] Elevated NLR has been associated with disease severity, prognosis, and mortality in a wide range of systemic inflammatory conditions, including cardiovascular disease, chronic kidney disease, inflammatory bowel disease, and familial Mediterranean fever.[23.] In OSAS, intermittent hypoxia, sympathetic activation, and oxidative stress resulting from repetitive apneic–hypopneic events may enhance neutrophil activation while suppressing lymphocyte function, leading to increased NLR values. Indeed, previous studies and meta-analyses have reported a stepwise increase in NLR with increasing OSAS severity, and the greatest elevations have been reported in patients with severe OSAS. [24-25]

Consistent with the existing literature, our results demonstrate increasing neutrophil counts and decreasing lymphocyte counts with advancing OSAS severity, resulting in significantly elevated NLR values, particularly in patients with moderate-to-severe OSAS. Importantly, NLR remained independently associated with moderate-to-severe OSAS even after adjustment for age, sex, and BMI, supporting its potential role as a marker of OSAS-associated cellular inflammation.

The SII is a biomarker that integrates neutrophils, platelets, and lymphocytes into a single formula, thereby providing a comprehensive assessment of inflammatory and immune status.[26] Neutrophils reflect inflammatory and oxidative stress responses, platelets play active roles in inflammation, endothelial dysfunction, and thrombosis, and lymphocytes represent adaptive and regulatory immune responses. The combined evaluation of these three cellular components renders SII a more inclusive index than individual parameters alone.[27] In OSAS, intermittent hypoxia–induced oxidative stress, sympathetic activation, and endothelial dysfunction may promote neutrophil activation and platelet reactivity while suppressing lymphocyte function, suggesting a biological basis for the relevance of SII in this condition.[28]

Although previous studies have examined the relationship between OSAS and SII, findings regarding the association between SII and disease severity have been inconsistent. While Topuz et al. reported a significant association between OSAS severity and SII, Gölen et al. failed to demonstrate such a relationship.[29-30] In the present study, SII was significantly associated with OSAS severity after adjustment for potential confounders such as age, sex, and BMI. Furthermore, the independent association of SII with moderate-to-severe OSAS in multivariable logistic regression analyses helps reconcile previous discrepancies and suggests that SII may serve as a clinically meaningful marker of systemic inflammatory burden in OSAS. Nevertheless, the limited discriminative performance of SII in ROC analyses indicates that this index may be more appropriate as an indicator of inflammatory load rather than as a stand-alone diagnostic tool.

Obesity represents one of the strongest risk factors for OSAS and is itself associated with chronic low-grade systemic inflammation. In the present study, BMI increased progressively with OSAS severity, raising the possibility that obesity may have contributed to the observed elevations in inflammatory biomarkers. To address this issue, all multivariable regression models were adjusted for BMI, which remained a significant predictor of moderate-to-severe OSAS in all models. Nevertheless, residual confounding related to obesity cannot be

completely excluded, and the observed associations should be interpreted within this context.

Several limitations should be acknowledged. The retrospective and single-center design may limit the generalizability of our findings. In addition, comorbid conditions commonly associated with OSAS, including hypertension, diabetes mellitus, cardiovascular disease, and dyslipidemia, were not systematically evaluated in the analyses. As these conditions may independently influence systemic inflammatory biomarkers, their potential confounding effects cannot be completely excluded. Furthermore, although OSAS severity was classified according to the apnea–hypopnea index (AHI), which remains the standard parameter for disease grading, other polysomnographic variables reflecting nocturnal hypoxemia and sleep fragmentation, such as the oxygen desaturation index, minimum oxygen saturation, T90, and arousal index, were not evaluated. Future studies incorporating these parameters may provide a more comprehensive assessment of the relationship between sleep-disordered breathing and systemic inflammation. Additionally, the optimal cut-off values identified for CAR, NLR, and SII were derived from a single retrospective cohort and were not externally validated. Therefore, these thresholds should be interpreted with caution and require confirmation in independent prospective cohorts before routine clinical implementation. Nevertheless, the relatively large sample size and the use of multivariable models with appropriate adjustments enhance the robustness and reliability of the results.

In summary, this study provides a comprehensive evaluation of the relationship between OSAS severity and systemic inflammation using both humoral and cellular biomarkers. Our data demonstrate that increasing OSAS severity is characterized by progressive elevations in both cellular (NLR, SII) and humoral (CAR) inflammatory burden. The identification of these markers as independent predictors of moderate-to-severe OSAS, together with the superior diagnostic performance of CAR, highlights their potential clinical utility.

Collectively, these findings support the concept that OSAS is not merely a mechanical disorder of upper airway obstruction but a multisystem disease accompanied by systemic inflammation, oxidative stress, and immune dysregulation. Given that a substantial

proportion of undiagnosed OSAS patients initially present to primary care settings, easily obtainable and low-cost inflammatory biomarkers such as CAR, NLR, and SII may provide supportive information regarding inflammatory burden and disease severity. These markers may be considered adjunctive biomarkers alongside established clinical and polysomnographic assessments. Future prospective studies are warranted to clarify their prognostic value and to determine their potential role in clinical practice.

Ethical Considerations: This study was reviewed and approved by the Kayseri City Hospital Ethics Committee (Approval No. 628, November 4, 2025). Given that data were analyzed retrospectively, individual informed consent was not obtained.

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

References

1. Liu Y, Tan H, Yu Y, Zeng Y, Xiao L. Analysis of clinical characteristics and polysomnography indicators of obstructive sleep apnea–hypopnea syndrome patients based on sleep perception types. Front Neurol 2020;11:988.
2. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MSM, Morrell MJ, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. Lancet Respir Med 2019;7(8):687-98.
3. Imani MM, Sadeghi M, Farokhzadeh F, Khazaie H, Brand S, et al. Evaluation of blood levels of C-reactive protein marker in obstructive sleep apnea: a systematic review, meta-analysis and meta-regression. Life (Basel) 2021;11(4):362.
4. Huang T, Goodman M, Wang H, Sofer T, Tworoger SS, et al. Genetic predisposition to elevated C-reactive protein and risk of obstructive sleep apnea. Am J Respir Crit Care Med 2024;209(3):329-31.
5. Li X, Hu R, Ren X, He J. Interleukin-8 concentrations in obstructive sleep apnea syndrome: a systematic review and meta-analysis. Bioengineered 2021;12(2):10666-81.
6. Algul FE, Unver FB. Relationship between trigeminal neuralgia and the C-reactive protein/albumin ratio. Turk J Neurol 2024;30(3):134-40.
7. Oh TK, Song IA, Lee JH. Clinical usefulness of C-reactive protein to albumin ratio in predicting 30-day mortality in critically ill patients: a retrospective analysis. Sci Rep 2018;8(1):14977.
8. Yıldız F, Karabuga B, Karatlı S, Yalınkılıç M, Deliktaş Onur İ, Bilgic Koylu E, et al. Clinicopathological predictors of pathological complete response in HER2-positive breast cancer treated with pertuzumab-based neoadjuvant therapy: a multicenter real-world study. Medicina 2026;62(4):763.
9. Karagoz I, Peltek Ozer S, Ozer B, Aktas G. Prognostic nutritional index could serve as a reliable prognostic marker in intensive care population. Med Sci 2025;13(2):59.

10. Fest J, Ruiter TR, Groot Koerkamp B, Rizopoulos D, Ikram MA, et al. The neutrophil-to-lymphocyte ratio is associated with mortality in the general population: the Rotterdam Study. Eur J Epidemiol 2019;34(5):463-70.
11. Pau MC, Zinellu A, Mangoni AA, Paliogiannis P, Lacana MR, et al. Evaluation of inflammation and oxidative stress markers in patients with obstructive sleep apnea (OSA). J Clin Med 2023;12(12):3935.
12. Lira AB, de Sousa Rodrigues CF. Evaluation of oxidative stress markers in obstructive sleep apnea syndrome and additional antioxidant therapy: a review article. Sleep Breath 2016;20(4):1155-60.
13. Walmsley SR, Print C, Farahi N, Peyssonnaux C, Johnson RS, et al. Hypoxia-induced neutrophil survival is mediated by HIF-1alpha-dependent NF-kappaB activity. J Exp Med 2005;201(1):105-15.
14. Dyugovskaya L, Polyakov A, Cohen-Kaplan V, Lavie P, Lavie L. Bax/Mcl-1 balance affects neutrophil survival in intermittent hypoxia and obstructive sleep apnea: effects of p38MAPK and ERK1/2 signaling. J Transl Med 2012;10:211.
15. Song R, Baker TL, Watters JJ, Kumar S. Obstructive sleep apnea–associated intermittent hypoxia–induced immune responses in males, pregnancies, and offspring. Int J Mol Sci 2024;25(3):1852.
16. Shah A, Mukherjee S, McArdle N, Singh B, Ayas N. Circulating C-reactive protein levels in patients with suspected obstructive sleep apnea. J Clin Sleep Med 2022;18(4):993-1001.
17. Lundetræ RS, Lehmann S, Saxvig IW, Saeed S, Gislason T, Bjorvatn B. Severity of obstructive sleep apnea is related to C-reactive protein levels: the influence of comorbidities and self-reported sleep duration. Sleep Med 2025;131:106529.
18. Sharma SK, Mishra HK, Sharma H, Goel A, Sreenivas V, et al. Obesity, and not obstructive sleep apnea, is responsible for increased serum hs-CRP levels in patients with sleep-disordered breathing in Delhi. Sleep Med 2008;9(2):149-56.
19. Sarıçam G, Ege F, Aslanyavrusu M. C-reactive protein/albumin ratio as a predictor of comorbidity in migraine and restless legs syndrome. Arch Basic Clin Res 2025;7(3):186-92.

20. Park JE, Chung KS, Song JH, et al. The C-reactive protein/albumin ratio as a predictor of mortality in critically ill patients. J Clin Med 2018;7(10):333.
21. Hizli O, Cayir S, Coluk Y, Kayabasi S, Yildirim G. The novel indicators of moderate to severe sleep apnea: fibrinogen to albumin ratio vs. CRP to albumin ratio. Eur Arch Otorhinolaryngol 2021;278(3):851-5.
22. Imtiaz F, Shafique K, Mirza SS, Ayoob Z, Vart P, et al. Neutrophil lymphocyte ratio as a measure of systemic inflammation in prevalent chronic diseases in Asian population. Int Arch Med 2012; 5:2.
23. Altintas N, Çetinoğlu E, Yuceege M, Acet AN, Ursavas A, Firat H, et al. Neutrophil-to-lymphocyte ratio in obstructive sleep apnea; a multi center, retrospective study. Eur Rev Med Pharmacol Sci 2015;19(17):3234-40.
24. Rha MS, Lee JE, Kim JH, Cho JH, Lee SH, et al. Association between the neutrophil-to-lymphocyte ratio and obstructive sleep apnea: a meta-analysis. Sci Rep 2020;10(1):10862.
25. Uygur F, Tanriverdi H, Aktop Z, Erboy F, Altinsoy B, et al. The neutrophil-to-lymphocyte ratio in patients with obstructive sleep apnoea syndrome and its relationship with cardiovascular disease. Heart Lung 2016;45(2):121-5.
26. Peng X, Wang X, Hua L, Yang R. Prognostic and clinical value of the systemic immune-inflammation index in biliary tract cancer: a meta-analysis. J Immunol Res 2022; 2022:6988489.
27. Ye Z, Hu T, Wang J, Xiao R, Liao X, et al. Systemic immune-inflammation index as a potential biomarker of cardiovascular diseases: a systematic review and meta-analysis. Front Cardiovasc Med 2022; 9:933913.
28. Güneş ZY, Günaydın FM. The relationship between the systemic immune inflammation index and obstructive sleep apnea. Sleep Breath 2024; 28:311-7.
29. Topuz MF, Ture N, Akdag G, Arik O, Gulhan PY. The importance of systemic immune-inflammation index in obstructive sleep apnea syndrome. Eur Arch Otorhinolaryngol 2022;279(10):5033-8.

30. Gölen MK, Işık ŞM, Arıkan V. Is there a relationship between the severity of obstructive sleep apnea syndrome and the systemic immune inflammation index? Eur Arch Otorhinolaryngol 2024;281(9):5007-13.

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):500-514 // 10.5505/amj.2026.64933

ELEVATED SYSTEMIC IMMUNE-INFLAMMATION INDEX AND ITS CORRELATION WITH TSH LEVELS IN PATIENTS WITH SUBCLINICAL HYPOTHYROIDISM

Mustafa Onmaz[1], Burak Altındağ[1], Ahmet Emre Hatır[1]

[1]Department of Family Medicine, Bandırma Onyedi Eylul University, Balıkesir, Türkiye

Correspondence:
Mustafa Onmaz (e-mail: monmaz@bandirma.edu.tr)

Submitted: 08.03.2026 // Accepted: 24.06.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: To investigate whether hemogram-derived inflammatory indices differ between patients with subclinical hypothyroidism and euthyroid controls and whether these indices are associated with thyroid-stimulating hormone (TSH) levels.

Materials and Methods: In this retrospective cross-sectional study, 300 adults were included: 100 euthyroid individuals and 200 patients with subclinical hypothyroidism. Neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, monocyte-to-lymphocyte ratio, and systemic immune-inflammation index (SII) were calculated from complete blood count parameters. Correlations between TSH and laboratory variables were assessed in each group, and multivariable logistic regression identified variables independently associated with subclinical hypothyroidism.

Results: SII values were significantly higher in the subclinical hypothyroidism group than in the euthyroid group (p=0.047). However, the between-group difference in SII showed a small effect size (Cohen's d=0.29). In the subclinical hypothyroidism group, TSH was positively correlated with neutrophil count (r=0.201, p=0.045), creatinine (r=0.188, p=0.046), NLR (r=0.310, p=0.002), MLR (r=0.239, p=0.017), and SII (r=0.329, p<0.001). Older age (OR 1.058, 95% CI 1.035–1.082; p<0.001), female sex (OR 1.940, 95% CI 1.125–3.345; p=0.017), and higher body mass index (OR 1.055, 95% CI 1.006–1.107; p=0.048) were independently associated with subclinical hypothyroidism, whereas SII was not (OR 1.001, 95% CI 0.999–1.002; p=0.084).

Conclusion: Subclinical hypothyroidism was associated with modestly higher SII values and positive correlations between TSH and selected hemogram-derived inflammatory indices. However, the small effect size of SII and its lack of independent significance after adjustment suggest that SII should be interpreted as a supportive and nonspecific marker of low-grade inflammation rather than an independent indicator of subclinical hypothyroidism.

Keywords: Subclinical hypothyroidism, thyrotropin, inflammation, blood cell count.

Introduction

Subclinical hypothyroidism is characterized by an increase in serum thyroid-stimulating hormone (TSH) concentrations while free thyroxine (fT4) levels remain within the reference range. It is frequently encountered in primary care settings and is often identified incidentally, as many affected individuals do not exhibit clear or specific clinical manifestations.[1] Although it is generally considered a mild form of thyroid dysfunction, accumulating evidence indicates that subclinical hypothyroidism may be linked to adverse metabolic changes, increased cardiovascular risk, and persistent low-grade inflammation.[2,3]

Interest in the inflammatory background of subclinical hypothyroidism has increased in recent years. Therefore, simple and readily obtainable biomarkers of inflammation have gained attention.[4–6] These include the neutrophil-to-lymphocyte ratio (NLR; NEU/LYM), platelet-to-lymphocyte ratio (PLR; PLT/LYM), monocyte-to-lymphocyte ratio (MLR; MONO/LYM), and systemic immune-inflammation index (SII; [PLT × NEU]/LYM), all of which are derived from routine complete blood count parameters and have been suggested as practical surrogate markers of systemic inflammation in endocrine and metabolic disorders.[7–9]

Previous studies have demonstrated the clinical relevance of these hematological inflammatory markers in conditions such as diabetes mellitus, overt thyroid disorders, and various metabolic abnormalities.[5,7–10] However, their potential role in subclinical hypothyroidism has not yet been fully clarified. Therefore, this study aimed to determine whether hemogram-derived inflammatory indices, including NLR, PLR, MLR, and SII, are altered in patients with subclinical hypothyroidism compared with euthyroid controls and whether these indices are associated with TSH levels.

Materials and Methods

This study was designed as a retrospective observational study evaluating laboratory data obtained during routine health check-ups of individuals who attended the Family Medicine

Outpatient Clinic, Bandırma Training and Research Hospital between January 1, 2025 and December 31, 2025. Data were extracted from the hospital information management system.

Study population

Participants were assigned to two groups on the basis of thyroid function test results. According to the recommendations of the Turkish Society of Endocrinology and Metabolism (TEMD), subclinical hypothyroidism was defined as a serum TSH concentration greater than 4.0 mIU/L with an fT4 value remaining within the laboratory reference interval, whereas euthyroidism was defined as a TSH level between 0.5 and 4.0 mIU/L accompanied by a normal fT4 level.[6] In our laboratory, the reference interval was 0.5–4.0 mIU/L for TSH and 0.80–1.80 ng/dL for fT4. Individuals under 18 years of age were excluded. To limit potential confounding effects on hemogram-derived inflammatory indices, participants with acute infection, chronic inflammatory or autoimmune disease, malignancy, hematologic disorder, pregnancy, ongoing corticosteroid or immunosuppressive therapy, previously diagnosed overt hypothyroidism or hyperthyroidism, current thyroid-specific treatment, or non-fasting blood samples were also excluded. Records lacking data required for group allocation or statistical analyses were omitted, and no imputation procedure was applied. The sample size was determined based on an a priori power analysis performed using G*Power software version 3.1.9.7, together with the availability of eligible records during the study period. Assuming a two-sided alpha level of 0.05, 80% statistical power, a small-to-moderate between-group effect size, and a 2:1 allocation ratio, a total sample size of 300 participants was considered sufficient. Accordingly, medical records were reviewed until 200 patients with subclinical hypothyroidism and 100 euthyroid controls meeting the inclusion and exclusion criteria were included in the final analysis.

Laboratory parameters and inflammatory indices

Laboratory measurements were obtained retrospectively from the hospital electronic database. To improve standardization, only fasting blood samples collected during the morning period (08:00–12:00) were included. Evaluated hematologic parameters

comprised hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, mean platelet volume, white blood cell count, neutrophil count, lymphocyte count, monocyte count, and platelet count. Biochemical variables included creatinine, total cholesterol, triglycerides, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, TSH, and fT4. Hemogram-derived inflammatory indices were calculated as follows: NLR = NEU/LYM, PLR = PLT/LYM, MLR = MONO/LYM, and SII = (PLT × NEU)/LYM.[7–10]

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 21.0 (IBM Corp., Armonk, NY, USA). The distribution of continuous variables was assessed using the Shapiro–Wilk test and visual inspection of histograms. Continuous variables are presented as mean ± standard deviation, whereas categorical variables are summarized as numbers and percentages. Between-group comparisons were performed using the independent-samples t test for continuous variables and the chi-square test for categorical variables. For independent-samples t tests, absolute test statistics (t-values) were reported, and Cohen's d was calculated to assess the magnitude of between-group differences—with a specific focus on the systemic immune-inflammation index (SII)—to support the interpretation of clinical relevance. Within each group, the associations of TSH with hematological, biochemical, and hemogram-derived inflammatory variables were evaluated using Spearman correlation analysis.

A multivariable logistic regression model was constructed to determine whether SII was independently associated with subclinical hypothyroidism after adjustment for potential confounders. Age, sex, and body mass index were included in the model as covariates because they are clinically relevant variables that may influence both thyroid function and baseline inflammatory status. SII was selected as the specific hemogram-derived inflammatory index for the regression analysis because it was the only index demonstrating a significant between-group difference and the strongest correlation with TSH in preliminary analyses. NLR, PLR, and MLR were deliberately excluded from the multivariable

model to avoid potential multicollinearity arising from mathematical overlap and shared hematological components. Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated, and a two-sided p-value <0.05 was considered statistically significant.

Results

A total of 300 adults were included, comprising 100 euthyroid individuals and 200 patients with subclinical hypothyroidism. Age (p=0.318), sex distribution (p=0.742), BMI (p=0.067), systolic blood pressure (p=0.184), and diastolic blood pressure (p=0.713) were similar between the groups (Table 1).

Table 1. Comparison of baseline demographic and clinical characteristics between euthyroid and subclinical hypothyroidism groups

Variable	Euthyroid	Subclinical	p-value
Age (years)	50.10 ± 12.05	51.55 ± 11.40	0.318
Sex, n (%)			0.742
-Female	53 (53.0%)	110 (55.0%)	
-Male	47 (47.0%)	90 (45.0%)	
BMI (kg/m^2)	26.0 ± 3.6	27.0 ± 4.2	0.067
Systolic BP (mmHg)	123.16 ± 16.44	125.94 ± 11.74	0.184
Diastolic BP (mmHg)	76.8 ± 9.6	77.2 ± 9.4	0.713

Data are presented as mean ± standard deviation or n (%). An independent-samples t-test was used for continuous variables. Chi-square test was used for categorical variables. BMI, Body Mass Index.

When hematological, biochemical, and inflammatory parameters were compared, TSH levels were significantly higher in the subclinical hypothyroidism group than in the euthyroid group (t=20.76, p=0.008). Among the hemogram-derived inflammatory indices, only SII showed a statistically significant between-group difference, with higher values observed in the subclinical hypothyroidism group (t=2.36, p=0.047). However, the magnitude of this difference was small, as indicated by a Cohen's d value of 0.29 (Table 2).

Table 2. Comparison of hematological parameters and inflammatory indices between euthyroid and subclinical hypothyroidism groups

Variable	Euthyroid	Subclinical hypothyroidism	t-value	p-value
	mean ± SD	mean ± SD		
HGB (g/dL)	15.12 ± 0.81	15.05 ± 1.02	0.60	0.645
HCT (%)	45.10 ± 3.85	45.74 ± 4.28	1.26	0.214
MCV (fL)	86.70 ± 4.80	87.10 ± 4.20	1.69	0.092
MCH (pg)	29.60 ± 1.70	29.50 ± 1.80	1.26	0.210
MCHC (g/dL)	33.10 ± 1.10	32.90 ± 1.00	1.44	0.150
MPV (fL)	9.70 ± 1.10	9.60 ± 1.10	1.34	0.180
WBC (10^3/µL)	7.03 ± 2.18	7.46 ± 2.26	1.57	0.094
NEU (10^3/µL)	4.31 ± 1.84	4.84 ± 1.92	1.88	0.061
LYM (10^3/µL)	2.06 ± 0.47	2.22 ± 0.51	2.63	0.067
MONO (10^3/µL)	0.47 ± 0.15	0.58 ± 0.18	5.26	0.074
PLT (10^3/µL)	266.40 ± 88.70	272.10 ± 91.20	0.52	0.212
Creatinine (mg/dL)	0.94 ± 0.12	0.95 ± 0.13	0.64	0.405
TC (mg/dL)	191.85 ± 34.10	198.72 ± 36.45	1.57	0.108
TG-C (mg/dL)	158.40 ± 54.80	167.72 ± 58.60	1.33	0.176
LDL-C (mg/dL)	113.20 ± 24.60	118.66 ± 27.10	1.70	0.081
HDL-C (mg/dL)	50.42 ± 11.84	48.36 ± 11.62	1.44	0.149
TSH (mIU/L)	2.18 ± 0.81	5.36 ± 1.42	20.76	**0.008**
fT4 (ng/dL)	1.15 ± 0.14	1.13 ± 0.15	1.11	0.214
NLR	2.09 ± 0.88	2.18 ± 0.91	0.82	0.082
PLR	129.41 ± 41.76	122.58 ± 39.84	1.38	0.168
MLR	0.23 ± 0.08	0.26 ± 0.09	2.82	0.058
SII	536.84 ± 214.37	603.72 ± 238.95	2.36	**0.047**

Variables were compared using the independent-samples t-test. NLR = NEU / LYM; PLR = PLT / LYM; MLR = MONO / LYM; SII = (PLT × NEU) / LYM.

Spearman correlation analysis showed no significant correlations between TSH levels and the evaluated hematological, biochemical, or inflammatory parameters in the euthyroid group (all $p>0.05$). In the subclinical hypothyroidism group, TSH levels were positively correlated with neutrophil count ($r=0.201$, $p=0.045$), creatinine ($r=0.188$, $p=0.046$), NLR ($r=0.310$, $p=0.002$), MLR ($r=0.239$, $p=0.017$), and SII ($r=0.329$, $p<0.001$). (Table 3).

Multivariable logistic regression analysis showed that older age, female sex, and higher BMI were independently associated with subclinical hypothyroidism, whereas SII did not reach statistical significance as an independent determinant in the adjusted model (age: OR 1.058, 95% CI 1.035–1.082, $p<0.001$; female sex: OR 1.940, 95% CI 1.125–3.345, $p=0.017$; BMI: OR 1.055, 95% CI 1.006–1.107, $p=0.048$; SII: OR 1.001, 95% CI 0.999–1.002, $p=0.084$) (Table 4).

Table 3. Correlation of TSH levels with hematological parameters and inflammatory indices in euthyroid and subclinical hypothyroidism groups

	Euthyroid		Subclinical hypothyroidism	
	r	**p**	**r**	**p**
HGB	0.029	0.778	-0.045	0.528
HCT	0.057	0.575	-0.091	0.202
MCV	0.035	0.731	-0.069	0.332
MCH	-0.012	0.906	0.003	0.968
MCHC	-0.043	0.674	-0.027	0.709
MPV	0.062	0.537	0.099	0.161
WBC	0.014	0.849	0.145	0.150
NEU	0.010	0.892	0.201	**0.045**
LYM	0.020	0.782	0.098	0.333
MONO	0.022	0.762	0.132	0.062
PLT	0.118	0.421	0.031	0.663
Creatinine	0.032	0.654	0.188	**0.046**
TC	0.086	0.395	0.091	0.201
TG -C	0.080	0.427	0.067	0.346
LDL-C	0.103	0.306	0.100	0.160
HDL-C	-0.001	0.990	-0.021	0.768
NLR	-0.053	0.459	0.310	**0.002**
PLR	0.010	0.886	0.186	0.064
MLR	-0.006	0.928	0.239	**0.017**
SII	-0.061	0.392	0.329	**<0.001**

Correlations were analyzed using Spearman's correlation test.

Table 4. Multivariable Logistic Regression Analysis for Subclinical Hypothyroidism

Variables	OR (95%CI)	p-value
Age	1.058 (1.035–1.082)	**<0.001**
Female sex	1.940 (1.125–3.345)	**0.017**
BMI	1.055 (1.006–1.107)	**0.048**
SII	1.001 (0.999–1.002)	0.084

OR, odds ratio; CI, confidence interval; BMI, Body Mass Index.

Discussion

In this study, we aimed to evaluate the levels of hemogram-derived inflammatory indices in patients with subclinical hypothyroidism and their associations with basic clinical and biochemical parameters. As expected, the coexistence of elevated TSH and preserved fT4 levels in the subclinical hypothyroidism group reflects the biochemical nature of this clinical condition, supporting the appropriateness of our case classification. The similarity between the groups in terms of baseline clinical characteristics, such as age, sex distribution, body mass index (BMI), and blood pressure values, allows the observed differences in inflammatory indices to be interpreted with reduced concern regarding major baseline demographic imbalances.

One of the main findings of our study was that the SII was significantly higher in the subclinical hypothyroidism group than in the euthyroid group and showed a positive correlation with TSH levels. Previous studies and meta-analytic evidence have similarly linked hypothyroid states to low-grade systemic inflammation and oxidative stress.[4,11,12] Our findings support this association, suggesting that subclinical hypothyroidism may be accompanied by a mild inflammatory profile; however, this inflammatory signal appears to be modest and is not reflected uniformly across all hemogram-derived indices.

This heterogeneous behavior of inflammatory indices is consistent with the complex, multifactorial nature of subclinical hypothyroidism, which is frequently associated with

autoimmune thyroiditis and influenced by demographic or metabolic factors.[13,14] Prior literature demonstrates that indices like the Neutrophil-to-Lymphocyte Ratio (NLR) and Platelet-to-Lymphocyte Ratio (PLR) can show highly variable patterns, often reflecting nonspecific inflammation related to underlying thyroid autoimmunity rather than thyroid hormone status alone.[15–17] Therefore, the absence of statistically significant differences for NLR, PLR, and Monocyte-to-Lymphocyte Ratio (MLR) in our study does not necessarily indicate the absence of inflammation; rather, it highlights that these specific markers are highly sensitive to disease severity, metabolic profiles, and patient-specific characteristics.

Regarding biochemical parameters, creatinine levels showed a weak positive correlation with TSH levels in the subclinical hypothyroidism group, potentially reflecting subtle changes in renal hemodynamics associated with mild thyroid dysfunction.[18] Similarly, lipid parameters did not differ significantly between the groups, although LDL-cholesterol levels tended to be higher in the subclinical hypothyroidism group, indicating a relatively mild and variable metabolic risk profile in our study population.[14]

In the multivariable logistic regression analysis, older age, female sex, and higher BMI were independently associated with subclinical hypothyroidism, whereas SII did not retain statistical significance after adjustment. These demographic and anthropometric variables were selected as covariates because they are clinically relevant potential confounders that may influence both thyroid function and low-grade inflammatory status. SII was included in the adjusted model because it was the only index showing a significant between-group difference and the strongest correlation with TSH, while NLR, PLR, and MLR were excluded to reduce the risk of multicollinearity arising from shared hematological components. Therefore, the final model was constructed to provide a clinically justified and statistically cautious assessment, demonstrating that the unadjusted difference observed for SII should not be interpreted as an independent diagnostic effect.

The small effect size observed for SII (Cohen's d = 0.29) further supports this cautious interpretation. From a clinical perspective, SII should be viewed as an adjunctive, nonspecific inflammatory marker rather than an independent diagnostic indicator. Its practical value is

limited to providing complementary information regarding the overall inflammatory status in selected patients, and it should not replace standard thyroid function tests or thyroid autoimmunity assessments.

Several limitations should be considered when interpreting our findings. First, the retrospective and cross-sectional design of the study precludes causal inference; thus, we cannot determine whether subclinical hypothyroidism directly drives systemic inflammation or vice versa. Second, thyroid autoimmunity markers, including anti-TPO and anti-Tg, were not available in our dataset. This is a crucial limitation because autoimmune thyroiditis is a common cause of subclinical hypothyroidism and a well-known contributor to chronic low-grade inflammation. Consequently, we could not isolate whether the elevated SII reflected thyroid hormone imbalance, underlying autoimmune activity, or both. Third, despite strict exclusion criteria for acute infections and overt inflammatory diseases, the potential effects of unmeasured subclinical metabolic abnormalities or residual confounding cannot be entirely ruled out.

Nevertheless, the study is strengthened by its relatively large sample size, the inclusion of a euthyroid comparison group, and the simultaneous evaluation of multiple hemogram-derived inflammatory indices. The post hoc sensitivity power analysis showed that the final sample size of 300 participants had 80% power to detect a minimum effect size of Cohen's d = 0.34. As the observed effect size for SII was smaller, its borderline statistical significance should be interpreted cautiously and should not be considered evidence of a strong or independent clinical effect.

In conclusion, our findings demonstrate that SII is higher in patients with subclinical hypothyroidism and correlates positively with TSH levels. However, this association is characterized by a small effect size and loses independent significance after adjustment for age, sex, and BMI. These results suggest that SII may reflect a mild, low-grade inflammatory burden accompanying subclinical hypothyroidism, but its utility as an independent clinical marker is limited. Prospective studies incorporating thyroid autoimmunity markers are required to further clarify these relationships.

ANKARA MEDICAL JOURNAL

Ethical Considerations: Ethical approval was obtained from the Non-Interventional Research Ethics Committee of Bandırma Onyedi Eylül University Faculty of Medicine (Meeting No: 2026-01, File No: 2026-01-08, 13 January 2026). The study was carried out in accordance with the Declaration of Helsinki.

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

References

1. Biondi B, Cappola AR, Cooper DS. Subclinical hypothyroidism: a review. *JAMA*. 2019;322(2):153-60. doi:10.1001/jama.2019.9052
2. Moon S, Kim MJ, Yu JM, Yoo HJ, Park YJ. Subclinical hypothyroidism and the risk of cardiovascular disease and all-cause mortality: a meta-analysis of prospective cohort studies. *Thyroid.* 2018;28(9):1101-10. doi:10.1089/thy.2017.0414
3. Tezcan D, Onmaz DE, Sivrikaya A, et al. Assessment of serum neopterin and calprotectin as biomarkers for subclinical inflammation in patients with familial Mediterranean fever. Ir J Med Sci. 2023;192(4): 2015-22. doi: 10.1007/s11845-022-03173-w.
4. Tellechea ML. Meta-analytic evidence for increased low-grade systemic inflammation and oxidative stress in hypothyroid patients. Can levothyroxine replacement therapy mitigate the burden? *Endocrine.* 2021;72(1):62-71. doi:10.1007/s12020-020-02484-1
5. Vural S, Muhtaroğlu A, Güngör M. Systemic immune-inflammation index: a new marker in differentiation of different thyroid diseases. *Medicine (Baltimore).* 2023;102(31):e34596. doi:10.1097/MD.0000000000034596
6. Öz Gül Ö, Cander S, Ersoy C. Subklinik hipotiroidi nedir? Tedavi edilmeli midir? *Uludag Univ Tip Fak Derg.* 2018;44(2):129-34. doi:10.32708/uutfd.435351
7. Mertoglu C, Gunay M. Neutrophil-lymphocyte ratio and platelet-lymphocyte ratio as useful predictive markers of prediabetes and diabetes mellitus. *Diabetes Metab Syndr.* 2017;11(Suppl 1):S127-31. doi: 10.1016/j.dsx.2016.12.021.
8. Chen H, Ju JQ, Qian XW, Zhu ZT, Zhao CZ, Liu Z. The correlation between thyroid parameters and the ratios of neutrophil/lymphocyte and platelet/lymphocyte in euthyroid type 2 diabetic patients. *Diabetes Metab Syndr Obes.* 2023;16:3763-71. doi:10.2147/DMSO.S431941
9. Yurtdaş M, Asoğlu R, Özaydın A, Doğan Z. Lymphocyte to monocyte ratio may predict increased carotid intima-media thickness in patients with subclinical

hypothyroidism. *Clin Hemorheol Microcirc*. 2021;77(2):133-42. doi:10.3233/CH-200820

10. Hu B, Yang XR, Xu Y, et al. Systemic immune-inflammation index predicts prognosis of patients after curative resection for hepatocellular carcinoma. *Clin Cancer Res*. 2014;20(23):6212-22. doi:10.1158/1078-0432.CCR-14-0442
11. Erinc O, Yesilyurt S, Senat A. A comprehensive evaluation of hemogram-derived inflammatory indices in Hashimoto thyroiditis and non-immunogenic hypothyroidism. *Acta Endocrinol (Buchar).* 2023;19(4):435-40. doi:10.4183/aeb.2023.435
12. Gupta G, Sharma P, Kumar P, Itagappa M. Study on subclinical hypothyroidism and its association with various inflammatory markers. *J Clin Diagn Res*. 2015;9(11):BC04-6. doi:10.7860/JCDR/2015/14640.6806
13. Gosi SKY, Kaur J, Garla VV. Subclinical Hypothyroidism. [Updated 2024 Feb 15]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https: //www.ncbi.nlm.nih.gov/books/NBK536970/. (Accessed: 07.04.2026).
14. Jiang L, Du J, Wu W, Fang J, Wang J, Ding J. Sex differences in subclinical hypothyroidism and associations with metabolic risk factors: a health examination-based study in mainland China. *BMC Endocr Disord*. 2020;20(1):100. doi:10.1186/s12902-020-00586-5
15. Onalan E, Dönder E. Neutrophil and platelet to lymphocyte ratio in patients with hypothyroid Hashimoto's thyroiditis. *Acta Biomed*. 2020;91(2):310-4. doi:10.23750/abm.v91i2.8592
16. Bilge M, Yesilova A, Adas M, Helvaci A. Neutrophil- and platelet-to-lymphocyte ratio in patients with euthyroid Hashimoto's thyroiditis. *Exp Clin Endocrinol Diabetes.* 2019;127(8):545-9. doi: 10.1055/a-0723-3441.
17. Keskin H, Kaya Y, Cadirci K, Kucur C, Ziypak E, Simsek E, et al. Elevated neutrophil-lymphocyte ratio in patients with euthyroid chronic autoimmune thyreotidis. *Endocr Regul.* 2016;50(3):148-53. doi:10.1515/enr-2016-0017

18. den Hollander JG, Wulkan RW, Mantel MJ, Berghout A. Correlation between severity of thyroid dysfunction and renal function. *Clin Endocrinol (Oxf).* 2005;62(4):423-7. doi:10.1111/j.1365-2265.2005.02236.x

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):515-532 // 10.5505/amj.2026.65049

GENERATIVE AI AND LEXICAL SHIFTS IN TURKISH MEDICAL PUBLISHING (2010–2025)

Muhammed İnan[1], Cenk Aypak[1]

[1]Department of Family Medicine, University of Health Sciences, Ankara Etlik City Hospital, Ankara, Türkiye

Correspondence:
Muhammed İnan (e-mail: muhammedinan.md@gmail.com)

Submitted: 18.03.2026 //Accepted: 21.08.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

Abstract

Objectives: The adoption of Large Language Models (LLMs) raises concerns about scientific writing integrity, affecting clinical guidelines and patient safety. While studies track AI-generated text in English databases, impacts on bilingual ecosystems remain underexplored. This study investigates the "Linguistic Equalizer" hypothesis by analyzing lexical shifts in TR Dizin, determining if LLMs are primarily utilized for content generation or linguistic refinement.

Material and Methods: We conducted a longitudinal bibliometric analysis of 591,819 abstracts (2010–2025), divided into "General" and "Medical" datasets. We monitored five LLM marker word groups in English and their Turkish equivalents. We evaluated statistical significance via a large-sample pooled two-proportion test (primary analysis, using risk ratios [RR] with 95% CI), a yearly-aggregate Mann-Whitney U test (secondary/directional), and segmented regression analysis to identify post-2022 trend slopes.

Results: A divergence emerged. In the General Dataset, English markers exhibited a marked increase ("underscore" rose from 0.02% to 1.80% of abstracts; RR=78.73, 95% CI 64.49–96.10), while the Turkish equivalent ("vurgula") showed a more modest, though substantial, increase (RR=2.54, 95% CI 2.47–2.61). In the Medical Dataset, English "underscore" rose more sharply (RR=87.98, 95% CI 62.51–123.84), while Turkish usage showed a moderate increase (RR=2.02, 95% CI 1.90–2.15).

Conclusion: The disproportionate rise in English AI markers alongside comparatively stable Turkish vocabulary is consistent with the possibility that researchers utilize LLMs primarily as translation or editing tools. These findings lend empirical support to GenAI's potential role as a 'Linguistic Equalizer,' consistent with strategic adaptation to global publishing pressures rather than content fabrication.

Keywords: Research integrity, artificial intelligence, bibliometrics, authorship, linguistic Equity

Introduction

The public release of ChatGPT in November 2022 transformed scientific writing, offering new drafting possibilities while raising ethical concerns.[1,2] Large Language Models (LLMs) also exhibit linguistic patterns, including overrepresented terms such as “delve” and “underscore.”[3-5]

This shift impacts non-native English-speaking (NNES) researchers. AI text detectors exhibit false-positive biases against NNES authors.[6] LLMs may function as “linguistic equalizers” for NNES researchers seeking to meet international publication standards.[7-10]

However, AI integration is not without risk: unverified AI-generated content—such as hallucinated data or inaccurate clinical narratives—could compromise the reliability of studies that inform clinical guidelines, warranting editorial scrutiny. Concerns over homogenized scientific voices, "tortured phrases," and peer review integrity have prompted publishers to revise authorship guidelines.[11,12] AI-generated text has infiltrated sensitive areas, including peer review reports[13] and residency personal statements,[14] blurring authentic and synthetic authorship. While global studies have analyzed databases like PubMed, Web of Science, and Dimensions,[3,15,16] there is a scarcity of longitudinal data focusing on national citation indices in non-Western contexts, where the incentive to use AI for translation and polishing is stronger.

This study addresses this gap through a bibliometric analysis of TR Dizin, Türkiye's national citation index. Using over 590,000 abstracts from 2010 to 2025, we examined whether five semantic marker groups reveal a structural lexical shift following LLM adoption, with medical sciences as a case study.

Materials and Methods

This study is a bibliometric analysis based on publicly available bibliographic data retrieved from the TR Dizin database (https://trdizin.gov.tr/) on August 6, 2026. As the research did not involve human participants, animal subjects, or the use of private or confidential patient data, ethics committee approval was not required.

Records were retrieved using field-restricted Boolean queries applied to abstracts. Each of the five marker groups (Table 1) was searched separately in English and Turkish within the General and Medical datasets. The Medical subset was defined using the “Subject: Medicine” filter. As the Medical Dataset is a subset of the General Dataset, between-dataset comparisons are descriptive, not statistically independent. Exact TR Dizin search queries are reported in Supplementary Table S1.

The study defined two distinct datasets for analysis: a 'General Dataset' encompassing all academic disciplines (591,819 abstracts, 2010–2025) and a domain-specific 'Medical Dataset' restricted to health sciences (197,232 abstracts, 2010–2025). Because 2026 indexing was incomplete at data retrieval, 2026 data are presented descriptively but excluded from primary statistical comparisons.

Five LLM-associated marker groups were examined using English terms and Turkish equivalents. For each English marker, four commercial LLMs were independently prompted to provide the most natural Turkish academic-register equivalent. Terms supported by at least three models formed the primary narrow/consensus set; when agreement was low (≤2/4), the authors selected the most contextually appropriate term from the LLM-generated pool. A broad sensitivity set comprising the union of all unique terms proposed by the four models was also constructed to assess whether findings depended on specific term selection. The five terms with the largest effects were subsequently re-tested in separate context-free sessions, with all models converging on the same Turkish roots/stems (Supplementary Table S2).

Table 1. Classification of LLM Marker Words, Search Queries, and Literature Basis

Marker Group	English terms	Turkish – Narrow/Consensus (primary)	Turkish – Broad/All-LLM (sensitivity)	Rationale and Literature Findings
Group 1: Deep Investigation	Delve, Delves, Delving, Deep dive	Derinlemesine	Derinlemesine, Derinlemesine incele*, Mercek tut*	The verb "delve" is identified as the single strongest signal of LLM usage post-2023. [3,4]
Group 2: Emphasis	Underscore, Underscoring, Underscores	Vurgula*	Vurgula*, Altını çiz*	LLMs persistently prefer "underscore" over standard alternatives like "highlight". [3,4]
Group 3: Exaggerated Importance	Crucial, Pivotal, Imperative, Indispensable, Paramount	Kritik, Kilit, Zorunlu, Vazgeçilmez, En önemli	Kritik, Hayati, Kilit, Kilit rol, Zorunlu, Elzem, Vazgeçilmez, En önemli, Son derece önemli, Büyük önem taşıyan	Generative models tend to inflate the significance of studies using adjectives like "pivotal" and "crucial". [3,9]
Group 4: Complexity & Scope	Intricate, Meticulous, Comprehensive, Multifaceted, Nuanced	Karmaşık, Titiz, Kapsamlı, Çok yönlü, İncelikli	Karmaşık, Titiz, Kapsamlı, Çok yönlü, Çok boyutlu, İncelikli, Nüanslı, Ayrıntılı, İnce ayrıntılı	Adjectives such as "intricate" and "meticulous" are favored by LLMs to describe methodologies. [3,4,9]
Group 5: Future Potential	Unleash, Unlock, Potential, Paradigm shift, Groundbreaking, Transformative, Pave the way.	Ortaya çıkar*, Potansiyel, Paradigma değişimi, Çığır açan, Dönüştürücü, Yol aç*	Ortaya çıkar*, Açığa çıkar*, Aydınlat*, Potansiyel*, Paradigma değişimi, Yeni bir dönem, Çığır açan, Çığır açıcı, Dönüştürücü, Yol aç*, Zemin hazırla*, Önünü aç*	"Potential" exhibits one of the largest frequency gaps between human and AI text. [3,6]

Note: The asterisk (*) denotes a wildcard capturing morphological variants (e.g., "vurgula*" captures "vurgulamaktadır," "vurgular"). The Narrow/Consensus column lists terms on which a majority (≥3/4) of four independently queried LLMs agreed and constitutes the primary Turkish query set; derived from majority LLM agreement (≥3/4) or manual author selection from the generated pool in cases of low agreement (≤2/4). The Broad/All-LLM column lists the full union of terms suggested by any of the four models and constitutes the sensitivity-analysis query set. Reported statistics in Results reflect the full OR-aggregated group for each set.

Each query returned the number of unique abstracts containing at least one queried term; duplicate matches within a query were handled natively by TR Dizin. A stratified random sample of 100 abstracts was manually reviewed, confirming substantive use of the target terms in all cases (Supplementary Table S3). Abstracts could contribute to multiple marker groups, with each group independently normalized against the same annual denominator.

TR Dizin conventionally indexes both a Turkish-language and an English-language abstract for the majority of articles, irrespective of the article's primary language of publication — a bilingual-abstract convention common to Turkish national journals. Consequently, this study does not distinguish between English-medium and Turkish-medium articles; rather, English-language queries were matched against the English abstract field and Turkish-language queries against the Turkish abstract field of the same underlying corpus, with both normalized against the same total-article denominator for each year (Section 2.1). As a result, the English- and Turkish-language marker prevalence estimates are not drawn from independent article samples but instead reflect two language-field views of an overlapping set of articles.

Statistical Analysis

Data were processed and analyzed using Python (v3.10) with pandas for data manipulation and SciPy for statistical testing. To account for the varying number of abstracts published annually, raw word counts were normalized by the total number of abstracts in the respective dataset (General or Medical) to calculate annual prevalence:

Prevalence % = (Word Count / Total Abstracts) × 100

The primary Post-LLM period was 2023–2025; 2026 was excluded because of incomplete indexing and included only in sensitivity analyses. Annual prevalence values were compared using a one-tailed Mann–Whitney U test, reported as a directional indicator because only three post-LLM years were available. The primary statistical analysis was a pooled two-

proportion z-test based on abstract-level counts, leveraging the full corpus. Risk ratios (RRs) with 95% confidence intervals were the primary effect-size measures, and Holm–Bonferroni correction was applied across 15 comparisons per dataset. Annual counts, prevalence estimates, and 95% Wilson confidence intervals are reported in Supplementary Tables S4 and S5.

Linguistic shift was quantified using Fold Change, the ratio of mean prevalence in the Post-LLM to Pre-LLM era, with significance set at $p < 0.05$. Given the right-skewed, non-normal distribution of lexical prevalence across abstracts, the non-parametric Mann–Whitney U test was used instead of parametric alternatives to avoid inflated Type I error rates. To assess differences independent of sample size, effect sizes were calculated using the standard formula $r=Z/\sqrt{N}$. In addition, segmented regression analysis was conducted to identify and isolate the structural break associated with the late-2022 release of ChatGPT, enabling a comparison of pre-2022 and post-2022 trend slopes (annual percentage change) rather than assuming a uniform linear trajectory. All results were visualized using Seaborn and Matplotlib to illustrate longitudinal patterns.

95% confidence intervals for regression slopes are reported where at least three data points were available per segment; post-2022 slope CIs could be estimated for the primary analysis (n=3 post-LLM years, one residual degree of freedom), though estimates remain imprecise given the small sample. As a sensitivity check, all analyses were repeated with 2026 included; the pooled test was materially unaffected (e.g., Medical 'vurgula': RR=2.02 primary vs. RR=2.10 with 2026 included), while the yearly-aggregate test was more sensitive to this choice (e.g., Medical 'vurgula' significance shifted from p=0.095 to p=0.030), further supporting the pooled test as the primary basis for our conclusions (Supplementary Table S6a). Additionally, analyses were performed separately for English, narrow Turkish, and broad Turkish term sets. The broad-to-narrow fold-change ratio was used to assess sensitivity to Turkish term selection (Supplementary Table S6b).

Results

General Dataset Trends and Statistical Significance

The General Dataset showed a marked post-ChatGPT shift, predominantly in English abstracts. During the pre-LLM period (2010–2022), marker prevalence in both languages was stable or increased only marginally, whereas English markers rose sharply after 2023 (Figure 1). Segmented regression confirmed this structural break, showing substantially steeper post-2022 slopes for most English marker groups than their Turkish equivalents (Table 2).

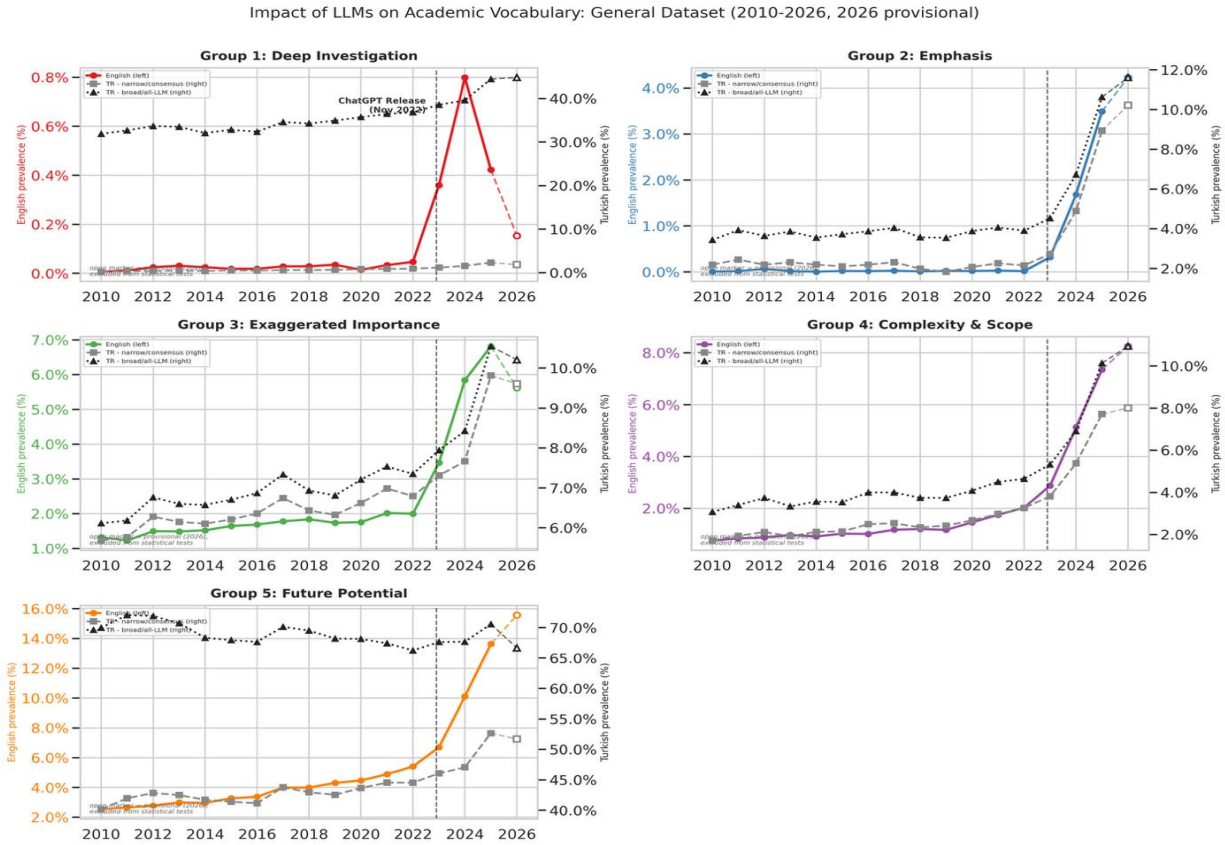


Figure 1. Longitudinal Trends of LLM Marker Groups in General Dataset Abstracts (2010–2026)

Longitudinal Trends of LLM Marker Groups in General Dataset Abstracts (2010–2026). Prevalence of five marker groups (see Table 1 for constituent terms) in English (left axis, solid) and Turkish (right axis, dashed) abstracts. The vertical dashed line indicates the public release of ChatGPT (November 2022). Open markers and dashed connecting lines for 2026 indicate provisional data, excluded from statistical comparisons due to

likely incomplete indexing (Section 2). (A) Group 1: Deep Investigation. (B) Group 2: Emphasis. (C) Group 3: Exaggerated Importance. (D) Group 4: Complexity & Scope. (E) Group 5: Future Potential.

Table 2. Segmented Regression Analysis: Pre-2022 and Post-2022 Trend Slopes for LLM Marker Words

Dataset	Marker Group	Group Term (Language)	Pre-2022 Slope (95% CI)	Post-2022 Slope (95% CI)
General	Group 1: Deep Investigation	Delve (ENG)	0.00 (0.00 to 0.00)	0.03 (-2.96 to 3.02)
		Derinlemesine (TR)	0.06 (0.05 to 0.06)	0.57 (-0.65 to 1.80)
	Group 2: Emphasis	Underscore (ENG)	0.00 (0.00 to 0.00)	1.59 (-0.05 to 3.22)
		Vurgula (TR)	-0.02 (-0.04 to 0.01)	3.10 (-3.79 to 10.00)
	Group 3: Exaggerated Importance	Crucial (ENG)	0.06 (0.05 to 0.07)	1.67 (-3.44 to 6.79)
		Kritik (TR)	0.09 (0.06 to 0.12)	1.25 (-5.31 to 7.82)
	Group 4: Complexity & Scope	Intricate (ENG)	0.09 (0.06 to 0.11)	2.24 (2.16 to 2.31)
		Karmaşık (TR)	0.10 (0.08 to 0.13)	1.95 (-0.92 to 4.82)
	Group 5: Future Potential	Potential (ENG)	0.23 (0.20 to 0.26)	3.47 (3.03 to 3.91)
		Potansiyel (TR)	0.26 (0.12 to 0.40)	3.30 (-13.65 to 20.24)
Medical	Group 1: Deep Investigation	Delve (ENG)	0.00 (0.00 to 0.00)	0.05 (-0.94 to 1.05)
		Derinlemesine (TR)	0.02 (0.01 to 0.03)	0.14 (0.09 to 0.20)
	Group 2: Emphasis	Underscore (ENG)	0.00 (0.00 to 0.00)	1.97 (-0.50 to 4.44)
		Vurgula (TR)	-0.13 (-0.15 to -0.10)	2.37 (-3.52 to 8.26)
	Group 3: Exaggerated Importance	Crucial (ENG)	0.07 (0.05 to 0.09)	2.03 (-2.90 to 6.97)
		Kritik (TR)	-0.03 (-0.08 to 0.02)	0.95 (-5.86 to 7.75)
	Group 4: Complexity & Scope	Intricate (ENG)	0.08 (0.05 to 0.11)	1.76 (-0.40 to 3.91)
		Karmaşık (TR)	0.06 (0.03 to 0.08)	1.34 (-1.07 to 3.75)
	Group 5: Future Potential	Potential (ENG)	0.29 (0.23 to 0.34)	4.45 (0.39 to 8.52)
		Potansiyel (TR)	-1.05 (-1.31 to -0.79)	1.95 (-10.06 to 13.97)

Note: Trend slopes represent the annual percentage change in marker prevalence. The structural break for the segmented regression was set at the release of ChatGPT (late 2022). Post-2022 slopes are calculated based on the primary analysis period (2023–2025), with 2026 excluded due to incomplete indexing. Turkish (TR) terms reflect the narrow/consensus set identified by the majority of LLMs. CI = Confidence Interval.

Group 2 (Emphasis) showed the largest increase: “underscore” rose from 0.02% to 1.80% of English abstracts (RR=78.73, 95% CI 64.49–96.10, Holm-adjusted p<0.001), compared with RR=2.54 (95% CI 2.47–2.61) for Turkish “vurgula.” “Similar English–Turkish disparities occurred for Group 1 (Deep Investigation: “delve,” RR=18.68, 95% CI 15.45–22.58 vs. “derinlemesine,” RR=2.56, 95% CI 2.43–2.70), Group 4 (Complexity & Scope: RR=3.88, 95% CI 3.75–4.01 vs. RR=2.23, 95% CI 2.16–2.29), and Group 3 (Exaggerated Importance: RR=3.06, 95% CI 2.96–3.15 vs. RR=1.27, 95% CI 1.25–1.30). Group 5 (Future Potential) showed the smallest gap, likely because Turkish “potansiyel” had a high baseline prevalence (>60%), limiting further increase (English RR=2.50, 95% CI 2.45–2.55; Turkish RR=1.13, 95% CI 1.12–1.14; Holm-adjusted p<0.001). Of the 15 Holm-Bonferroni–corrected comparisons in the General Dataset, 14 remained significant; only the broad-set Group 5 (Future Potential) comparison did not. Annual counts, prevalence estimates, and 95% Wilson confidence intervals are reported in Supplementary Table S4.

Translation-sensitivity analyses using the broad Turkish term set produced directionally consistent results for all groups, although effect magnitudes differed for Groups 1 and 5. Broad/narrow fold-change ratios were 0.76–1.01 for Groups 2–4, supporting robustness to term-list breadth. For Group 1, the effect remained significant but was substantially attenuated (narrow RR=2.56 vs. broad RR=1.18, ratio=0.41). For Group 5, the broad-set result was not significant (RR=1.00, 95% CI 1.00–1.00, Holm-adjusted p=0.445), consistent with the near-ceiling prevalence of “potansiyel.”

Medical Dataset Specifics

In the Medical Dataset, English markers showed similarly pronounced post-LLM increases, led by “underscore” (RR=87.98, 95% CI 62.51–123.84) and “delve” (RR=45.75, 95% CI 20.01–104.56); “crucial” (RR=3.73, 95% CI 3.52–3.95) and “intricate” (RR=3.89, 95% CI 3.65–4.15) also increased substantially. All five English comparisons remained significant after Holm-Bonferroni correction (all p<0.001). Annual counts, prevalence estimates, and 95% Wilson confidence intervals are reported in Supplementary Table S5.

Turkish medical markers were more heterogeneous. “Derinlemesine” showed the largest increase (RR=3.15, 95% CI 2.67–3.73, adjusted p<0.001), followed by “vurgula” (RR=2.02, 95% CI 1.90–2.15) and “karmaşık” (RR=2.07, 95% CI 1.95–2.20); “kritik” increased modestly (RR=1.13, 95% CI 1.08–1.19), whereas “potansiyel” declined slightly (RR=0.94, 95% CI 0.93–0.96). Of the ten Turkish comparisons (five narrow-set and five broad-set), seven remained significant after Holm-Bonferroni correction — narrow-set "derinlemesine," "vurgula," "kritik," and "karmaşık," plus broad-set "vurgula" and "karmaşık" (all p<0.001). The remaining three did not reach significance: broad-set "derinlemesine" and both the narrow- and broad-set "potansiyel" (each Holm-adjusted p=1.000),

Sensitivity analyses were consistent for most Turkish medical markers (broad/narrow ratio 0.77–1.01). However, “derinlemesine” was highly sensitive to term-list breadth (narrow RR=3.15 vs. broad RR=0.96; ratio=0.24), indicating that this finding is not robust and should be interpreted cautiously.

Discussion

These findings indicate a marked post-2022 lexical shift in Turkish academic abstracts. Marker prevalence was broadly stable during 2010–2022 but increased from 2023 onward in both datasets, consistent with global reports of growing LLM-associated vocabulary in scientific writing.[3]

A marked discrepancy emerged between English and Turkish abstracts. While English markers exhibited a marked increase post-2023—exemplified by a risk ratio of 78.73 for 'underscore' within the General Dataset—their Turkish equivalents showed a more moderate increase (e.g., 'vurgula,' RR=2.54). In the Medical Dataset, this gap widened further, with 'delve' showing a markedly larger relative increase in English (RR=45.75, 95% CI 20.01–104.56) than in Turkish, where the corresponding rise was more modest and sensitive to term-list choice (RR=3.15, but not robust to broader term sets; Supplementary Table S6b). Our data are consistent with the possibility that the core scientific ideas and

narratives are still being conceived and drafted in the authors' native language (Turkish), preserving the organic lexical distribution. The footprint may emerge during translation or polishing.[3,4]

The observed decline in some English markers after 2024 (Figures 1A and 2A) was more pronounced in the General Dataset (Group 1 "delve," peaking near 0.8% of English abstracts in 2024 before falling to approximately 0.43% in 2025 and 0.15% in 2026) than in the Medical Dataset. Possible explanations include stylistic saturation or deliberate avoidance of this well-publicized AI marker due to increased editorial scrutiny, though the present data cannot distinguish between these. Only 2026 data (Figures 1–2) remain provisional and are excluded from statistical comparisons.

The present design cannot distinguish whether elevated English marker prevalence reflects LLM-assisted content generation, LLM-assisted translation or editing of author-drafted text, or a combination of both; this distinction cannot be resolved through lexical frequency analysis alone.

One notable exception to this robustness pattern was 'derinlemesine' (Group 1) in the Medical dataset, where narrow- and broad-set results diverged sharply. This suggests that the apparent increase in this specific term-context combination is partly an artifact of which Turkish synonyms are included, rather than a robust lexical shift, and this particular result should be interpreted with more caution than the other Group 1–4 findings.

While the trend is widespread, the qualitative nature of the shift differs between the two datasets. The Medical Dataset showed a distinct preference for Group 4 (Complexity and Scope) markers, with terms like "intricate," "meticulous," and "comprehensive" nearly doubling in prevalence. This can be attributed to the specific pressures of medical publishing, where conveying methodological rigor is paramount. Medical researchers' writing patterns are consistent with using LLMs to polish the language of their study protocols and scope.[16]

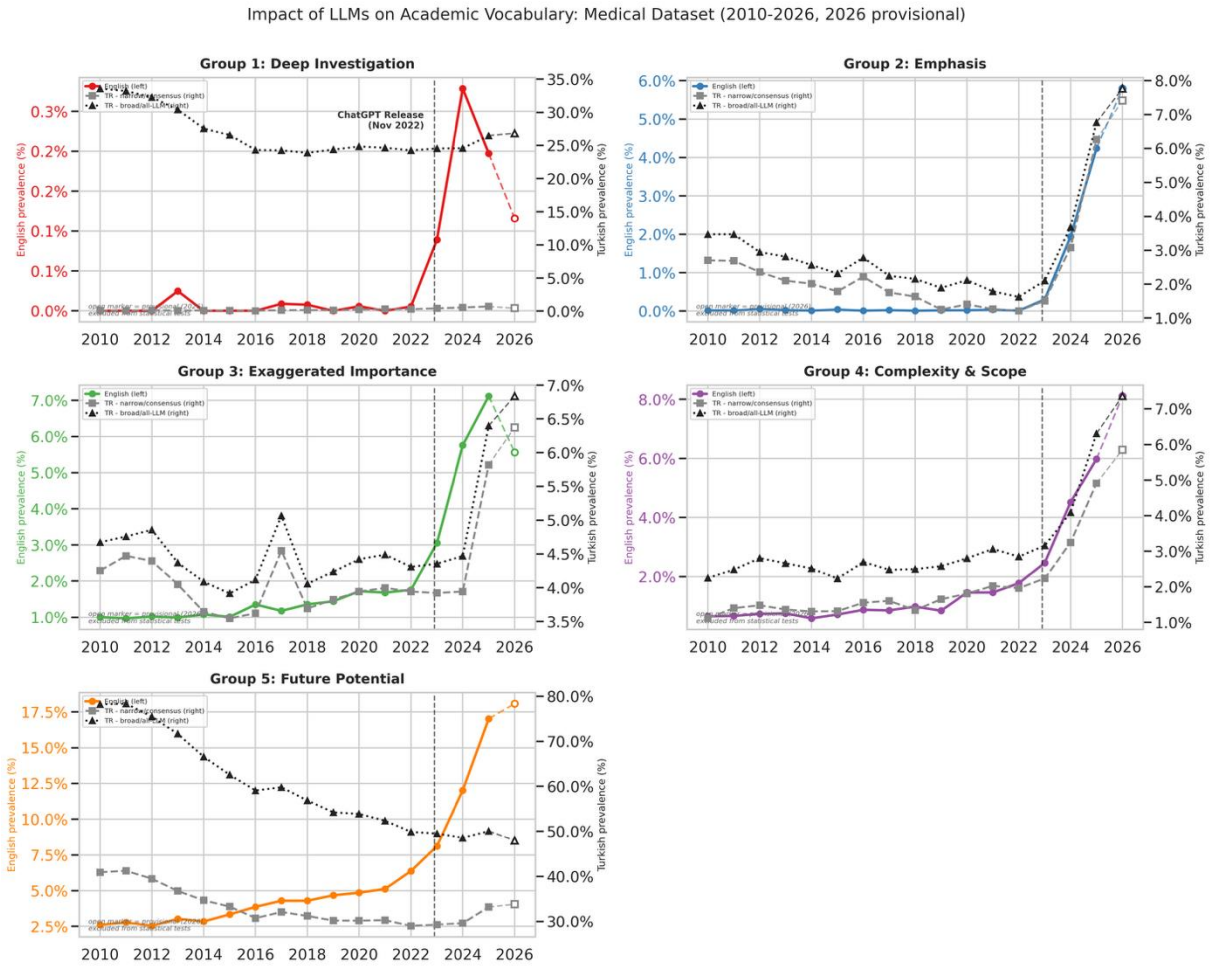


Figure 2. Longitudinal Trends of LLM Marker Groups in Medical Dataset Abstracts (2010–2026)

Longitudinal Trends of LLM Marker Groups in Medical Dataset Abstracts (2010–2026). Prevalence of five marker groups in English abstracts (left axis, solid line, circle markers) and Turkish abstracts (right axis) under two Turkish query sets: narrow/consensus (dashed line, square markers; primary analysis) and broad/all-LLM (dotted line, triangle markers; sensitivity analysis). 2026 data are shown as provisional (open markers) and excluded from statistical comparisons (Section 2.1). (A) Group 1: Deep Investigation. (B) Group 2: Emphasis. (C)–(E) Groups 3–5. English markers show larger relative increases than Turkish equivalents across most groups (Section 3.2); the narrow- and broad-set Turkish lines diverge notably in Panel A, reflecting the term-list sensitivity noted for "derinlemesine" in this dataset (Section 4.1, Supplementary Table S6b).

In contrast, the General Dataset showed sharper increases in connective and emphatic terms (Group 1: Delve and Group 2: Underscore). This may suggest that medical researchers may use AI to enhance technical description, whereas the broader academic community uses it to

improve narrative flow and cohesion. This distinction highlights that AI adoption is consistent with researchers deploying these tools to address domain-dependent writing challenges.

The word group 'delve' increased substantially in the post-LLM period (RR=18.68, General; RR=45.75, Medical). Juzek and Ward identify "delve" as a prime example of lexical overrepresentation driven by RLHF tuning, where models are incentivized to use "sophisticated" but often vacuous transition words.[4] The ubiquity of such terms threatens to homogenize scientific discourse. If a significant portion of the national literature converges on the limited stylistic patterns of a single model family (e.g., GPT-3.5/4), we risk losing the lexical diversity and unique authorial voices that characterize vibrant scientific debate.[3,17]

Our comparative data provide additional empirical support for the "Linguistic Equalizer" hypothesis proposed by Lin et al.[7] The stability of Turkish vocabulary alongside the substantial increase in English complexity markers is consistent with NNES researchers strategically deploying GenAI to bridge the 'native speaker premium' gap. The nearly four-fold relative increase in complex English adjectives (e.g., 'intricate,' RR=3.88, 95% CI 3.75–4.01) alters perceived sophistication. This represents a survival mechanism in a 'publish or perish' ecosystem rather than inherent scientific misconduct.[6] However, lacking validated Turkish excess vocabulary lists, the exact extent of native-language content generation remains open.

The risk ratios observed in this study suggest a potential long-term structural shift in academic production, but these markers do not imply scientific misconduct. Unlike the “tortured phrases” described by Cabanac et al. in predatory journals,[12] TR Dizin markers represent standard, albeit overused, academic English. The concern is therefore not marker presence itself, but insufficient transparency about AI use. As research integrity depends on disclosure,[18] AI-assisted drafting must balance efficiency with strict ethical oversight,[19] particularly given the risk of a “Bot Delusion,” in which researchers unknowingly propagate AI-hallucinated concepts.[20]

Blanket bans on AI are neither practical nor equitable. Because AI detectors have high false-positive rates among NNES authors,[6] studies should not be judged solely on lexical markers. Instead, journals and national indices such as TR Dizin should adopt COPE-aligned disclosure standards that distinguish legitimate AI-assisted linguistic editing from AI-generated content that may compromise research integrity.[23,11]

In clinical research, human-led verification of data and rigorous back-translation following AI-assisted translation are particularly important to preserve clinical fidelity and patient safety.

These findings extend beyond bibliometric trends. LLMs can generate fabricated citations and plausible but erroneous medical information, with risks increasing in specialized medical contexts,[24] while they are also being evaluated for direct clinical applications, including geriatric deprescribing and polypharmacy management in Family Medicine.[25] If AI-hallucinated data or fabricated clinical narratives enter the peer-reviewed literature, they may contaminate meta-analyses and evidence-based guidelines used by both clinicians and future medical AI systems. Rigorous transparency in AI use is therefore not only an issue of research integrity, but also a clinical safeguard against patient-safety risks.[26]

This study has several limitations. Lexical markers are probabilistic proxies rather than definitive evidence of AI use. [21,22] This may partly reflect lexical entrainment from widespread AI exposure.[5] Although marker-based tracking offers a transparent alternative to commercial AI detectors, which show high false-positive rates among NNES authors,[6] they cannot distinguish AI-assisted translation or language polishing from generative content creation. Reliance on Western LLM lexicons may also underestimate native-language AI use, as validated Turkish "excess vocabulary" lists are unavailable. Our 100-abstract manual review confirmed substantive term usage, but cannot exclude low background false-positive rates across the full corpus; moreover, candidate Turkish terms lacked independent professional linguistic validation. Furthermore, English and Turkish marker terms may differ functionally: vurgula- is relatively neutral in Turkish academic prose, whereas "underscore" is more marked and LLM-associated. Such rhetorical asymmetry may partly

explain the differing increases between languages. However, results were robust in nine of ten group/dataset combinations when analyses used the complete union of equivalents suggested by four LLMs (Supplementary Table S1 and S6b). The use of a single national index may introduce TR Dizin-specific translation inconsistencies and indexing delays, potentially affecting 2024–2026 dataset saturation. In addition, the blurred boundary between generative AI and advanced neural machine translation means that some observed fluency gains may reflect translation rather than generative drafting. Finally, because Turkish abstracts became optional in English-medium journals after September 2015, although Turkish titles, keywords, and abstracts continued to be encouraged to facilitate access for national researchers and enhance the visibility and impact of national publications,[27] changes in the Turkish-language denominator may partly contribute to the observed English–Turkish divergence. Future studies should examine full texts, citation outcomes, and other non-English academic ecosystems to further test the "equalizer" hypothesis.

In conclusion, this bibliometric analysis provides evidence consistent with Generative AI–related lexical change in a bilingual national index. The sharper post-2022 increase in English markers than in Turkish equivalents is consistent with LLMs functioning as "Linguistic Equalizers," particularly for translation and language polishing. However, lexical markers cannot establish AI authorship or scientific misconduct. In medicine, responsible AI use should therefore combine transparent disclosure with strict human verification of clinical content to protect research integrity and patient safety.

Ethical Considerations: Ethics committee approval was not required for this study as it involves the bibliometric analysis of publicly available, non-identifiable data retrieved from the TR Dizin database and does not involve human participants or animal experiments.

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

Note: Additional tables related to this article are available from the editorial office upon request at info@ankaramedicaljournal.com

References

1. Thorp HH. ChatGPT is fun, but not an author. Science. 2023;379(6630):313 doi:10.1126/science.adg7879.
2. Van Dis EAM, Bollen J, Zuidema W, van Rooij R, Bockting CL. ChatGPT: five priorities for research. Nature. 2023;614(7947):224-6 doi:10.1038/d41586-023-00288-7.
3. Kobak D, González-Márquez R, Horvát EÁ, Lause J. Delving into LLM-assisted writing in biomedical publications through excess vocabulary. Sci Adv. 2025;11(27):eadt3813 doi:10.1126/sciadv.adt3813.
4. Juzek TS, Ward ZB. Why Does ChatGPT "Delve" So Much? Exploring the Sources of Lexical Overrepresentation in Large Language Models. In: Proceedings of the 31st International Conference on Computational Linguistics. 2025:6397-412.
5. Weidinger L, Uesato J, Rauh M, et al. Taxonomy of Risks posed by Language Models. In: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22). New York: Association for Computing Machinery; 2022:214-29 doi:10.1145/3531146.3533088.
6. Liang W, Yuksekgonul M, Mao Y, Wu E, Zou J. GPT detectors are biased against non-native English writers. Patterns (NY). 2023;4(7):100779 doi:10.1016/j.patter.2023.100779.
7. Lin D, Zhao N, Tian D, Li J. ChatGPT as Linguistic Equalizer? Quantifying LLM-Driven Lexical Shifts in Academic Writing. arXiv [Preprint]. 2025; https://doi.org/10.48550/arXiv.2504.12317. (Accessed: 17.02.2026).
8. Moorhouse BL, Wan Y, Wu C, Wu M, Ho T. Generative AI Tools and Empowerment in L2 Academic Writing. System. 2025;133:103779 doi:10.1016/j.system.2024.103779.
9. Kudina O, Ballsun-Stanton B, Alfano M. The use of large language models as scaffolds for proleptic reasoning. Asian J Philos. 2025;4(1):24 doi:10.1007/s44204-025-00247-1.
10. Akyon SH, Akyon FC. Digital transformation of clinical education through artificial intelligence: a strengths, weaknesses, opportunities, and threats (SWOT) analysis. Ankara Med J. 2025;(1):96-118 doi:10.5505/amj.2025.63373.

11. Flanagin A, Kendall-Taylor J, Bibbins-Domingo K. Guidance for Authors, Peer Reviewers, and Editors on Use of AI, Language Models, and Chatbots. JAMA. 2023;330(8):702-3 doi:10.1001/jama.2023.15606.
12. Cabanac G, Labbé C, Magazinov A. Tortured phrases: A dubious writing style emerging in science. arXiv [Preprint]. 2021; https://doi.org/10.48550/arXiv.2107.06751. (Accessed: 17.02.2026).
13. Shen S, Wang K. Detecting AI-Generated Content in Academic Peer Reviews. arXiv [Preprint]. 2026; https://doi.org/10.48550/arXiv.2602.00319. (Accessed: 17.02.2026).
14. Stern JM, Fernandez-Perez AM, Cruz-Ossa N, Hernandez VH, McNamara CA, D'Apuzzo MR. Detecting Artificial Intelligence-Generated Text in Personal Statements of Adult Reconstruction Fellowship Applicants. J Arthroplasty. 2026;41(3):958-62 doi:10.1016/j.arth.2025.07.050.
15. Gray A. Estimating the prevalence of LLM-assisted text in scholarly writing. arXiv [Preprint]. 2025; https://doi.org/10.48550/arXiv.2512.01560. (Accessed: 17.02.2026).
16. Gencer G, Gencer K. Large Language Models in Healthcare: A Bibliometric Analysis and Examination of Research Trends. J Multidiscip Healthc. 2025;18:223-38 (doi:10.2147/JMDH.S460980).
17. Brainard J. Journals take up arms against AI-written text. Science. 2023;379(6634):740-1 doi:10.1126/science.adk0061.
18. Peh WCG, Saw A. ChatGPT in Medical Research and Scientific Writing. Malays Orthop J. 2023;18:58-60 doi:10.51866/cm0006.
19. Agarwal A, et al. Artificial Intelligence in Scientific Writing: Balancing Innovation and Efficiency with Integrity. World J Mens Health. 2025;43:e19 doi:10.5534/wjmh.230005.
20. Wieczorek O, Steinhardt I, Schmidt R, Mauermeister S, Schneijderberg C. The Bot Delusion. Large language models and anticipated consequences for academics' publication and citation behavior. Futures. 2025;166:103537 doi:10.1016/j.futures.2024.103537.

21. Giglio AD, Costa MUPD. The use of artificial intelligence to improve the scientific writing of non-native English speakers. Rev Assoc Med Bras. 2023;69(9):e20230560 doi:10.1590/1806-9282.20230560.
22. Chávez-Martínez O. ¿Cómo detectar textos científicos generados con inteligencia artificial? [How to detect scientific texts generated with artificial intelligence?]. Rev Med Inst Mex Seguro Soc. 2026;64(1):e6952 doi:10.5281/zenodo.17477984.
23. Lund BD, Wang T, Mannuru NR, Nie B, Shimray S, Wang Z. ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. J Assoc Inf Sci Technol. 2023;74(5):570-81 doi:10.1002/asi.24750.
24. Linardon J, Jarman HK, McClure Z, Anderson C, Liu C, Messer M. Influence of Topic Familiarity and Prompt Specificity on Citation Fabrication in Mental Health Research Using Large Language Models: Experimental Study. *JMIR Ment Health*. 2025;12:e80371 doi:10.2196/80371.
25. Rao A, Kim J, Lie W, Pang M, Fuh L, Dreyer KJ, Succi MD. Proactive Polypharmacy Management Using Large Language Models: Opportunities to Enhance Geriatric Care. *J Med Syst*. 2024;48(1):41. doi: 10.1007/s10916-024-02058-y.
26. Pathak JL, Zhang K, Jiang H, et al. Generative AI adoption and ethical perceptions: a comparative study of medical and non-medical researchers in Chinese universities. *BMC Med Ethics*. 2026;27(1):79 doi:10.1186/s12910-026-01415-5.
27. Yabancı Dildeki Makalelerde Türkçe Zorunluluğu [Turkish Language Requirement for Articles in Foreign Languages]. Updated 14 September 2015. Available at: https://trdizin.gov.tr/yabanci-dildeki-makalelerde-turkce-zorunlulugu/Accessed: 10 August 2026.

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):533-554 // 10.5505/amj.2026.06777

ASSOCIATION BETWEEN NUTRITIONAL RISK AND ANTICHOLINERGIC BURDEN AMONG INTERNAL MEDICINE INPATIENTS: A RETROSPECTIVE CROSS-SECTIONAL STUDY

Serap Balaban Barta[1], Rıdvan Erdin[2], Hilal Şimşek[3], Rıfat Bozkuş[2], Emin Gemcioğlu[2]

[1]Department of Nutrition and Dietetics, Faculty of Health Sciences, Gaziantep University, Gaziantep, Türkiye

[2]Department of Internal Medicine, Ankara Etlik City Hospital, Ankara, Türkiye

[3]Department of Nutrition and Dietetics, Bor Faculty of Health Sciences, Niğde Ömer Halisdemir University, Niğde, Türkiye

Correspondence:

Serap Balaban Barta (e-mail: dyt.balaban@gmail.com)

Submitted: 14.05.2026 // Accepted: 24.08.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: Nutritional risk is common among hospitalized patients, and anticholinergic burden may impair nutritional status. This study aimed to evaluate the association between anticholinergic burden and nutritional risk among adult patients hospitalized in an internal medicine ward.

Materials and Methods: This retrospective cross-sectional study included 1570 internal medicine inpatients. Nutritional risk was assessed using the Nutritional Risk Screening 2002 (NRS-2002), with scores ≥3 indicating high nutritional risk. Anticholinergic burden was calculated using the Anticholinergic Cognitive Burden (ACB) Scale and categorized as no burden, low–moderate burden, and high burden. Binary logistic regression identified factors associated with high nutritional risk.

Results: Among the participants, 51.2% had no anticholinergic burden, 32.3% had low–moderate burden, and 16.5% had high burden. High nutritional risk increased progressively across anticholinergic burden groups, from 47.3% in the no-burden group to 69.2% in the low–moderate burden group and 88.8% in the high-burden group ($p<0.001$). NRS-2002 scores also increased significantly with increasing anticholinergic burden (2.64±1.81, 3.55±1.82, and 4.25±1.48, respectively; $p<0.001$). In the adjusted logistic regression model, each one-point increase in ACB score was independently associated with higher odds of high nutritional risk (OR=1.96, 95% CI: 1.71–2.34, $p<0.001$), after controlling for potential confounders. Lower albumin level was also independently associated with high nutritional risk (OR=0.92; 95% CI: 0.89–0.95; $p<0.001$).

Conclusion: A higher anticholinergic burden was independently associated with higher odds of nutritional risk among internal medicine inpatients. These findings suggest that anticholinergic exposure may be a clinically relevant and modifiable factor in nutritional risk assessment.

Keywords: Anticholinergic burden; nutritional risk; internal medicine; polypharmacy; hospitalized patients.

Introduction

Nutritional risk is a common health problem that substantially increases morbidity and mortality, prolongs hospital stay, and raises treatment costs, particularly among hospitalized individuals. This condition may adversely affect patients' recovery processes and lead to a decline in their quality of life.[1,2] In a previous study, the prevalence of nutritional risk among hospitalized patients in an internal medicine clinic was reported to be 44.2%, increasing to 46.5% among older adults.[3] Consistent with these findings, a recent study conducted in the internal medicine ward of a tertiary hospital in Türkiye found that 61.4% of patients were at high risk of nutritional risk.[4] This condition may directly impair patients' recovery processes and lead to a substantial decline in their quality of life.[5] Therefore, early identification of nutritional risk is clinically important for improving prognosis in hospitalized patients. The Nutritional Risk Screening 2002 (NRS-2002) is a clinically valuable screening tool in hospital settings because it evaluates both nutritional status and disease severity. Previous studies have shown that higher NRS-2002 scores are associated with adverse short- and long-term clinical outcomes.[6,7]

In addition, several factors may contribute to an increased nutritional risk among internal medicine ward patients, including advanced age, polypharmacy, comorbidity burden, functional dependence, inadequate energy intake, and prolonged hospital stay.[8,9] Among the factors affecting nutritional risk, medication use plays a critical role, especially in older adults and individuals with multiple comorbidities. Anticholinergic medications may compromise nutritional status by limiting food intake and nutrient absorption through adverse effects such as dry mouth, impaired swallowing function, and gastrointestinal disturbances.[10] Polypharmacy is associated with drug-induced reductions in nutritional intake and impairment of oral and gustatory functions, particularly among older and hospitalized individuals.[11,12] Polypharmacy is commonly defined as the concurrent use of five or more medications.[12] A study reported that 57% of individuals aged 70 and over receiving home care services were at nutritional risk, and 9.3% were malnourished. The same study reported a mean number of 7.3 regularly used medications, while 27.8% of

participants had experienced moderate to severe reductions in food intake and weight loss during the preceding three months.[13] Another study indicated that the prevalence of eating or swallowing difficulties was 20.9% among older adults with polypharmacy, compared with 6.7% among those without polypharmacy.[14] Taken together, these findings suggest that polypharmacy may adversely affect nutritional status and that anticholinergic medications, in particular, warrant further consideration in relation to nutritional risk, given their potential association with adverse effects such as dry mouth, gastrointestinal symptoms, and swallowing difficulties.[15,16] Although previous studies have reported significant associations between anticholinergic burden and nutritional risk[17,18], further evidence is needed to clarify the extent of this relationship and its interplay with other clinical variables.[17] This study aimed to retrospectively evaluate the association between nutritional risk and anticholinergic burden in a large cohort of patients hospitalized in an internal medicine ward.

Materials and Methods

Study design and population

This retrospective cross-sectional study was conducted to evaluate the association between anticholinergic burden and nutritional risk among adult patients hospitalized in the Internal Medicine Department of Etlik City Hospital in Ankara, Türkiye. Etlik City Hospital is a large tertiary care public hospital with a high-volume and heterogeneous inpatient population. The Internal Medicine Department provides care for patients with various acute and chronic medical conditions, including multimorbidity and complex medication regimens.

The study population consisted of patients aged 18 years and older who were hospitalized in the Internal Medicine ward between June 1, 2025, and December 31, 2025. Data were collected retrospectively from the hospital electronic medical records, patient files, medication records, nutritional screening forms, and routine laboratory results. Demographic characteristics, clinical information, comorbidities, medication use, laboratory parameters, hospitalization-related variables, and NRS-2002 scores recorded during

hospital admission were reviewed (Figure 1). The sample size was determined using an a priori power analysis performed with G*Power version 3.1. The calculation was based on the effect size reported in a similar study evaluating nutritional risk defined by NRS-2002. An odds ratio of 1.23 for anticholinergic burden, a nutritional risk prevalence of 32.6%, a two-sided α level of 0.05, and 80% statistical power were assumed. Based on a two-tailed binary logistic regression model, the minimum required sample size was calculated as 848 patients. Considering the possibility of missing or incomplete data, the target sample size was increased by approximately 10-15%, corresponding to at least 932-975 participants.[15]

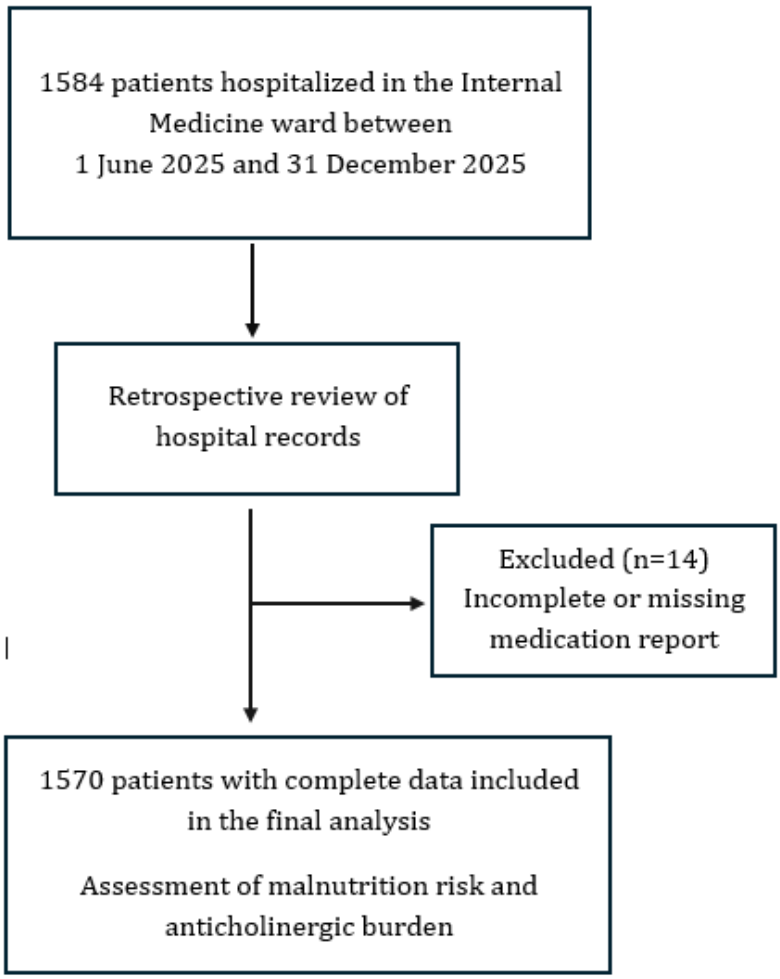


Figure 1. Flow diagram of participants included in the study

Patients with incomplete nutritional screening data, missing or unclear medication records, missing key clinical or laboratory data, day-case admissions, terminal or palliative care status, severe cognitive impairment, and severe dysphagia that prevented oral intake were excluded from the analysis. The study was approved by the Etlik City Hospital Ethics Committee with approval number AEŞH-BADEK1-2026-360. The study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the relevant ethics committee prior to data collection.

Data collection

In this study, data were collected retrospectively from hospital electronic medical records, patient files, medication records, nutritional screening forms, and routine laboratory results. The sociodemographic variables included age, sex, and educational status. The clinical data included the type of admission for hospitalization (classified using ICD-10 codes), hospitalization duration in the past year, length of stay in the internal medicine ward, hospitalization duration after enrolment, 30-day discharge or survival status, Charlson Comorbidity Index score, total number of medications, and polypharmacy status. Medication lists recorded at hospital admission were reviewed to calculate the anticholinergic burden. Nutritional risk data were obtained from the NRS-2002 forms completed during routine clinical assessment at admission. In addition, anthropometric measurements, including body weight, height, and body mass index, were recorded from patient files. Routine laboratory parameters obtained during hospitalization were also collected for analysis, including albumin (g/L), hemoglobin (g/dL), C-reactive protein (CRP, mg/L), urea (mg/dL), uric acid (mg/dL), and creatinine (mg/dL). These biochemical parameters were analyzed as continuous variables without predefined cut-off points.

Nutritional risk assessment

Nutritional risk was evaluated using the NRS-2002 scale, which is a valid and reliable screening tool widely used to assess nutritional risk among hospitalized patients. The initial screening included questions regarding low BMI, recent weight loss, reduced dietary intake, and severe illness. If any item is positive, final screening is performed by scoring the impaired nutritional status and disease severity from 0 to 3 points each. Following the initial screening, patients with at least one positive response underwent final screening, in which impaired nutritional status and disease severity were each scored from 0 to 3 points. An additional point was added for patients aged ≥70 years. A total score of ≥3 indicates nutritional risk and the need for a nutritional care plan. Therefore, patients were categorized as having a low nutritional risk when the NRS-2002 score was <3 and a high nutritional risk when the score was ≥3.[19]

Anticholinergic Cognitive Burden (ACB) Scale

Anticholinergic burden was assessed using the Anticholinergic Cognitive Burden (ACB) Scale. The ACB Scale is a commonly used tool that quantifies anticholinergic exposure by assigning each medication a score according to its anticholinergic activity. In this scale, each medication is scored as 0, 1, 2, or 3, and the total ACB score is calculated by summing the individual scores of all medications used by each patient. Based on the total ACB score, patients were categorized into three groups: no burden (score = 0), low–moderate burden (score = 1 and 2), and high burden (score ≥3).[20] The ACB Scale was used because it is a widely used and clinically practical tool for estimating anticholinergic burden based on routinely recorded medication data. Since this study had a retrospective design, the ACB Scale was considered appropriate because it does not require detailed dose information and can be applied using medication lists available in hospital records. In addition, the use of the ACB Scale was consistent with the routine assessment approach used in the study setting. [20]

Comorbidity assessment

The comorbidity burden was assessed using the Charlson Comorbidity Index (CCI), a widely used scoring system based on 19 predefined chronic conditions. A weighted score was assigned to each condition, and higher CCI total scores indicated a higher burden of comorbidities and poorer overall health status.[21]

Anthropometric Measurements

Anthropometric data, including body weight and height, were obtained from the patients' records. The body mass index (BMI) was calculated as body weight in kilograms divided by height in meters squared (kg/m^2). For categorical analyses, BMI was classified as underweight (<18.5 kg/m^2), normal weight (18.5–24.9 kg/m^2), and overweight/obese (≥25.0 kg/m^2).[22]

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA). The normality of continuous variables was assessed using the Kolmogorov–Smirnov test, visual inspection of histograms, and skewness and kurtosis values. Variables with skewness and kurtosis values within ±1.5 were considered to be approximately normally distributed.[23] Continuous variables were presented as mean ± standard deviation for normally distributed data and median with interquartile range for non-normally distributed data. Categorical variables were expressed as numbers and percentages. Patients were categorized into three groups according to anticholinergic burden: no burden (ACB=0), low–moderate burden (ACB=1 and 2), and high burden (≥3). Categorical variables were compared using Pearson's chi-square test. For categorical variables with significant overall differences, Bonferroni-adjusted pairwise comparisons of column proportions were performed. Continuous variables were compared across anticholinergic burden groups using one-way ANOVA for normally distributed variables and the Kruskal–Wallis test for non-normally distributed variables. When significant differences were detected, Bonferroni post-hoc tests were used for parametric analyses, and Dunn–Bonferroni post-hoc tests were used for nonparametric analyses. Bivariate correlation

analyses were conducted to examine the relationships among ACB score, NRS-2002 score, demographic and anthropometric variables, hospitalization-related variables, and biochemical parameters. Pearson correlation analysis was used for normally distributed variables, whereas Spearman correlation analysis was used for non-normally distributed variables. The results of the bivariate correlation analyses were visualized using a heatmap. High nutritional risk was defined as an NRS-2002 score ≥3. Univariable logistic regression analyses were first conducted to calculate unadjusted odds ratios. Multivariable logistic regression models were then used to estimate adjusted odds ratios after controlling for potential confounders, including sex, age, educational status, number of medications, BMI, CCI, hospitalization in the past year (days), length of stay in the IM ward (days), and hospitalization duration after enrolment (days). Logistic regression results were presented as odds ratios (ORs) with 95% confidence intervals (CIs). The adjusted model showed adequate fit (Hosmer–Lemeshow p=0.078) and explained 27.9% of the variance in high nutritional risk (Nagelkerke R^2=0.279). Age and BMI were included in the main multivariable model because they are clinically important potential confounders of the association between anticholinergic burden and nutritional risk. However, because the NRS-2002 includes BMI-related nutritional impairment and an additional point for patients aged ≥70 years, a sensitivity analysis excluding age and BMI was also performed. Patients with missing data for the variables required for the regression analyses were excluded, and all regression models were performed using a complete-case approach. No imputation was applied for missing data. A p-value of <0.05 was considered statistically significant. P values were reported as exact values where possible, and values below 0.001 were reported as p<0.001. This retrospective observational study was reported in accordance with the STROBE guidelines.

Results

The sociodemographic, clinical, anthropometric, and biochemical characteristics of the patients according to ACB are shown in Table 1. Among the inpatients, 51.2% (n=804) had no anticholinergic burden, 32.3% (n=507) had low–moderate anticholinergic burden, and

16.5% (n=259) had high anticholinergic burden. Educational status showed a significant distributional difference across anticholinergic burden groups (p<0.001). Primary school graduates represented the largest subgroup overall (45.9%). Post-hoc comparisons indicated that illiterate participants were underrepresented in the no-burden group (13.6%) and overrepresented in the low–moderate burden group (25.0%), whereas participants with high school or university-level education were more frequently observed in the no-burden group. Thirty-day clinical status differed significantly according to anticholinergic burden (p<0.001). Patients without anticholinergic burden had a higher discharge rate (93.0%) and lower mortality rate (1.6%) than those with low–moderate or high burden. Pairwise comparisons showed that both burden groups differed significantly from the no-burden group, while low–moderate and high-burden groups did not differ from each other. High nutritional risk according to NRS-2002 increased stepwise across anticholinergic burden groups, from 47.3% in the no-burden group to 69.2% in the low–moderate burden group and 88.8% in the high-burden group (p<0.001). Pairwise comparisons indicated significant differences between all burden groups. ICU transfers were significantly more frequent in both burden groups, whereas outpatient admissions were significantly more frequent in the no-burden group. Emergency admissions and transfers from other wards did not differ significantly between groups. BMI category distribution differed significantly across anticholinergic burden groups (p=0.006). Although overweight/obesity was numerically most frequent in the high-burden group, post-hoc comparisons indicated that the significant difference was primarily due to the lower proportion of overweight/obese patients in the low–moderate burden group (55.6%).

ANKARA MEDICAL JOURNAL

Table 1. Socio-Demographic characteristics, anthropometric measurements, scale scores, and biochemical results of participants by anticholinergic burden

	No Burden (n=804)	Low-moderate Burden (n=507)	High Burden (n=259)	Total (n=1570)	p value
Socio-Demographic Characteristics					
Sex, *n(%)*					0.130[d]
Female	417 (52.9)	290 (57.2)	146 (56.4)	853 (54.3)	
Male	387 (48.1)	217 (42.8)	113 (43.6)	717 (45.7)	
Age (year) ± SD	60.08±18.15[a]	70.36±14.39[b]	68.68±15.73[b]	64.82±17.32	<0.001[e]
Education Status, *n(%)*					<0.001[d]
Illiterate	109 (13.6)	127 (25.0)	57 (22.0)	293 (18.7)	
Literate	56 (7.0)	35 (6.9)	13 (5.0)	104 (6.6)	
Primary school	364 (45.3)	227 (44.8)	129 (49.8)	720 (45.9)	
Middle school	90 (11.2)	41 (8.1)	22 (8.5)	153 (9.7)	
High school	115 (14.3)	50 (9.9)	21 (8.1)	186 (11.8)	
University	70 (8.6)	27 (5.3)	17 (6.6)	114 (7.3)	
Clinical Characteristic					
Hospitalization in past year (days)	3.0 (10.0)[a]	8.0 (21.0)[b]	5.0 (15.0)[c]	4.0 (14.0)	<0.001[f]
Hospitalization in IM ward (days)	7.0 (8.0)[a]	10.0 (9.5)[b]	9.0 (9.0)[b]	9.0 (9.0)	<0.001[f]
Hospitalization after enrolment (days)	8.0 (8.5)[a]	11.0 (13.0)[b]	11.0 (11.0)[b]	10.0 (11.0)	<0.001[f]
Number of medications ± SD	6.10±3.01[a]	8.21±2.96[b]	8.29±3.19[b]	7.20±3.21	<0.001[e]
Polypharmacy[g], *n(%)*	467 (66.5)	452 (89.5)	223 (86.8)	1142 (78.0)	<0.001[d]
Survival within 30 days follow-up					<0.001[d]
Inpatient	43 (5.4)	46 (9.1)	23 (8.9)	112 (7.1)	
Discharged	748 (93.0)	435 (85.8)	220 (84.9)	1403 (89.4)	
Exitus	13 (1.6)	26 (5.1)	16 (6.2)	55 (3.5)	
Admission to the ward, *n(%)*	487 (60.6)	296 (58.4)	140 (54.1)	923 (58.8)	<0.001[d]
Emergency	12 (1.5)	6 (1.2)	3 (1.2)	21 (1.3)	
Transfer	157 (19.5)	151 (29.8)	84 (32.4)	392 (25.0)	
ICU	148 (18.4)	54 (10.7)	32 (12.4)	234 (14.9)	
Outpatient clinic					
Anthropometric Measurement					
Weight (kg) ± SD	74.44±15.99	72.61±17.24	74.21±16.78	73.81±16.54	0.137[e]
BMI (kg/m^2) ± SD	27.45±9.18	26.98±5.92	27.46±5.90	27.30±7.76	0.522[e]
Underweight, *n(%)*	27 (3.4)	15 (3.0)	14 (5.4)	56 (3.6)	0.006[d]
Normal, *n(%)*	273 (34.0)	210 (41.4)	77 (29.7)	560 (35.7)	
Overweight and obese, *n(%)*	504 (62.7)	282 (55.6)	168 (64.9)	954 (60.8)	

Scale Scores					
ACB ± SD	0.0±0.0[a]	1.19±0.39[b]	3.40±0.76[c]	1.16±1.36	<0.001[e]
ACB, *n(%)*	804 (51.2)	507 (32.3)	259 (16.5)	1570 (100.0)	
CCI ± SD	4.26±2.99[a]	6.11±3.01[b]	5.82±3.05[b]	5.11±3.13	<0.001[e]
NRS-2002 ± SD	2.64±1.81[a]	3.55±1.82[b]	4.25±1.48[c]	3.20±1.87	<0.001[e]
Low nutritional risk, *n(%)*	424 (52.7)	156 (30.8)	29 (11.2)	609 (38.8)	<0.001[d]
High nutritional risk, *n(%)*	380 (47.3)	351 (69.2)	230 (88.8)	961 (61.2)	
Biochemical results					
Hemoglobin (g/dL) ± SD	11.67±2.77[a]	10.38±2.47[b]	10.97±2.57[c]	11.14±2.70	<0.001[e]
Albumin (g/L) ± SD	38.11±5.58[a]	34.48±5.78[b]	35.51±6.02[b]	35.99±5.84	<0.001[e]
Urea (mg/dL) ± SD	50.39±35.87[a]	76.68±51.54[b]	66.54±48.26[c]	61.54±45.17	<0.001[e]
Uric acid (mg/dL) ± SD	5.61±2.23[a]	6.93±2.95[b]	6.13±2.68[c]	6.12±2.62	<0.001[e]
CRP (mg/L)	13.49(51.81)[a]	25.27(59.36)[b]	22.85(60.68)[c]	19.13(55.45)	0.001[f]
Creatinine (mg/dL)	0.91 (0.81)[a]	1.38 (1.24)[b]	1.11 (1.06)[c]	1.05 (1.07)	<0.001[f]

Note: Numerical data and categorical variables are presented as mean ± SD or median (IQR) and n (%). Abbreviations: BMI, body mass index; ACB, anticholinergic burden; CCI, Charlson Comorbidity Index; CRP, C-reactive protein; ICU, intensive care unit; IM, internal medicine; NRS-2002, Nutritional Risk Screening; SD, standard deviation. [a, b, c] Different superscript letters within the same row indicate statistically significant differences between anticholinergic burden groups. [d] Pearson's $\chi 2$ test (Bonferroni-adjusted pairwise comparisons of column proportions). [e] One-way ANOVA test (Bonferroni post-hoc test). [f] Kruskal-Wallis test (Dunn-Bonferroni post-hoc test). [g] Defined as $\geq$5 medication usage daily. * $p < 0.05$

Univariable and multivariable logistic regression analysis was used to determine the predictors of high nutritional risk (Table 2). In the unadjusted analyses, the ACB score was significantly associated with high nutritional risk (OR=1.873; 95% CI: 1.665–2.106; p<0.001). Albumin, uric acid, and CRP levels were also significantly associated with the nutritional risk, whereas hemoglobin, urea, and creatinine levels were not. In the multivariable logistic regression model, ACB score remained independently associated with high nutritional risk after adjustment for age, sex, educational status, number of medications, BMI, CCI, biochemical parameters, and hospitalization-related variables. Each one-point increase in ACB score was associated with 1.96-fold higher odds of high nutritional risk (OR=1.956; 95% CI: 1.710–2.237; p<0.001). In the adjusted model, lower albumin level was also independently associated with high nutritional risk (OR=0.918; 95%

CI: 0.890–0.947; p<0.001). Age, male sex, number of medications, uric acid, CRP, and hospitalization in the past year were also significantly associated with high nutritional risk in the adjusted model.

Bivariate correlation analysis was performed to examine the relationships among ACB score, NRS-2002 score, demographic and clinical variables, hospitalization-related parameters, and biochemical markers (Figure 2). The ACB score showed a significant positive correlation with NRS-2002 score (r=0.315, p<0.001), indicating that a higher anticholinergic burden was associated with higher nutritional risk. ACB score was also positively correlated with age (r=0.198, p<0.001), Charlson Comorbidity Index (r=0.197, p<0.001), number of medications (r=0.299, p<0.001), hospitalization in the past year (ρ=0.188, p<0.001), hospitalization in the internal medicine ward (ρ=0.193, p=0.002), and hospitalization after enrolment (ρ=0.203, p<0.001). In contrast, ACB score showed weak negative correlations with hemoglobin (r=-0.112, p<0.001) and albumin (r=-0.115, p<0.001). Similarly, NRS-2002 score was positively correlated with age, comorbidity burden, hospitalization-related variables, CRP, and urea, and negatively correlated with body weight, BMI, hemoglobin, and albumin.

ANKARA MEDICAL JOURNAL

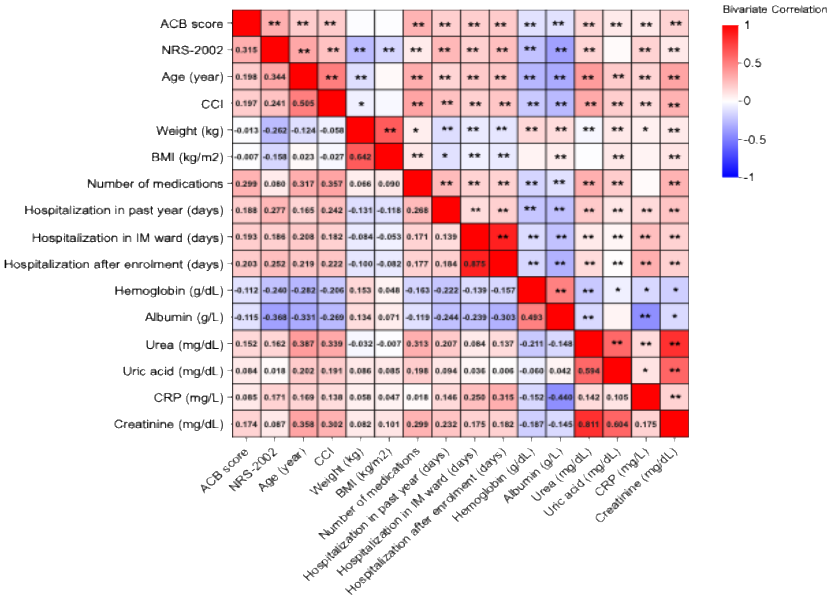


Figure 2. Bivariate correlation heatmap of ACB, NRS-2002, and key covariates. Bivariate correlation heatmap showing the relationships among ACB score, NRS-2002, demographic and anthropometric variables, hospitalization-related variables, and laboratory parameters. Correlation coefficients are displayed in the lower triangle, and color intensity represents the magnitude and direction of the correlations (red = positive; blue = negative). Abbreviations: CCI: Charlson comorbidity index; BMI: body mass index; IM: internal medicine; NRS-2002: The Nutritional Risk Screening 2002; CRP: C-reactive protein. Pearson and Spearman correlation was used, *p<0.05; **p<0.01

Sensitivity analyses were performed to assess the possible overlap of age and BMI with the NRS-2002 assessment (Table 3). ACB score remained significantly associated with high nutritional risk when age, BMI, or both were excluded from the model.

Table 2. Univariable and multivariable logistic regression analysis of factors associated with high nutritional risk

		Nutritional Risk				
		Unadjusted Model (R^2= 0.260)			Adjusted Model (R^2= 0.279) [c]	
Independent Variables	**β**	**OR (95% CI)**	***p value***	**β**	**OR (95% CI)**	***p value***
ACB	0.627	1.873 (1.665-2.106)	**<0.001***	0.671	1.956 (1.710-2.237)	**<0.001***
Hemoglobin (g/dL)	-0.026	0.975 (0.929-1.023)	0.302	-0.043	0.958 (0.902-1.018)	0.164
Albumin (g/L)	-0.120	0.887 (0.864-0.911)	**<0.001***	-0.085	0.918 (0.890-0.947)	**<0.001***
Urea (mg/dL)	-0.005	1.005 (1.001-1.009)	0.026*	0.004	1.004 (0.998-1.004)	0.161
Uric acid (mg/dL)	-0.074	0.928 (0.878-0.982)	**0.014***	-0.081	0.922 (0.864-0.984)	**0.014***
CRP (mg/L)	-0.003	0.997 (0.996-0.999)	**0.006***	-0.003	0.997 (0.995-0.999)	**0.006***
Creatinine (mg/dL)	-0.053	0.948 (0.816-1.101)	0.484	-0.067	0.935 (0.773-1.130)	0.487
Age (year)	_	_	_	0.018	1.019 (1.007-1.030)	**<0.001***
Sex[a]	_	_	_	-0.345	0.709 (0.525-0.957)	**0.025***

Educational status[b]	_	_	_	0.002	1.002 (0.554-1.812)	0.995
Number of medications (n)	_	_	_	-0.077	0.926 (0.882-0.972)	**0.002***
BMI (kg/m²)	_	_	_	-0.019	0.981 (0.959–1.003)	0.092
CCI	_	_	_	-0.030	0.971 (0.919–1.025)	0.287
Hospitalization in past year (days)	_	_	_	0.012	1.012 (1.003–1.020)	**0.005***
Hospitalization in IM ward (days)	_	_	_	-0.018	0.983 (0.955–1.011)	0.227
Hospitalization after enrolment (days)	_	_	_	0.022	1.023 (1.000–1.046)	0.050

OR: odds ratio; CI: confidence interval; ACB: Anticholinergic Cognitive Burden; BMI: body mass index; CCI: Charlson Comorbidity Index; CRP: C-reactive protein; R^2: Nagelkerke R-squared. a. Reference category: female; b. Reference category: university. c. The adjusted model was adjusted for sex, age, educational status, number of medications, BMI, CCI, hospitalization in the past year, in the IM ward, and after enrolment. The adjusted model included all variables shown in the table. Hosmer–Lemeshow test p=0.078. *p<0.05

Table 3. Sensitivity analyses for the association between ACB score and high nutritional risk

		Nutritional Risk		
Model	**Excluded variable(s)**	**β**	**OR (95% CI)**	***p value***
Sensitivity Model 1	Age excluded	0.677	1.968 (1.721-2.250)	**<0.001***
Sensitivity Model 2	BMI excluded	0.672	1.959 (1.713-2.241)	**<0.001***
Sensitivity Model 3	Age and BMI excluded	0.678	1.970 (1.723-2.252)	**<0.001***

OR: odds ratio; CI: confidence interval; BMI: body mass index; ACB: Anticholinergic Cognitive Burden. Sensitivity Model 1 excluded age from the final adjusted model; Sensitivity Model 2 excluded BMI from the final adjusted model; and Sensitivity Model 3 excluded both age and BMI from the final adjusted model. All models were adjusted for the remaining covariates included in the final model. *p<0.05

Discussion

In this retrospective cross-sectional study of 1570 patients hospitalized in the internal medicine ward, the anticholinergic burden was significantly associated with nutritional risk. The nutritional risk increased stepwise across the anticholinergic burden groups, reaching 88.8% among patients with high ACB scores. Similarly, the NRS-2002 scores progressively

increased with higher ACB. Patients with anticholinergic exposure also had higher comorbidity burden, higher medication use, longer hospitalization-related indicators, lower discharge rates, and higher 30-day mortality. A key finding of this study was that the ACB score remained independently associated with a high nutritional risk after adjustment for age, sex, educational status, number of medications, BMI, the Charlson Comorbidity Index, and hospitalization-related variables. These findings suggest that anticholinergic burden may represent an important and potentially modifiable medication-related factor in the nutritional risk profile of internal medicine inpatients.

Our findings are consistent with those of previous studies, suggesting that anticholinergic burden is associated with nutritional risk in older and clinically vulnerable populations. Previous studies by Naharci et al. and Ates-Bulut et al. identified a significant association between anticholinergic burden and the presence of malnutrition or nutritional risk in older adults.[18,24] Nastri et al. found a significant correlation between ACB and nutritional status in patients admitted to a geriatric unit and reported that the ACB score was an independent predictor of nutritional risk.[10] Recently, Simsek et al. reported that a high anticholinergic burden was a positive predictor of nutritional risk according to the GLIM diagnostic criteria, MNA-SF, and NRS-2002.[15] Consistent with previous studies, our study demonstrated a stepwise increase in high nutritional risk across ACB groups. [18,24] Previous studies have reported that a high anticholinergic burden results in a 1.20–2.21-fold increase in nutritional risk, as defined by the MNA-SF.[15,18,24]. Specifically, regarding the NRS-2002 screening tool, when the risk increase reported by Şimşek et al. (OR=1.23, 95% CI: 1.01–1.50) is considered together with the 1.96-fold risk increase (OR=1.96) identified in our study, the negative impact of high anticholinergic burden on nutritional risk is consistently and strongly confirmed with different screening tools. In contrast to numerous prior studies that predominantly concentrated on geriatric populations, the current study broadens this evidence by examining a substantial cohort of adult internal medicine inpatients using the NRS-2002 for assessment of nutritional risk.

Anticholinergic medication exposure may be associated with higher nutritional risk, primarily through their peripheral side effects, including xerostomia, decreased gastrointestinal motility, and appetite loss. These side effects may impair nutrient intake and absorption, often resulting in reduced food consumption.[25,26] Given that decreased food intake is a critical factor in nutritional screening[19,27], the presence of these side effects underscores the importance of monitoring the nutritional status of individuals prescribed anticholinergic medications to identify nutritional risk early.

The findings of the present study suggest that higher albumin levels are associated with a reduced nutritional risk (OR: 0.92), and this relationship is closely linked to patients' clinical profile. Notably, patients without an anticholinergic burden exhibited higher albumin levels and a more favorable overall clinical trajectory (evidenced by a lower CCI and hospitalization) than the other groups. This observation underscores the strong association between optimal nutritional status and clinical well-being.[28] Furthermore, the overall clinical status may mitigate the adverse effects of anticholinergic burden by improving nutritional status and alleviating common anticholinergic burden-related adverse effects.[25,29]

In the literature, the relationship between inflammation and metabolic changes with nutritional risk is well-characterized;[30] however, the small effect size of CRP and uric acid in predicting nutritional risk in this study underscores that these parameters have limited clinical importance and should not be interpreted as strong predictors of nutritional risk on their own. Nevertheless, the negative correlation of ACB score with albumin and its positive correlation with CRP and uric acid may reflect the broader clinical and inflammatory burden of hospitalized patients. Furthermore, the negative association between anticholinergic burden and albumin levels has been previously confirmed by Ates-Bulut et al.[24] The observed associations of ACB score with albumin, CRP, and uric acid may reflect the greater clinical complexity, inflammatory burden, and multimorbidity of patients with higher anticholinergic exposure. However, albumin is influenced by several factors in hospitalized patients, including inflammation, acute illness, hydration status, and hepatic or renal

function; therefore, it should not be interpreted as a specific marker of nutritional status. Although age and BMI may partially overlap with the NRS-2002 assessment, the sensitivity analyses showed that excluding these variables did not materially change the association between ACB score and high nutritional risk.

This study had several limitations. The retrospective and cross-sectional design limits the assessment of reverse causality, thereby constraining the establishment of causal relationships between the anticholinergic burden and the nutritional risk or malnutrition-related diseases. Moreover, specific anticholinergic side effects related to nutrition, such as appetite changes, were not evaluated, restricting our understanding of the direct mechanisms influencing nutritional status. Additionally, reliance on clinical records may have introduced information bias, and unmeasured confounding factors could have influenced our results. Despite these limitations, this large-scale study offers valuable insights into the association between anticholinergic burden and nutritional risk in an adult inpatient population. However, it should be noted that different anticholinergic burden scales may classify medications differently because they vary in their drug lists, scoring systems, and consideration of dose-related exposure. Therefore, the use of only the ACB Scale may limit direct comparability with studies using other anticholinergic burden tools.

In conclusion, this study showed that higher anticholinergic burden was significantly associated with higher nutritional risk among patients hospitalized in internal medicine wards. These findings suggest that anticholinergic burden may be a clinically relevant and potentially modifiable medication-related factor to consider in nutritional risk assessment. Considering clinical parameters such as length of hospital stay and need for intensive care in future studies to better understand the relationship between anticholinergic burden and nutritional risk is essential for more accurate risk assessment and targeted interventions. From a clinical perspective, these findings also have important implications for primary health care, where anticholinergic medications are frequently prescribed and renewed. Routine medication reviews and early nutritional risk screening may help identify nutritional risk in patients before hospitalization. In this context, assessing the

anticholinergic burden with nutritional assessment, such as recent weight loss, reduced food intake, loss of appetite, dry mouth, difficulty in swallowing, and gastrointestinal symptoms, may support timely referral, medication optimization, and nutritional monitoring.

Ethical Considerations: Ethical approval was obtained from the Non-Interventional Clinical Research Ethics Committee of Etlik City Hospital with decision number AEŞH-BADEK1-2026-360 (dated 15/04/2026). All procedures were conducted in accordance with the principles of the Declaration of Helsinki.

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

References

1. Veronese N, Torbahn G, Demurtas J, et al. Interest in meta-research in geriatric medicine: a survey of members of the European Geriatric Medicine Society. *Eur Geriatr Med.* 2020;11(6):1079-83. doi:10.1007/s41999-020-00369-0
2. Bellanti F, Lo Buglio A, Quiete S, Vendemiale G. Malnutrition in Hospitalized Old Patients: Screening and Diagnosis, Clinical Outcomes, and Management. *Nutrients.* 2022;14(4):910. doi:10.3390/nu14040910
3. Saka B, Altınkaynak M, Erten SN, et al. Malnutrition prevalence measurement and nutritional interventions in internal medical departments of Turkish hospitals: results of the LPZ study. *Clin Sci Nutr.* 2020;2(3):97-102. doi:10.5152/ClinSciNutr.2020.21-2-3
4. Bozkuş R, Balaban Barta S. Malnutrition risk and associated risk factors in patients hospitalised in the internal medicine ward: a prospective observational study from a tertiary hospital in Türkiye. *Intern Med J.* 2026:811-21. doi:10.1111/imj.70346
5. D'Amelio P, Rosso B, Fornelli G, et al. Malnutrition reduces quality of life and performance in hospitalized elderly. *OJEMD.* 2014;4:147-57. doi:10.4236/ojemd.2014.4601
6. Vong T, Yanek LR, Wang L, et al. Malnutrition increases hospital length of stay and mortality among adult inpatients with COVID-19. *Nutrients.* 2022;14(6):1310. doi:10.3390/nu14061310
7. Chen Z, Jiang H, He W, et al. The association of nutritional risk screening 2002 with 1-year re-hospitalization and the length of initial hospital stay in patients with heart failure. *Front Nutr.* 2022;9:849034. doi:10.3389/fnut.2022.849034
8. Amelia WR, Leoni AP, Syauqy A, et al. Malnutrition At Hospital Admission And Its Associated Factors In Internal Medicine Inpatients. *JGI.* 2023;11(2):128-39. doi:10.14710/jgi.11.2.128-139

9. Gonen O, Erbakan AN, Cakir IB, Mesci B, Keskinler MV, Oguz A. Malnutrition and inadequate food consumption in hospitalized patients. *Ann Med Res.* 2022:920-4. doi:10.5455/annalsmedres.2022.01.021
10. Nastri M, Bartoli G, De Colle P. Impact of anticholinergic burden on cognitive impairment, disability and malnutrition: a cross-sectional study among hospitalized older patients. *Geriatr Care.* 2019;5(2):52-6. doi:10.4081/gc.2019.8473
11. Halasabalu Kalgeri S, Bheemasamudra Balaraj S, Shivakumar AT, et al. Impact of polypharmacy on oral health in the elderly: challenges and management. *Front Dent Med.* 2026;7:1758771. doi:10.3389/fdmed.2026.1758771
12. Heuberger RA, Caudell K. Polypharmacy and nutritional status in older adults: a cross-sectional study. *Drugs Aging.* 2011;28(4):315-23. doi: 10.2165/11587670-000000000-00000
13. Fiske M, Moen A, Mdala I, Straand J. Malnutrition and Polypharmacy in Older Adult Patients Receiving Home Care Nursing Services: A Cross-sectional Study. *J Am Med Dir Assoc.* 2024;25(3):526-31. doi:10.1016/j.jamda.2023.11.016
14. Lee D-H, Kim B-R, Jun M-K. Effects of polypharmacy on sensory, oral, and dietary functions in older adults. *Explor Med.* 2026;7:1001380. doi:10.37349/emed.2026.1001380
15. Simsek H, Bozkus R, Oguz EO, Ucar A, Naharcı MI. From pills to plate: Anticholinergic burden and malnutrition in older adults across three nutrition assessment tools: A cross-sectional study. *Nutr Clin Pract.* 2026. doi:10.1002/ncp.70090
16. Kok WE, Haverkort EB, Algra YA, et al. The association between polypharmacy and malnutrition(risk) in older people: A systematic review. *Clin Nutr ESPEN.* 2022;49:163-71. doi:10.1016/j.clnesp.2022.03.007

17. Barbosa AAdO, Vicentini AP, Langa FR. Comparison of NRS-2002 criteria with nutritional risk in hospitalized patients. *Cien Saude Colet.* 2019;24:3325-34. doi:10.1590/1413-81232018249.25042017
18. Naharci MI, Katipoglu B, Tasci I. Association of anticholinergic burden with undernutrition in older adults: A cross-sectional study. *Nutr Clin Pract.* 2022;37(5):1215-24. doi:10.1002/ncp.10821
19. Kondrup J, Rasmussen HH, Hamberg O, Stanga Z. Nutritional risk screening (NRS 2002): a new method based on an analysis of controlled clinical trials. *Clin Nutr.* 2003;22(3):321-36. doi:10.1016/s0261-5614(02)00214-5
20. Lisibach A, Benelli V, Ceppi MG, Waldner-Knogler K, Csajka C, Lutters M. Quality of anticholinergic burden scales and their impact on clinical outcomes: a systematic review. *Eur J Clin Pharmacol.* 2021;77(2):147-62. doi:10.1007/s00228-020-02994-x
21. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40(5):373-83. doi: 10.1016/0021-9681(87)90171-8
22. World Health Organization (WHO). Body mass index. [Internet]. 2026; https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/body-mass-index. (Accessed: 10.02.2026).
23. Tabachnick BG, Fidell LS, Ullman JB. *Using multivariate statistics.* Vol 5. 7th ed: Pearson Boston, MA; 2007.
24. Ates Bulut E, Erken N, Kaya D, Dost FS, Isik AT. An Increased Anticholinergic Drug Burden Index Score Negatively Affect Nutritional Status in Older Patients Without Dementia. *Front Nutr.* 2022;9:789986. doi:10.3389/fnut.2022.789986

25. Nishtala PS, Salahudeen MS, Hilmer SN. Anticholinergics: theoretical and clinical overview. *Expert Opin Drug Saf.* 2016;15(6):753-68. doi:10.1517/14740338.2016.1165664

26. Hilmer SN, Gnjidic D. The anticholinergic burden: from research to practice. *Aust Prescr.* 2022;45(4):118-20. doi:10.18773/austprescr.2022.031

27. Serón-Arbeloa C, Labarta-Monzón L, Puzo-Foncillas J, et al. Malnutrition Screening and Assessment. *Nutrients.* 2022;14(12):2392. doi:10.3390/nu14122392

28. Bretschera C, Boesiger F, Kaegi-Braun N, et al. Admission serum albumin concentrations and response to nutritional therapy in hospitalised patients at malnutrition risk: Secondary analysis of a randomised clinical trial. *EClinicalMedicine.* 2022;45:101301. doi:10.1016/j.eclinm.2022.101301

29. Pasina L, Colzani L, Cortesi L, et al. Relation Between Delirium and Anticholinergic Drug Burden in a Cohort of Hospitalized Older Patients: An Observational Study. *Drugs Aging.* 2019;36(1):85-91. doi:10.1007/s40266-018-0612-9

30. Stumpf F, Keller B, Gressies C, Schuetz P. Inflammation and Nutrition: Friend or Foe? *Nutrients.* 2023;15(5):1159. doi:10.3390/nu15051159

ANKARA MEDICAL JOURNAL

Research Article

Ankara Med J, 2026;(3):555-571 // 10.5505/amj.2026.82765

APPROACHES TO PEDIATRIC COUGH: A CROSS-SECTIONAL SURVEY ON DIAGNOSIS AND TREATMENT PRACTICES

Umut Durak[1], Yusuf Deniz[2]

[1]Department of Developmental Pediatrics, University of Health Sciences, Beyhekim Training and Research Hospital, Konya, Türkiye

[2]Department of Pediatrics, Karabük University Faculty of Medicine Training and Research Hospital, Karabük, Türkiye

Correspondence:
Umut Durak (e-mail: dr.durak@hotmail.com)

Submitted: 16.07.2026 // Accepted: 24.08.2026

Ankara Yıldırım Beyazıt University Faculty of Medicine
Department of Family Medicine

ANKARA MEDICAL JOURNAL

Abstract

Objectives: In Türkiye, no nationally recognized guideline exists for the management of cough in children. The absence of standardized protocols may lead to heterogeneous clinical approaches. This study aimed to describe the diagnostic and therapeutic practices of physicians working in pediatric practice and their perceived need for a common guideline.

Materials and Methods: A cross-sectional digital survey was administered in Türkiye between June and December 2023 using snowball sampling. The 30-item form covered demographic characteristics, clinical experience, diagnostic approaches, and treatment preferences. Data were analyzed using descriptive statistics, with exploratory Fisher's exact comparisons.

Results: Sixty-eight physicians working in pediatric practice completed the survey. Most participants (85.29%) preferred observation before initiating treatment. Among the 55 respondents who selected exactly three medications, 36 (65.45%) selected antibiotics among their three most frequently used treatments. Definitions of acute and chronic cough varied widely. Although viral infection was the most reported cause of cough (92.65%), 96.92% reported some use of chest radiography, mostly in a small proportion of patients. Only 14.71% perceived that a national guideline on pediatric cough existed, and 85.29% reported needing further education or resources on cough management.

Conclusion: This descriptive study documents variability in the self-reported practices of pediatricians and a widely reported need for further guidance, in a setting without a national pediatric cough guideline. These hypothesis-generating findings do not establish that any specific practice is unnecessary or inappropriate. They may nonetheless be informative for other settings lacking such a guideline and support the development of standardized, age-specific guidance.

Keywords: Cough, child, practice guidelines as topic, health care surveys, pediatrics.

Introduction

Cough is a major reason for childhood presentations to health care facilities. The creation of pediatric cough guidelines is essential for its management. However, because factors such as the patient's age, pediatric-specific issues, and risk–benefit ratios can affect management, developing a comprehensive guideline remains challenging.[1] For this reason, existing guidelines often define pediatric cough as referring to patients up to 12–18 years of age.[2]

Clinical guidelines increase patient safety by providing physicians with evidence-based, standardized approaches to diagnostic and therapeutic processes. In the United States, the American Academy of Pediatrics (AAP) and the American College of Chest Physicians (ACCP) have published guidelines focusing on the systematic evaluation and management of chronic cough in children.[3,4] The UK NICE and Australia's Royal Children's Hospital have also produced widely used guidelines for childhood cough.[5,6] These guidelines guide symptom assessment, differential diagnosis, treatment, and the need for advanced investigations.

In Türkiye, however, there is no nationally accepted clinical guideline specifically focusing on pediatric cough. Although some guidelines published by the Turkish National Pediatrics Association include general topics related to respiratory system diseases, there is no structured clinical algorithm dedicated to cough.[7] Comparable associations likewise lack a guideline dedicated to cough.[8] Reviews in DergiPark and the TR Index discuss childhood cough but do not qualify as guidelines.[9-11] The Turkish Thoracic Society's adult cough algorithm has limited applicability to children.[12] Furthermore, adapting guidelines developed for adults directly to children poses both theoretical and practical problems.[13,14]

The absence of a shared guideline may lead to heterogeneity in clinical practice, which can result in unnecessary investigations and treatments, resource waste, and, in some cases, diagnostic delay.[15] This study aimed to assess whether there is a need for a common clinical guideline in the management of pediatric cough in Türkiye by collecting expert opinions through a survey. The results may also be relevant to other countries that do not use a shared guideline.

Materials and Methods

Study design and duration

This study was designed as a cross-sectional, descriptive survey. Data collection commenced in June 2023 and prospectively covered six months (June 1, 2023 – December 1, 2023).

Study population and sample

The study population consisted of physicians actively working in pediatric practice in Türkiye, including pediatric residents/research assistants and specialists. Snowball sampling was used: pediatricians in different institutions were directly contacted, invited, and asked to forward the survey to colleagues. The form was distributed digitally via Google Forms. No a priori sample size calculation was performed, as the study was an exploratory, descriptive survey rather than a hypothesis test; the sample reflects the pediatricians who voluntarily completed the questionnaire (n=68), and the resulting limitation is addressed below.

Participant criteria

Inclusion criteria were as follows: actively practicing as a pediatrician in Türkiye; being engaged in clinical or academic practice; and completing the survey voluntarily and in full. Exclusion criteria were as follows: physicians working in medical specialties other than pediatrics; retired or non-practicing pediatricians; and incomplete or invalid survey responses.

Survey development

The survey form used in this study was adapted from the original survey developed by Vogelberg et al. to evaluate approaches to cough treatment in pediatric patients.[2] That survey was published in an open-access journal under a Creative Commons Attribution 4.0 (CC BY 4.0) license, which permits adaptation with appropriate credit; it was therefore

adapted with attribution to the original authors, and no separate permission was required. It was revised for our context, taking into account Türkiye's health care system and clinical practices. The questionnaire was administered in Turkish. Because it was used as a practical data-collection tool rather than a validated scale, no formal forward–backward translation, content validity index calculation, or pilot testing was performed; content appropriateness was assessed qualitatively by three experienced pediatricians. Compared with the original instrument, several items were reworded for local relevance, and new context-specific items (for example, on traditional and complementary practices and on national guideline awareness) were added.

During adaptation, some questions were restructured for cultural and professional relevance, and additional questions were developed to address local practices and the research objectives, in line with clinical approaches and diagnostic algorithms in the literature.

The content validity of the survey form was evaluated by three independent academics, each with at least 10 years of clinical and academic experience in pediatrics, who were not included in the study. Expert opinion was sought regarding content appropriateness, and adjustments were made accordingly before ethical approval was obtained. Because the instrument was adapted by revising existing items and adding new context-specific questions rather than administering the original survey unchanged, the psychometric properties (formal validity and reliability) of the adapted version were not re-established through a separate validation study. This adaptation was intended to descriptively capture current practice patterns rather than to produce a formally validated measurement scale, and the resulting limitation is acknowledged in the Limitations section.

The final 30-item form covered demographics, clinical experience, diagnostic approaches, treatment preferences, and traditional/complementary practices, using multiple-choice, Likert-type, proportional, and open-ended items. The survey aimed to describe pediatricians' self-reported diagnostic and treatment practices and their perceptions regarding the availability of and need for national guidance on pediatric cough.

Data collection

Data were collected using a structured form on Google Forms, shared via e-mail, social media groups, and professional communication networks. Most items were mandatory, largely preventing blank responses. E-mail addresses were collected only to restrict submissions to one per participant and removed before analysis; the dataset contains no identifying information. No reminders were sent. Because of the snowball distribution, the number of physicians who received or opened the link could not be determined, and a participation or completion rate could not be calculated. The study was reported with reference to the CHERRIES recommendations for web-based surveys where applicable.[16]

Ethical considerations

The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Ethical approval was obtained from the Karabük University Non-Interventional Clinical Research Ethics Committee (Decision No: 2023/1411, dated May 16, 2023). Informed consent was obtained from all participants before their inclusion in the study.

Statistical analysis

The collected data were analyzed using Intellectus Statistics online software. Descriptive statistics of continuous variables were presented as mean, standard deviation, minimum, and maximum values. The distribution of categorical variables was reported as frequencies and percentages (%). There were no blank responses. Unless otherwise specified, percentages were calculated over the full sample (n=68); diagnostic-test percentages over the item-specific valid responses (n=63–67), and medication-preference percentages among the 55 respondents who selected exactly three options. The two open-ended items, on indications for advanced diagnostic evaluation and criteria for initiating treatment, were coded into recurring themes by the authors (not independent coders); answers not fitting these were grouped as “Other / not classified.” For the interval-based diagnostic-test items, responses containing more than one interval were considered invalid for that specific test

and excluded item-wise, and interval midpoints were used to estimate median patient percentages. The item asking for the three most frequently used medications was analyzed among the 55 respondents who selected exactly three options. Institution names were grouped into three anonymized categories (state/training and research, university, private); unclassifiable responses are reported separately. In addition, as an exploratory analysis, associations between selected practice-related outcomes (self-reported need for further education or resources; inclusion of antibiotics among the three most frequently used treatments; and awareness of a guideline) and physician characteristics (years of experience, dichotomized as ≤5 vs >5 years; and academic title, dichotomized as resident/research assistant vs specialist and above) were examined using Fisher's exact test, given the small sample size. To account for multiple comparisons, p-values were adjusted using the Benjamini-Hochberg false discovery rate (FDR) procedure. These analyses were considered exploratory and hypothesis-generating, and adjusted p-values <0.05 were regarded as statistically significant. The study was reported in accordance with the STROBE guideline for cross-sectional studies.[17]

Results

Sociodemographic characteristics of participants

Analysis: The age distribution of participating physicians showed that 42.65% were aged 30–40 years, 30.88% were 40–50 years, 23.53% were under 30 years, and 2.94% were over 50 years.

The distribution of academic titles was 47.06% specialists, 27.94% residents/research assistants, 10.29% subspecialists/subspecialty fellows, 7.35% assistant professors, 5.88% associate professors, and 1.47% professors. Institutions were mainly training and research hospitals or state hospitals (48.53%) and university hospitals (42.65%); 5.88% were private practice, and two (2.94%) were unclassifiable.

Regarding publications in the field of pediatric cough, 92.65% had not published, while 7.35% had published at least one study. The rate of participation in congresses related to pediatric cough was 57.35%, while 42.65% had not attended. The distribution of years in pediatrics showed that 35.29% had more than 10 years, 17.65% had 5–10 years, 13.24% had 3–5 years, 26.47% had 1–3 years, and 7.35% had less than 1 year (Table 1).

Table 1. Sociodemographic characteristics and clinical experience of the participating pediatricians

Variable	Category	n (%)
Age distribution	Under 30 years	16 (23.53)
	30–40 years	29 (42.65)
	40–50 years	21 (30.88)
	Over 50 years	2 (2.94)
Academic title	Specialist	32 (47.06)
	Resident/research assistant	19 (27.94)
	Subspecialist/subspecialty fellow	7 (10.29)
	Assistant professor	5 (7.35)
	Associate professor	4 (5.88)
	Professor	1 (1.47)
Institution type	State/training and research hospital	33 (48.53)
	University hospital	29 (42.65)
	Private hospital/clinic	4 (5.88)
	Other / not classified	2 (2.94)
Publication on pediatric	Yes	5 (7.35)
	No	63 (92.65)
Attendance at pediatric	Yes	39 (57.35)
	No	29 (42.65)
Years of experience in	Less than 1 year	5 (7.35)
	1–3 years	18 (26.47)
	3–5 years	9 (13.24)
	5–10 years	12 (17.65)
	More than 10 years	24 (35.29)

Data are presented as numbers (percentages). n: number of participants.

Clinical experience and observations of pediatricians regarding pediatric cough

In the distribution of patients presenting with cough within the past year, 60.29% of physicians reported having seen more than 300 patients, 14.71% between 100 and 300, 11.76% between 50 and 100, and 13.24% fewer than 50 patients.

For patients under two years of age, 55.88% reported having seen more than 100, 14.71% between 30 and 60, 13.24% between 60 and 100, and 16.18% fewer than 30 patients. The most common age group encountered in clinical practice was children aged 2–5 years, at 57.35%.

The most frequent cause of cough was reported as viral infection (92.65%), followed by allergic responses (5.88%) and asthma (1.47%). A total of 86.76% stated that distinguishing between wet and dry cough was useful in diagnosis.

Regarding the duration of acute cough, 38.24% considered it shorter than 2 weeks, 26.47% shorter than 1 week, and 20.59% shorter than 4 weeks. For chronic cough, 47.06% defined it as longer than 4 weeks, 19.12% longer than 3 weeks, and 14.71% longer than 2 weeks.

The most important criteria for advanced diagnostic evaluation were reported as accompanying findings (35.29%), duration of cough (30.88%), recurrent or high-risk conditions (17.65%), lack of treatment response (8.82%), and growth and developmental delay (7.35%).

Diagnostic methods, treatment approaches, and clinical protocols

Respondents indicated the approximate percentage of cough patients in whom each test was requested. Responses selecting more than one band were excluded per test (valid n = 63–67). Among valid responses, the proportion reporting any use (>0% of patients) was: chest radiography 96.92%, viral panel 86.15%, allergy testing 85.71%, CT 72.73%, serology 73.13%, sputum culture 59.70%, spirometry 53.03%, quadruple test 46.97%, and bronchoscopy 41.79%. However, the median percentage of patients tested was low for most

tests (Table 2), indicating that, apart from chest radiography, these investigations were requested in only a small proportion of patients.

Table 2. Diagnostic and treatment approaches of the participating pediatricians to pediatric cough

Clinical parameter	Response option	n (%)
Diagnostic tests, any use (>0% of patients)	Chest radiography	63/65 (96.92)
	Viral panel	56/65 (86.15)
	Serology	49/67 (73.13)
	Allergy testing	54/63 (85.71)
	Thoracic CT	48/66 (72.73)
	Sputum culture	40/67 (59.70)
	Spirometry	35/66 (53.03)
	Quadruple test	31/66 (46.97)
	Bronchoscopy	28/67 (41.79)
Perceived existence of a national guideline	Exists	10 (14.71)
	Does not exist	32 (47.06)
	Not sure	25 (36.76)
	Uses the adult algorithm	1 (1.47)
Institutional protocol presence	Exists	16 (23.53)
	Does not exist	38 (55.88)
	Not sure	14 (20.59)
Initial treatment approach	Observation first	58 (85.29)
	Immediate treatment	7 (10.29)
	Other / not classified	3 (4.41)
Criteria for starting treatment	Accompanying findings	30 (44.12)
	Lack of treatment response	18 (26.47)
	Quality of life affected	11 (16.18)
	Other / not classified	9 (13.24)
Age group receiving direct treatment	All ages (still prefer observation)	40 (58.82)
	Direct treatment in 0–3 months group	20 (29.41)
	Direct treatment in 2–5 years group	3 (4.41)
	Direct treatment in 0–1 year group	2 (2.94)
	Direct treatment in 0–6 months group	2 (2.94)

Clinical parameter	Response option	n (%)
	Direct treatment in 1–2 years group	1 (1.47)
Medication use preferences	Bronchodilators	37 (67.27)
	Antihistamines	33 (60.00)
	Antibiotics	36 (65.45)
	Nasal/inhaled corticosteroids	28 (50.91)
	Antitussives	5 (9.09)
Perceived need for further education or resources	Yes	58 (85.29)
	No	8 (11.76)
	Undecided	2 (2.94)

Data are presented as numbers (percentage), calculated over the full sample (n=68) unless otherwise noted; there were no blank responses, but non-conforming responses to the interval-based and three-item questions were handled as described below and in the Statistical Analysis section. For diagnostic tests, respondents reported the approximate percentage of their cough patients in whom each test was requested; a small number of respondents selected more than one percentage band for a given test, and these ambiguous responses were excluded on a per-test basis (valid n = 63–67); the figures shown are the number and proportion of valid respondents reporting any use (use in more than 0% of patients). The median percentage of patients tested was low for most tests (chest radiography 12.5%, allergy testing 7.5%, viral panel 7.5%, serology 3.0%, thoracic CT 3.0%, sputum culture 3.0%, spirometry 3.0%, quadruple test 0%, bronchoscopy 0%). Medication preferences reflect the three most frequently used treatments; this item was analyzed among the 55 respondents who selected exactly three medications, and percentages do not total 100%. For the open-ended item on criteria for starting treatment, free-text answers were coded into the main themes shown, and answers that did not fit these themes are reported as "Other / not classified"; for the initial treatment approach and the age group receiving direct treatment, all response categories are listed so that the relevant single-choice and coded open-ended variables total 68. CT: computed tomography; n: number of participants.

The item specifically asked whether a national clinical guideline for pediatric cough existed in the respondent's country; it did not assess awareness of international guidelines. In response, 47.06% believed no national guideline existed, 36.76% were unsure, 14.71% thought one existed, and 1.47% reported using the adult algorithm. As for the presence of a cough treatment protocol in their institutions, 23.53% reported that one existed, 55.88% reported that it did not, and 20.59% were unsure.

Regarding the initiation of treatment, 85.29% preferred to observe before starting therapy, while 10.29% initiated treatment immediately. The main criteria for initiating treatment were accompanying findings (44.12%), lack of treatment response (26.47%), and effects on

quality of life (16.18%). Among those who were asked about direct treatment without observation, 58.82% stated they would still observe in all age groups first, while 29.41% reported applying direct treatment in the 0–3-month age group.

For the three-most-used-medications item, 13 respondents selected a different number, so it was analyzed among the 55 compliant respondents. The most commonly selected were bronchodilators (67.27%), antibiotics (65.45%), and antihistamines (60.00%), followed by nasal/inhaled corticosteroids (50.91%); antitussives were selected by only 9.09%. In a separate item, 54.41% reported not using antitussives at all and 27.94% not using antibiotics.

When participants were asked whether they needed further education or resources regarding the management of pediatric cough, 85.29% (n=58) answered affirmatively, 11.76% (n=8) answered negatively, and 2.94% (n=2) were undecided (Table 2). Regarding traditional and complementary approaches, 83.82% (57/68) reported preferring traditional methods and 91.18% (62/68) reported using nasal irrigation.

Exploratory associations with physician characteristics

In exploratory analyses using Fisher's exact test across six comparisons with Benjamini-Hochberg correction, three associations remained statistically significant after adjustment. The self-reported need for further education or resources was more frequent among physicians with ≤5 years of experience than among those with >5 years (100% vs 76.5%; adjusted p=0.022), although it did not differ significantly by academic title (adjusted p=0.140). The inclusion of antibiotics among the three most frequently used treatments (n=55) was more frequent among less experienced physicians (82.1% vs 48.1%; adjusted p=0.022) and among residents/research assistants versus specialists and above (93.3% vs 55.0%; adjusted p=0.022). Awareness of a national guideline was not significantly associated with experience or title, and no other comparison reached significance after adjustment (Table 3). Given the small sample size and the exploratory nature of these analyses, these associations should be interpreted with caution and regarded as hypothesis-generating.

Table 3. Exploratory associations between physician characteristics and selected practice-related outcomes (Fisher's exact test)

Outcome/subgroup	Outcome present, n (%)	p	Adjusted p (FDR)
Need for further education or resources.	.		
≤5 years of experience (n=32)	32 (100)	0.005	0.022
>5 years of experience (n=34)	26 (76.5)		
Resident/research assistant (n=19)	19 (100)	0.093	0.140
Specialist and above (n=47)	39 (83.0)		
Antibiotics among top-3 treatments (n=55)			
≤5 years of experience (n=28)	23 (82.1)	0.011	0.022
>5 years of experience (n=27)	13 (48.1)		
Resident/research assistant (n=15)	14 (93.3)	0.010	0.022
Specialist and above (n=40)	22 (55.0)		
Perceived existence of a national guideline			
≤5 years of experience (n=32)	5 (15.6)	1.000	1.000
>5 years of experience (n=36)	5 (13.9)		
Resident/research assistant (n=19)	2 (10.5)	0.714	0.857
Specialist and above (n=49)	8 (16.3)		

Percentages refer to the proportion within each physician subgroup. Six comparisons (three outcomes × two physician characteristics) were tested, and p-values were adjusted for these six comparisons using the Benjamini-Hochberg false discovery rate procedure. For the outcome "need for further education or resources," the two undecided responses were excluded from this binary comparison (both in the specialist-and-above and >5-year groups), giving subgroups of 34 (>5 years) and 47 (specialist and above) for that outcome. The antibiotics outcome was analyzed among the 55 respondents who selected exactly three medications, as instructed. These analyses are exploratory and hypothesis-generating. Experience was dichotomized as ≤5 vs >5 years; academic title as resident/research assistant vs specialist and above (subspecialty fellows, who have completed pediatric residency, were classified as specialists). FDR: false discovery rate.

Discussion

One of the first stages in the management of pediatric cough is the classification of cough as acute or chronic. This classification is critically important for differential diagnosis and treatment planning, as it allows benign causes such as viral infections to be distinguished

from more serious underlying conditions.[18] In the guidelines published by the AAP and the ACCP, acute cough in children is defined as lasting less than 4 weeks, while chronic cough is defined as lasting more than 4 weeks, and diagnostic algorithms are structured according to these thresholds.[3,4]

However, our study identified a lack of consensus on these definitions: only 47.06% defined chronic cough as lasting more than 4 weeks, and only 20.59% stated that acute cough lasted less than 4 weeks, whereas guidelines provide clear duration-based definitions.[3-6]

This classification helps avoid unnecessary investigations; imaging is generally recommended only for chronic cough when clinically indicated.[3-6] In our survey, chest radiography was the only test that most respondents reported using in a substantial share of patients (any use 96.92%; median 12.5% of patients). For most other tests, the median percentage tested was low, suggesting these were reserved for selected patients rather than applied routinely. These self-reported figures cannot establish over-investigation; nevertheless, the relatively widespread use of chest radiography, where viral infection was the most reported cause (92.65%), may reflect diagnostic uncertainty, parental expectations, or medico-legal concerns rather than guideline-based practice.

Similarly, the clinical practice guidelines prepared by the Royal Children's Hospital in Australia recommend avoiding unnecessary imaging and laboratory tests.[6] In children with sore throat and viral symptoms, throat swabs are unlikely to change treatment.[19]

When making treatment decisions, most participants (85.29%) reported preferring observation before initiating therapy, yet antibiotics were selected among the three most frequently used treatments by 36 of the 55 compliant respondents (65.45%). The frequent reporting of antibiotics, where viral infection was the most common reported cause, may reflect diagnostic uncertainty, parental expectations, or local prescribing habits rather than guideline-based reasoning. For reference, the AAP and the American Academy of Family Physicians recommend that antibiotics be reserved for patients with suspected bacterial infection,[4,20] the Canadian Paediatric Society notes that antitussives provide no clinically

meaningful benefit in children and may be harmful,[21] and the ACCP guideline does not recommend their routine use for chronic cough in children.[1]

Notably, 85.29% of participants reported a need for additional education or resources, more often among less experienced physicians, who also selected antibiotics more frequently. Together with the lack of consensus in cough definitions and diagnostic evaluation, these observations are consistent with the absence of a guideline-based algorithm, though they cannot establish causation. The European Respiratory Society recommends that treatment decisions follow algorithms based on symptom duration, accompanying findings, and risk factors.[22] Age-specific, evidence-based guidance may help reduce this variability. The questionnaire also covered traditional and complementary approaches: a majority reported preferring traditional methods (57/68) and using nasal irrigation (62/68), as described in the Results.

Limitations

Several limitations should be noted. The snowball sampling method, based on participant referrals within their own networks, may have limited the representativeness of the sample, and the geographic distribution of respondents was not recorded. Because the number of physicians reached could not be determined, a participation rate could not be calculated, and selection bias cannot be excluded. As a self-reported survey not linked to prescription or patient records, the findings reflect stated rather than verified practice. The relatively small sample (n=68) further restricts precision, and the exploratory subgroup analyses were limited in statistical power; they should be regarded as hypothesis-generating rather than confirmatory. A further limitation concerns the instrument itself: because the questionnaire was adapted by revising existing items and adding new context-specific questions, the validity and reliability of the adapted version were not formally re-established through a dedicated validation study, which may have affected how some items were interpreted. For these reasons, the results should be read as descriptive indicators of self-reported practice, and it is difficult to claim that they represent all pediatricians in Türkiye.

In conclusion, within the constraints of its descriptive design and limited sample, this study suggests that heterogeneous practices may exist in the management of pediatric cough in Türkiye. The variability observed in the definitions, diagnostic preferences, and treatment approaches reported by participants may reflect differences in clinical interpretation, local protocols, case mix, access to diagnostic resources, or varying awareness of existing recommendations, rather than a single identifiable cause. Because the data are self-reported and were not linked to prescription records or patient-level outcomes, these findings should be interpreted as hypothesis-generating rather than confirmatory. Our findings may nonetheless be informative for other settings that also lack a comprehensive guideline for pediatric cough. Larger studies with broader participation from different regions are needed to substantiate these observations. If such studies yield consistent results, academic organizations, particularly the Turkish National Pediatrics Association, may consider prioritizing the development of age-specific and evidence-based guidance, which could in turn inform an internationally structured framework organized according to age groups.

Ethical Considerations: The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Ethical approval was obtained from the Karabük University Non-Interventional Clinical Research Ethics Committee (Decision No: 2023/1411, dated May 16, 2023). Informed consent was obtained from all participants before their inclusion in the study.

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

References

1. Chang AB, Oppenheimer JJ, Irwin RS; CHEST Expert Cough Panel. Managing chronic cough as a symptom in children and management algorithms: CHEST guideline and expert panel report. Chest. 2020;158:303-29 doi:10.1016/j.chest.2020.01.042.
2. Vogelberg C, Cuevas Schacht F, Watling CP, Upstone L, Seifert G. Therapeutic principles and unmet needs in the treatment of cough in pediatric patients: review and expert survey. BMC Pediatr. 2023;23:34 doi:10.1186/s12887-022-03814-0.
3. Kasi AS, Kamerman-Kretzmer RJ. Cough. Pediatr Rev. 2019;40:157-67 doi:10.1542/pir.2018-0116.
4. Chang AB, Glomb WB. Guidelines for evaluating chronic cough in pediatrics: ACCP evidence-based clinical practice guidelines. Chest. 2006;129(1 Suppl):260S-283S (doi:10.1378/chest.129.1_suppl.260S).
5. National Institute for Health and Care Excellence (NICE). Cough (acute): antimicrobial prescribing. NICE guideline NG120 [Internet]. 2019. https://www.nice.org.uk/guidance/ng120. (Accessed: 06.09.2025).
6. Royal Children's Hospital Melbourne. Cough: clinical practice guideline [Internet]. https://www.rch.org.au/clinicalguide/guideline_index/cough/. (Accessed: 06.09.2025).
7. Turkish National Pediatric Society. Guidelines [Internet]. https://millipediatri.org.tr/kategori/menu/klavuzlar-131. (Accessed: 06.09.2025).
8. Turkish Pediatric Association. Official documents and clinical resources [Internet]. https://turkpediatri.org.tr/. (Accessed: 06.09.2025).
9. İğde M, Öksüz BG. Çocuklarda öksürük: klinik yaklaşım. Güncel Pediatri. 2015;13:201-8 doi:10.4274/jcp.93064.
10. Tüsüz Önata E, Dikici Ü, Özdemir Ö. Çocuklarda kronik öksürüğe yaklaşım. Sağlık Bilimlerinde Değer. 2024;14:451-7 doi:10.33631/sabd.1409933.
11. Turgay Yağmur İ. Evaluation of patients presenting with chronic cough to the pediatric allergy clinic according to age groups. Genel Tıp Derg. 2025;35:155-61 doi:10.54005/geneltip.1582368.

12. Turkish Thoracic Society. Cough algorithm for adults [Internet]. https://toraks.org.tr/. (Accessed: 06.09.2025).
13. Jiang M, Guan WJ, Fang ZF, et al. A critical review of the quality of cough clinical practice guidelines. Chest. 2016;150:777-88 doi:10.1016/j.chest.2016.04.028.
14. Chang AB. Pediatric cough: children are not miniature adults. Lung. 2010;188(Suppl 1):S33-40 doi:10.1007/s00408-009-9166-2.
15. Chang AB, Robertson CF, Van Asperen PP, et al. A cough algorithm for chronic cough in children: a multicenter, randomized controlled study. Pediatrics. 2013;131:e1576-83 doi:10.1542/peds.2012-3318.
16. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004;6(3):e34 doi:10.2196/jmir.6.3.e34.
17. Cuschieri S. The STROBE guidelines. Saudi J Anaesth. 2019;13(Suppl 1):S31-4 doi:10.4103/sja.SJA_543_18.
18. Kang C. Acute cough remedies for children: so many options, not enough evidence. Drugs Ther Perspect. 2021;37:358-62 doi:10.1007/s40267-021-00848-9.
19. Choosing Wisely Canada. Paediatrics: recommendations [Internet]. Toronto: Canadian Paediatric Society; 2024. https://choosingwiselycanada.org/recommendation/paediatrics/. (Accessed: 09.09.2025).
20. Gonzales R, Bartlett JG, Besser RE, Hickner JM, Hoffman JR, Sande MA. Principles of appropriate antibiotic use for treatment of uncomplicated acute bronchitis: background. Am Fam Physician. 2012;86(2):153-5.
21. Goldman RD; Canadian Paediatric Society, Drug Therapy and Hazardous Substances Committee. Treating cough and cold: guidance for caregivers of children and youth. Paediatr Child Health. 2011;16(9):564-9 doi:10.1093/pch/16.9.564.
22. Morice AH, Millqvist E, Bieksiene K, et al. ERS guidelines on the diagnosis and treatment of chronic cough in adults and children. Eur Respir J. 2020;55(1):1901136 doi:10.1183/13993003.01136-2019.