e-ISSN: 2147-2092

GAZI MEDICAL JOURNAL

2026
April
Volume 37 • Issue 2

galenos yayınevi

medicaljournal.gazi.edu.tr

GAZI MEDICAL JOURNAL

medicaljournal.gazi.edu.tr

Editorial Team

GAZI MEDICAL JOURNAL

medicaljournal.gazi.edu.tr

Section Editors and Scientific Editorial Board

Ethical Board

Statistical Board

Scientific Advisory Board

GAZI MEDICAL JOURNAL

medicaljournal.gazi.edu.tr

Please refer to the journal's webpage (https://medicaljournal.gazi.edu.tr/) for "About the Journal" and "Submissions".

The editorial and publication process of Gazi Medical Journal are shaped in accordance with the guidelines of the International Committee of Medical Journal Editors (ICMJE), World Association of Medical Editors (WAME), Council of Science Editors (CSE), Committee on Publication Ethics (COPE), European Association of Science Editors (EASE), and National Information Standards Organization (NISO). The journal is in conformity with the Principles of Transparency and Best Practice in Scholarly Publishing.

Gazi Medical Journal is indexed in Emerging Sources Citation Index, Scopus, Directory of Open Access Journals, EuroPub, Islamic World Science Citation Center, ABCD Index.

The online published articles are freely available on the public internet.

Owner: Musa Yıldız on Behalf of Gazi University

Responsible Manager: Mehmet Ali Ergün

Publisher Contact

Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1 34093 İstanbul, Türkiye

Phone: +90 (539) 307 32 03

E-mail: info@galenos.com.tr/yayin@galenos.com.tr

Web: www.galenos.com.tr

Publisher Certificate Number: 14521

Publication Date: March 2026

e-ISSN: 2147-2092

International scientific journal published quarterly.

GAZI MEDICAL JOURNAL

medicaljournal.gazi.edu.tr

CONTENTS

Original Investigations - Özgün Araştırmalar

GAZI MEDICAL JOURNAL

medicaljournal.gazi.edu.tr

CONTENTS

Case Reports - Olgu Sunumları

GAZI MEDICAL JOURNAL

medicaljournal.gazi.edu.tr

CONTENTS

Letters to the Editor - Editöre Mektuplar

Literature Review with Cases | Olgularla Literatür İncelemeleri

Gazi Med J 2026;37(2):147-151

DOI: http://dx.doi.org/10.12996/gmj.2025.4476

Evaluation of the Relationship Between Postoperative Surgical Complications and Caudal Epidural Block in Pediatric Hypospadias Surgery

Pediatrik Hipospadias Cerrahisinde Postoperatif Cerrahi Komplikasyonlar ile Kaudal Epidural Blok İlişkisinin İncelenmesi

Cem Kaya[1], Gökhan Arkan[1], Gonca Güney Keçeli[1], Ramazan Karabulut[1], Zafer Türkyılmaz[1], Kaan Sönmez[1], Berrin Işık[2]

[1]Department of Pediatric Surgery, Gazi University Faculty of Medicine, Ankara, Türkiye
[2]Department of Anesthesiology and Reanimation, Gazi University Faculty of Medicine, Ankara, Türkiye

ABSTRACT

Objective: The use of caudal epidural block (CEB) for perioperative analgesia in hypospadias surgery has yielded conflicting results regarding its impact on postoperative complications. Our aim is to evaluate the relationship between postoperative surgical complications and the use of CEB in pediatric hypospadias surgery, specifically assessing the impact of CEB across different surgical techniques.

Methods: This retrospective study analyzed pediatric patients who underwent hypospadias surgery under general anesthesia, with or without CEB, at Gazi University between January 2017 and December 2021. Data from 174 patients, aged 7 to 143 months, were reviewed. Parameters included patient age, level of the ectopic external urethral meatus, surgical technique, postoperative complications, and anesthetic method. Statistical analyses were performed using IBM SPSS V22.0.

Results: Among the patients, 144 (82.8%) received CEB and 30 (17.2%) did not. No serious CEB-related complications were observed. Postoperative complications were recorded in 28.7% of cases. The study found no significant difference in complication rates between patients with and without CEB. The frequency of complications varied significantly with the level of the external urethral meatus. The study indicates that the choice of anesthetic method, including the use or nonuse of CEB, does not significantly affect postoperative complications following pediatric hypospadias surgery. While CEB

ÖZ

Amaç: Hipospadias cerrahisinde perioperatif analjezi amacıyla kaudal epidural blok (KEB) kullanımının, postoperatif komplikasyonlar üzerindeki etkisine ilişkin literatürde çelişkili sonuçlar bulunmaktadır. Bu çalışmanın amacı, pediatrik hipospadias cerrahisinde KEB kullanımı ile postoperatif cerrahi komplikasyonlar arasındaki ilişkiyi değerlendirmek ve farklı cerrahi tekniklerde KEB'nin etkisini incelemektir.

Yöntemler: Bu retrospektif çalışmada, Ocak 2017–Aralık 2021 tarihleri arasında Gazi Üniversitesi'nde genel anestezi altında, KEB uygulanarak veya uygulanmadan hipospadias cerrahisi geçiren pediatrik hastalar analiz edildi. Yaşları 7–143 ay arasında değişen toplam 174 hastanın verileri değerlendirildi. İncelenen parametreler; hasta yaşı, ektopik eksternal üretral meatusun seviyesi, uygulanan cerrahi teknik, postoperatif komplikasyonlar ve anestezi yöntemi idi. İstatistiksel analizler IBM SPSS V22.0 yazılımı kullanılarak yapıldı.

Bulgular: Hastaların 144'üne (%82,8) KEB uygulanırken, 30'una (%17,2) uygulanmadı. KEB'ye bağlı ciddi bir komplikasyon saptanmadı. Postoperatif komplikasyonlar olguların %28,7'sinde gözlendi. KEB uygulanan ve uygulanmayan hastalar arasında komplikasyon oranları açısından anlamlı bir fark bulunmadı. Komplikasyon sıklığının, eksternal üretral meatusun seviyesine göre anlamlı farklılık gösterdiği saptandı. Bulgular, KEB uygulanıp uygulanmamasının da dahil olduğu anestezi yönteminin, pediatrik hipospadias cerrahisi sonrası postoperatif komplikasyonları anlamlı düzeyde etkilemediğini göstermektedir. KEB

Cite this article as: Kaya C, Arkan G, Güney Keçeli G, Karabulut R, Türkyılmaz Z, Sönmez K, Işık B. Evaluation of the relationship between postoperative surgical complications and caudal epidural block in pediatric hypospadias surgery. Gazi Med J. 2026;37(2):147-151

Address for Correspondence/Yazışma Adresi: Gökhan Arkan, Department of Pediatric Surgery, Gazi University Faculty of Medicine, Ankara, Türkiye
E-mail / E-posta: gokhanarkan@gazi.edu.tr
ORCID ID: orcid.org/0000-0002-1429-4690

Received/Geliş Tarihi: 21.06.2025
Accepted/Kabul Tarihi: 03.12.2025
Epub: 02.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

This paper was presented as a poster at the 26th Annual European Pediatric Surgeons' Association (EUPSA) Congress, held in Dubrovnik, Croatia, on 21–24 May 2025.

ABSTRACT

provides effective analgesia, its association with complications is not statistically significant.

Conclusion: CEB is a beneficial analgesic technique in pediatric hypospadias surgery and is not associated with a significant increase in postoperative complications. Its use should be considered for its analgesic benefits.

Keywords: Anesthesia, caudal, epidural, conduction-blocking, hypospadias, postoperative complications

ÖZ

etkili bir analjezi sağlamakla birlikte, komplikasyonlarla istatistiksel olarak anlamlı bir ilişki göstermemiştir.

Sonuç: KEB, pediatrik hipospadias cerrahisinde yararlı bir analjezik tekniktir ve postoperatif komplikasyonlarda anlamlı bir artış ile ilişkili değildir. Analjezik faydaları göz önünde bulundurularak kullanımının değerlendirilmesi önerilir.

Anahtar Sözcükler: Anestezi, kaudal, epidural, iletim bloğu, hipospadias, postoperatif komplikasyonlar

INTRODUCTION

Hypospadias, one of the most common congenital anomalies in males, is a malformation in which the external urethral meatus is located proximally and ectopically (1). Hypospadias occurs in approximately 1 in 200–300 male births and is the second most common congenital anomaly in males, after undescended testes (2). The diagnosis of hypospadias is made by physical examination, during which chordee and preputial defects are frequently observed (3).

One of the most commonly used classifications for hypospadias was proposed by Duckett. According to this classification, hypospadias is divided into two groups: distal and proximal. Distal hypospadias is observed in 70% of cases, in which the ectopic meatus is located on the glans, corona, or distal shaft. Proximal hypospadias, seen in 30% of cases, is characterized by an ectopic meatus located on the mid-penile shaft, proximal shaft, penoscrotal region, or perineal region (4).

Caudal epidural block (CEB) is a commonly used regional anesthetic and analgesic technique that can provide perioperative and postoperative analgesia. It can be useful in pediatric patients undergoing sub-umbilical procedures, including inguinal hernia repair, hypospadias surgery, and lower extremity procedures. CEB can be used alone or serve as a supplement to general anesthesia for these procedures (5). Complications associated with CEB include headache, systemic toxicity from local anesthetics, transient neurological symptoms, cauda equina syndrome, and wound pain (6).

In hypospadias surgery, postoperative complication rates range from 5–10% for distal hypospadias to 15–90% for proximal hypospadias. The most common complications include fistula, meatal stenosis, dehiscence of the wound or glans, persistent chordee, and diverticulum (7).

Conflicting results regarding the effect of CEB on postoperative complications in hypospadias cases have been reported (8,9). In our study, we aimed to contribute to the literature by reporting data from our center and to evaluate separately the effects of different surgical techniques.

MATERIALS AND METHODS

After approval was obtained from the Clinical Research Ethics Committee of Gazi University, patients who underwent surgery for hypospadias between January 2017 and November 2021 were retrospectively evaluated. A total of 174 ASA I patients who underwent hypospadias surgery and received general anesthesia, with either CEB or penile block (PB), were included in the study. The patients were followed for a minimum of 6 months.

Patient data were reviewed to determine patients' ages, hypospadias levels, anesthesia methods, hypospadias repair types, and complications. Levels of hypospadias were classified as proximal, midpenile, subcoronal, coronal, or glanular in each patient. Anesthesia methods were divided into two groups: CEB and PB; both groups underwent general anesthesia.

According to our clinic's standard operating procedure, CEBs are performed by an experienced anesthesiologist in no more than two attempts. If CEB cannot be performed within the first two attempts, it is performed under USG guidance. PBs are performed by surgeons. For CEB, 0.8 mL/kg of 0.20% bupivacaine is administered, with a maximum total volume of 20 mL.

Surgery was performed using meatotomy, meatal advancement and glanuloplasty incorporated (MAGPI), tubularized incised plate urethroplasty (TIPU), the Mathieu procedure, skin-flap methods, and graft methods. It was evaluated whether postoperative complications such as recurrence, fistula, and urethral stricture were related to CEB.

Statistical Analysis

SPSS Version 22.0 for Windows (IBM Corp., Armonk, NY) was used for statistical analysis. Chi-square and Fisher's exact tests were used to compare categorical variables between patients who received CEB and PB, and between patients with and without complications. Chi-square and Fisher's exact tests were used to compare the presence or absence of complications between patients who received CEB and those who received PB.

RESULTS

Data from 174 patients were analyzed. Patient ages ranged from 7 to 143 months (mean ± standard deviation, 4.70 ± 2.61 years). The Ectopic external urethral meatus was glandular in 52 patients (29.9%), coronal in 55 (31.6%), subcoronal in 33 (19.0%), midpenile in 12 (6.9%), and penoscrotal in 22 (12.6%). The mean age of the 30 patients who received PB was 6.17 ± 2.87, whereas the mean age of the 144 patients who received CEB was 4.40 ± 2.46; this difference was statistically significant ($p < 0.01$). Patients were evaluated for postoperative surgical complications, including recurrence, fistula, and urethral stricture. Complications occurred in 7 of the 30 patients who received PB and in 43 patients who received CEB, with no statistically significant difference between the two groups ($p = 0.472$). When evaluating the relationship between CEB and PB and the frequency of complications across surgical techniques, no statistically significant differences were found for any individual

technique. The MAGPI procedure, one of the most frequently performed techniques, was associated with complications, including recurrence (n = 1), fistula (n = 2), and urethral stricture (n = 2), in 11.62% of the 43 patients who received CEB. No complications occurred in the 11 patients who received PB; the difference was not statistically significant (p = 0.571) (Table 1).

Ethical approval was obtained from the Gazi University Clinical Research Ethics Committee (decision number: 219, dated: 27.12.2021).

DISCUSSION

Our study found no significant association between anesthetic method (with or without CEB) and postoperative complications in pediatric hypospadias surgery. According to our literature review, a prospective study published in September 2011 was the first to suggest that CEB causes vasodilation and venous pooling, leading to increased bleeding and penile edema in the surgical field, increased tension in sutures, and impaired wound healing, thereby increasing urethral fistula formation (10). Later, Zaidi et al. (11) reported that, in a randomly selected group of operated hypospadias patients at their center, the use of CEB was not significantly associated with fistula development. In the next study, Kreysing and Höhne (12), who evaluated all hypospadias patients over a five-year period, reported that CEB was not statistically associated with fistula development. However, a subsequent study by Kim and colleagues reported that the use of CEB in hypospadias surgeries performed with the TIPU technique significantly increased the frequency of postoperative complications (13). Saavedra-Belaunde et al. (14), who examined the effect of CEB on complications in distal hypospadias surgery, similarly reported that it increased the frequency of postoperative complications over a five-year period. In the same year, Braga et al. (15) and colleagues reported that CEB did not alter the frequency of complications in patients undergoing the TIPU technique, whereas Taicher and colleagues reported that CEB increased postoperative complications thirteenfold. They further stated that its application should be carefully evaluated until further research is conducted, which may increase ethical dilemmas for clinicians (8).

The first meta-analysis on this subject reported no significant relationship between CEB and postoperative complications. However, two meta-analyses published approximately six months later reported that CEB increased the frequency of postoperative surgical complications (16-18).

While studies continued to report conflicting findings regarding whether CEB increases postoperative surgical complications of hypospadias, Ngoo et al. (19), in 2020, highlighted a different point by showing that PB significantly increased the need for postoperative revision surgery. Following this publication, Braga et al. (20) drew attention to the issue by publishing an article that discussed the dilemmas and attributed the existing contradictions to methodological issues such as selection bias, confounding, sample size, reliability, and generalizability.

Since then, five retrospective studies, one randomized controlled trial, and three meta-analyses have all reported that CEB has no effect on complications of hypospadias surgery (9,21-28). In a recent meta-analysis published in 2024, which included 3,201 patients and 10 studies involving 33 surgeons, the importance of study design, sample size, and power analysis was particularly emphasized, in addition to reporting the study's results, thereby addressing previous conflicting publications (28).

Table 1. Ages and complications based on surgical techniques with caudal epidural block or penile block

	Complication	Total (n = 174)	Caudal block (n = 144)	Penile block (n = 30)	p-value
Age (years)		4.70 ± 2.61	4.40 ± 2.46	6.17 ± 2.87	0.002[a]
Meatotomi	+	0	0	0	-
(n = 17)	-	17	13	4	
MAGPI	+	5	5	0	0.571[b]
(n = 54)	-	49	38	11	
TIPU	+	24	19	5	1.000[b]
(n = 53)	-	29	24	5	
Mathiue	+	10	9	1	1.000[b]
(n = 36)	-	26	23	3	
Onlay island flap	+	2	2	0	-
(n = 7)	-	5	5	0	
Dorsal inlay greft	+	6	5	1	1.000[b]
(n = 7)	-	1	1	0	

A Mann-Whitney U test was used for statistical analysis.
Fisher's exact test was used for statistical analysis.
MAGPI: Meatal advancement and glanuloplasty incorporated, TIPU: Tubularized incised plate urethroplasty.

Several authors have proposed that CEB may increase postoperative complications through vasodilation induced by decreased sympathetic activity, leading to venous pooling in penile tissues. This hemodynamic alteration has been suggested to promote intraoperative penile edema, increased bleeding within the surgical field, and subsequent tension on suture lines, mechanisms that are thought to impair tissue healing and to predispose patients to fistula formation or dehiscence, as initially described by Kundra et al. (10) and later supported by Kim et al. (13) and Saavedra-Belaunde et al. (14) and their colleagues. However, numerous subsequent retrospective studies, randomized trials, and meta-analyses have not confirmed a consistent association between CEB and impaired wound healing, calling into question the clinical relevance of this proposed physiological pathway (9,21–28). In line with these recent findings, the absence of a significant difference in complication rates in our study suggests that venous pooling is unlikely to be a predominant mechanism underlying postoperative complications.

Study Limitations

This study has several limitations that should be acknowledged. First, its retrospective design inherently carries a risk of selection bias and limits the ability to control for potential confounding factors. The relatively small sample size, particularly in the PB group, reduces the statistical power and restricts the generalizability of the findings. Finally, the single-center institutional nature of the data may not reflect practices or outcomes in other settings, reducing external validity.

Prospective, multicenter studies with larger sample sizes and detailed stratification based on hypospadias severity, surgical technique, and patient-specific factors are necessary to provide more definitive evidence on the impact of different regional anesthesia techniques on outcomes of hypospadias surgery.

CONCLUSION

Hypospadias surgery is one of the most challenging areas in pediatric urology because of the high prevalence of hypospadias, relatively high complication rates, and more than three hundred surgical techniques described in the literature. The application of CEB, a frequently used technique in this field, has often been investigated for its impact on surgical outcomes, resulting in conflicting findings in the literature. In our study, the lack of a significant difference in surgical complications associated with CEB suggests that its use, considering the associated risks, is not a contraindication to surgery in these patients. Considering its analgesic effectiveness, we believe CEB to be beneficial in pediatric hypospadias surgery.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Gazi University Clinical Research Ethics Committee (decision number: 219, dated: 27.12.2021).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: C.K., G.A., R.K., Z.T., K.S., B.I., Concept: R.K., Z.T., K.S., B.I., Design: R.K., Z.T., K.S., B.I., Data Collection or Processing: C.K., G.A., G.G.K., Analysis or Interpretation: C.K., G.A., G.G.K., Literature Search: C.K., G.A., R.K., Z.T., K.S., B.I., Writing: C.K., G.A., G.G.K., R.K., Z.T., K.S., B.I.

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

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

References

1. Bouty A, Ayers KL, Pask A, Heloury Y, Sinclair AH. The genetic and environmental factors underlying hypospadias. Sex Dev. 2015; 9: 239-59.
2. Blaschko SD, Cunha GR, Baskin LS. Molecular mechanisms of external genitalia development. Differentiation. 2012; 84: 261-8.
3. Baskin L. What Is Hypospadias? Clin Pediatr (Phila). 2017; 56: 409-18.
4. van der Horst HJ, de Wall LL. Hypospadias, all there is to know. Eur J Pediatr. 2017; 176: 435-41.
5. Hüppe T, Pattar G, Maass B. Kaudalanästhesie: Übersicht und praktische Handlungsempfehlungen [Caudal Anesthesia: Overview and Practical Recommendations]. Anasthesiol Intensivmed Notfallmed Schmerzther. 2022; 57: 724-36.
6. Xie L, Tao H, Bao F, Zhu Y, Fang F, Bao X, et al. Major complications of caudal block: A prospective survey of 973 cases in adult anorectal surgery. Heliyon. 2023; 9: e20759.
7. Wood D, Wilcox D. Hypospadias: lessons learned. An overview of incidence, epidemiology, surgery, research, complications, and outcomes. Int J Impot Res. 2023; 35: 61-6.
8. Taicher BM, Routh JC, Eck JB, Ross SS, Wiener JS, Ross AK. The association between caudal anesthesia and increased risk of postoperative surgical complications in boys undergoing hypospadias repair. Paediatr Anaesth. 2017; 27: 688-94.
9. Alizadeh F, Amraei M, Haghdani S, Honarmand A. The effect of caudal epidural block on the surgical complications of hypospadias repair in children aged 6 to 35 months: A randomized controlled trial. J Pediatr Urol. 2022; 18: 59.e1-6.
10. Kundra P, Yuvaraj K, Agrawal K, Krishnappa S, Kumar LT. Surgical outcome in children undergoing hypospadias repair under caudal epidural vs penile block. Paediatr Anaesth. 2012; 22: 707-12.
11. Zaidi RH, Casanova NF, Haydar B, Voepel-Lewis T, Wan JH. Urethrocutaneous fistula following hypospadias repair: regional anesthesia and other factors. Paediatr Anaesth. 2015; 25: 1144-50.
12. Kreysing L, Höhne C. A retrospective evaluation of fistula formation in children undergoing hypospadias repair and caudal anesthesia. Paediatr Anaesth. 2016; 26: 329-30.
13. Kim MH, Im YJ, Kil HK, Han SW, Joe YE, Lee JH. Impact of caudal block on postoperative complications in children undergoing tubularised incised plate urethroplasty for hypospadias repair: a retrospective cohort study. Anaesthesia. 2016; 71: 773-8.
14. Saavedra-Belaunde JA, Soto-Aviles O, Jorge J, Escudero K, Vazquez-Cruz M, Perez-Brayfield M. Can regional anesthesia have an effect on surgical outcomes in patients undergoing distal hypospadia surgery? J Pediatr Urol. 2017; 13: 45.e1-4.
15. Braga LH, Jegatheeswaran K, McGrath M, Easterbrook B, Rickard M, DeMaria J, et al. Cause and effect versus confounding-ıs there a true association between caudal blocks and tubularized ıncised plate repair complications? J Urol. 2017; 197: 845-51.
16. Zhu C, Wei R, Tong Y, Liu J, Song Z, Zhang S. Analgesic efficacy and impact of caudal block on surgical complications of hypospadias

repair: a systematic review and meta-analysis. Reg Anesth Pain Med. 2019; 44: 259-67.

17. Goel P, Jain S, Bajpai M, Khanna P, Jain V, Yadav DK. Does caudal analgesia increase the rates of urethrocutaneous fistula formation after hypospadias repair? Systematic review and meta-analysis. Indian J Urol. 2019; 35: 222-9.
18. Tanseco PP, Randhawa H, Chua ME, Blankstein U, Kim JK, McGrath M, et al. Postoperative complications of hypospadias repair in patients receiving caudal block vs. non-caudal anesthesia: A meta-analysis. Can Urol Assoc J. 2019; 13: E249-57.
19. Ngoo A, Borzi P, McBride CA, Patel B. Penile nerve block predicts higher revision surgery rate following distal hypospadias repair when compared with caudal epidural block: A consecutive cohort study. J Pediatr Urol. 2020; 16: 439.e1-6.
20. Braga LH, McGrath M, Farrokhyar F. Dorsal penile block versus caudal epidural anesthesia effect on complications post-hypospadias repair: Dilemmas, damned dilemmas and statistics. J Pediatr Urol. 2020; 16: 708-11.
21. Kandırıcı A, Mutlu M, Yiğit D. Hipospadias ameliyatında kullanılan blok anestezisi komplikasyon oranlarını arttırıyor mu? J Behcet Uz Child Hosp. 2021; 11: 328-32.
22. Zhang J, Zhu S, Zhang L, Fu W, Hu J, Zhang Z, et al. The association between caudal block and urethroplasty complications of distal tubularized incised plate repair: experience from a South China National Children's Medical Center. Transl Androl Urol. 2021; 10: 2084-90.
23. Karagözlü Akgül A, Canmemiş A, Eyvazov A, Hürel H, Kiyan G, Umuroğlu T, et al. Effects of caudal and penile blocks on the complication rates of hypospadias repair. Balkan Med J. 2022; 39: 239-45.
24. Fischer KM, Van Batavia J, Hyacinthe N, Weiss DA, Tan C, Zderic SA, et al. Caudal anesthesia is not associated with post-operative complications following distal hypospadias repair. J Pediatr Urol. 2023; 19: 374-9.
25. Adler AC, Bhatia VP, Chandrakantan A, Nathanson BH, Ouellette L, Austin PF. Association of analgesic block with the ıncidence of complications following hypospadias surgery; a meta-analysis. Urology. 2022; 166: 11-7.
26. Xia Y, Yang Z, Li J, Liu P, Song H, Sun N, et al. Urethrocutaneous fistula and glans dehiscence formation of hypospadias surgery in patients receiving caudal block vs. non-caudal block: A meta-analysis. J Pediatr Urol. 2024; 20: 227-36.
27. Hu JC, Belon C, Ravula NR, Durbin-Johnson B, Kurzrock EA. Impact of caudal block on revision rates after hypospadias repair: Multi-institution review. J Pediatr Urol. 2023; 19: 292.e1-7.
28. Adler AC, Austin PF. Caudal block for hypospadias repair: Unfolding the controversy through statistical analysis and how we can put it all to rest! Paediatr Anaesth. 2024; 34: 108-11.

Gazi Med J 2026;37(2):152-156

DOI: http://dx.doi.org/10.12996/gmj.2026.4633

Management of Pediatric Orbital Cellulitis: Impact of Surgery and Corticosteroid Use on Clinical Outcomes

Pediatrik Orbital Selülit Yönetimi: Cerrahi ve Kortikosteroid Kullanımının Klinik Sonuçlara Etkisi

Berçin Tarlan, Onur Konuk

Department of Ophthalmology, Gazi University, Faculty of Medicine, Ankara, Türkiye

ABSTRACT

Objective: To evaluate the clinical characteristics, treatment approaches including surgical intervention and systemic corticosteroid use, and clinical outcomes in pediatric patients with orbital cellulitis treated at a tertiary referral center.

Methods: Medical records of pediatric patients (aged 0–18 years) diagnosed with orbital cellulitis and treated between January 2015 and January 2025 were retrospectively reviewed. Demographic data, clinical findings, laboratory results, imaging features, microbiological culture results, treatment modalities (medical, surgical, and corticosteroid therapy), and clinical outcomes were analyzed. Statistical analyses were performed using SPSS version 26.0. Continuous variables were expressed as mean ± standard deviation or median [interquartile range (IQR)], and categorical variables as frequencies and percentages. A p-value ≤0.05 was considered statistically significant.

Results: Fifty-eight patients (58 eyes) were included: 26 (44.8%) were female and 32 (55.2%) were male. The mean age was 7.1 ± 4.52 years, with a median age of 6.8 years (IQR: 2.6–12.2). Thirty patients (51.7%) were managed with medical therapy alone, while 28 (48.3%) underwent surgical intervention. The mean age was significantly higher in the surgical group compared to the medical group (9.3 ± 6 vs. 6.5 ± 4.4 years, p = 0.034).

Subperiosteal abscesses were detected in 30 patients (51.7%), and 28 of these (93.3%) underwent surgical drainage. Abscesses were most commonly located at the medial orbital wall (73.3%). The overall surgical intervention rate was 48.3%. Mean hospital stay was significantly longer in the surgical group than in the medical group (8.8 ± 7.1 vs. 6.3 ± 3.2 days, p = 0.028). Systemic corticosteroids were administered to 24 patients (41.4%). Among surgically treated patients,

ÖZ

Amaç: Bu çalışmada üçüncü basamak sağlık merkezinde orbital selülit tanısıyla tedavi edilen pediatrik hastalarda klinik özelliklerin, cerrahi müdahale ve sistemik kortikosteroid kullanımını içeren tedavi yaklaşımlarının ve klinik sonuçların değerlendirilmesi amaçlandı.

Yöntemler: Ocak 2015 ile Ocak 2025 tarihleri arasında orbital selülit tanısı alarak tedavi edilen pediatrik hastaların (0–18 yaş) tıbbi kayıtları retrospektif olarak incelendi. Demografik veriler, klinik bulgular, laboratuvar sonuçları, görüntüleme özellikleri, mikrobiyolojik kültür sonuçları, tedavi yöntemleri (medikal, cerrahi ve kortikosteroid tedavisi) ve klinik sonuçlar analiz edildi. İstatistiksel analizler SPSS sürüm 26.0 kullanılarak gerçekleştirildi. Sürekli değişkenler ortalama ± standart sapma veya medyan [çeyrekler arası aralık (IQR)] olarak, kategorik değişkenler ise sayı ve yüzde olarak ifade edildi. İstatistiksel anlamlılık düzeyi p ≤ 0,05 olarak kabul edildi.

Bulgular: Çalışmaya 58 hasta (58 göz) dahil edildi; bunların 26'sı (%44,8) kadın, 32'si (%55,2) erkekti. Ortalama yaş 7,1 ± 4,52 yıl, medyan yaş ise 6,8 yıl (IQR: 2,6–12,2) idi. Otuz hasta (%51,7) yalnızca medikal tedavi ile izlenirken, 28 hastaya (%48,3) cerrahi müdahale uygulandı. Cerrahi grupta ortalama yaş, medikal gruba kıyasla anlamlı olarak daha yüksekti (9,3 ± 6 vs. 6,5 ± 4,4 yıl, p = 0,034).

Subperiostal apse 30 hastada (%51,7) saptandı ve bunların 28'ine (%93,3) cerrahi drenaj uygulandı. Apseler en sık medial orbital duvarda (%73,3) lokalizeydi. Genel cerrahi müdahale oranı %48,3 idi. Ortalama hastanede kalış süresi cerrahi grupta medikal gruba göre anlamlı olarak daha uzundu (8,8 ± 7,1 vs. 6,3 ± 3,2 gün, p = 0,028). Sistemik kortikosteroid 24 hastaya (%41,4) uygulandı. Cerrahi uygulanan hastaların 20'si (%71,4) kortikosteroid tedavisi aldı. Kortikosteroid verilen ve verilmeyen cerrahi hastalar arasında hastanede kalış süresi

Cite this article as: Tarlan B, Konuk O. Management of pediatric orbital cellulitis: impact of surgery and corticosteroid use on clinical outcomes. Gazi Med J. 2026;37(2):152-156

Address for Correspondence/Yazışma Adresi: Berçin Tarlan, Department of Ophthalmology, Gazi University, Faculty of Medicine, Ankara, Türkiye
E-mail / E-posta: bercintarlan@gazi.edu.tr
ORCID ID: orcid.org/0000-0001-5800-5451

Received/Geliş Tarihi: 27.01.2026
Accepted/Kabul Tarihi: 09.02.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

20 (71.4%) received corticosteroids. No significant difference in hospital stay was found between corticosteroid-treated and untreated surgical patients (8.2 ± 4.1 vs. 8.0 ± 3.9 days, p > 0.05). Leukocyte levels significantly decreased after treatment in both groups; however, no significant difference was found between the medical and surgical groups in pre- or post-treatment leukocyte counts (p > 0.05). Culture positivity was detected in 11 of 28 surgically treated patients (39.3%). The most common microorganisms were *Streptococcus pyogenes* and *Streptococcus pneumoniae*. Visual acuity significantly improved postoperatively in surgically treated patients. Complications occurred in 3 patients (5.2%): one had osteomyelitis and two had subdural empyema.

Conclusion: In pediatric orbital cellulitis, subperiosteal abscess and older age were associated with an increased need for surgical intervention and a longer hospitalization. Leukocyte count alone was not a reliable predictor of treatment approach. Systemic corticosteroid use did not significantly reduce the length of hospital stay among surgically treated patients. Multidisciplinary management in tertiary centers remains essential for optimal outcomes.

Keywords: Orbital cellulitis, pediatric, subperiosteal abscess, surgical treatment, corticosteroids, sinusitis

ÖZ

açısından anlamlı fark saptanmadı (8,2 ± 4,1 vs. 8,0 ± 3,9 gün, p > 0,05). Lökosit düzeyleri her iki grupta da tedavi sonrası anlamlı olarak azaldı; ancak medikal ve cerrahi gruplar arasında tedavi öncesi ve sonrası lökosit sayıları açısından anlamlı fark bulunmadı (p > 0,05). Kültür pozitifliği cerrahi tedavi uygulanan 28 hastanın 11'inde (%39,3) saptandı. En sık izole edilen mikroorganizmalar *Streptococcus pyogenes* ve *Streptococcus pneumoniae* idi. Cerrahi uygulanan hastalarda ameliyat sonrası görme keskinliğinde anlamlı iyileşme gözlendi. Üç hastada (%5,2) komplikasyon gelişti: birinde osteomiyelit, ikisinde subdural ampiyem saptandı.

Sonuç: Pediatrik orbital selülitte subperiostal apse varlığı ve ileri yaş, cerrahi müdahale gereksinimi ve daha uzun hastanede yatış süresi ile ilişkili bulundu. Lökosit sayısı tek başına tedavi yaklaşımını öngörmede güvenilir bir belirteç değildi. Sistemik kortikosteroid kullanımı, cerrahi uygulanan hastalarda hastanede kalış süresini anlamlı olarak azaltmadı. Optimal sonuçlar için üçüncü basamak merkezlerde multidisipliner yaklaşım önemini korumaktadır.

Anahtar Sözcükler: Orbital selülit, pediatrik, subperiosteal abse, cerrahi tedavi, kortikosteroid, sinüzit

INTRODUCTION

Preseptal and orbital cellulitis are common causes of hospital admission in the pediatric population, and orbital cellulitis requires careful evaluation because of its potential for life-threatening complications and vision loss. The orbital septum represents a critical anatomical boundary for distinguishing between these two clinical entities. Preseptal cellulitis refers to an infectious inflammation involving the soft tissues anterior to the orbital septum, whereas orbital cellulitis involves infectious inflammation of all tissues posterior to the septum (1-4). Early recognition of orbital cellulitis is crucial because of possible complications such as subperiosteal abscess, cavernous sinus thrombosis, subdural empyema, and leptomeningeal involvement (4,5).

Paranasal sinusitis is the most common etiological factor for orbital cellulitis; however, skin-derived infections, odontogenic infections, trauma, foreign bodies, exogenous factors, endogenous causes such as bacteremia, and ocular infections including endophthalmitis and dacryocystitis may also lead to orbital cellulitis (5,6).

The causative microorganism in orbital cellulitis is most reliably identified through microbiological culture of samples obtained during surgery. The most frequently isolated organisms are *Staphylococcus aureus* (*S. aureus*), *Streptococcus species*, and anaerobic bacteria (7,8). Due to geographical variability in the prevalence of methicillin-resistant *S. aureus* (MRSA), routine empirical coverage may not always be necessary; however, MRSA should be considered in patients with recent skin trauma or in cases resistant to treatment (4).

Management requires hospitalization, close monitoring, and a multidisciplinary approach. Treatment consists of broad-spectrum empirical parenteral antibiotics effective against *S. aureus* (including MRSA), streptococci, and gram-negative bacilli, and surgical intervention in cases with orbital abscess formation (2-4,7).

Routine use of glucocorticoids as adjunctive therapy in orbital cellulitis is not generally recommended. Although some oculoplastic surgeons advocate the adjunctive use of glucocorticoids, their clinical efficacy remains uncertain, and concerns exist that glucocorticoids may mask inflammatory signs and obscure disease progression.

MATERIALS AND METHODS

This retrospective cohort study reviewed the medical records of pediatric patients aged 0–18 years diagnosed with orbital cellulitis at our center from January 1, 2015, to January 1, 2025. In all patients, the clinical diagnosis and the presence of a subperiosteal abscess were confirmed by contrast-enhanced orbital computed tomography. Empirical antibiotic therapy targeting common pathogens, including MRSA, was initiated in accordance with national and international recommendations. A multidisciplinary team decided on surgical intervention in cases of subperiosteal abscess, inadequate response to medical therapy within 48–72 hours, or deterioration in visual function. Demographic characteristics, clinical findings, laboratory and imaging results, medical and surgical treatments, and treatment responses were evaluated. Patients with incomplete records were excluded.

Statistical analyses were performed using SPSS version 26.0 (Statistical Package for the Social Sciences). Continuous variables were expressed as mean ± standard deviation or median [interquartile range (IQR)], and categorical variables as frequencies and percentages. The normality of continuous variables was assessed using the Kolmogorov–Smirnov test. Either Student's t-test or the Mann–Whitney U test was applied, as appropriate. Categorical variables were compared using the chi-square test or Fisher's exact test. A p-value ≤0.05 was considered statistically significant. Ethical approval was obtained from the Gazi University Faculty of Medicine Clinical Research Ethics Committee (approval number: E-77082166-604.01-1284873, date: 10.07.2025). The study was conducted in

accordance with the Declaration of Helsinki. Due to the retrospective design, informed consent was waived by the ethics committee.

RESULTS

Demographic and clinical characteristics of the patients are presented in Table 1. Fifty-eight eyes from who were 58 patients diagnosed with orbital cellulitis were included. Twenty-six patients (44.8%) were female and 32 (55.2%) were male. The mean age was 7.1 ± 4.52 years, with a median age of 6.8 years (IQR: 2.6–12.2). Thirty patients (51.7%) were managed with medical therapy alone, while 28 patients (48.3%) underwent surgical intervention. The mean age in the medical group was 6.5 ± 4.4 years, compared to 9.3 ± 6 years in the surgical group. The mean age of surgically treated patients was significantly higher than that of the medically treated patients ($p < 0.05$).

A comparison of demographic and clinical characteristics between the medical and surgical groups is shown in Table 2.

Evaluation of etiological factors revealed sinusitis in 38 patients (65.5%), odontogenic infection in 8 patients (13.8%), trauma in 6 patients (10.3%), skin barrier disruption due to insect bite in 2 patients (3.4%), conjunctivitis in 2 patients (3.4%), and skin infection in 2 patients (3.4%). All patients who developed a subperiosteal abscess had concomitant sinusitis.

Mean hospital stay for the entire cohort was 7.5 ± 5.4 days. The surgical group had a significantly longer hospital stay than the medical group (8.8 ± 7.1 vs. 6.3 ± 3.2 days, $p = 0.028$).

Table 1. Demographic and clinical characteristics of patients.

Variable	Entire cohort (n = 58)
Age (mean ± SD), years	7.1 ± 4.52
Median age (IQR), years	6.8 (2.6–12.2)
Sex	
Female	26 (44.8%)
Male	32 (55.2%)
Presumed etiologic factors	
Sinusitis	38 (65.5%)
Odontogenic infection	8 (13.8%)
Trauma	6 (10.3%)
Skin barrier disruption due to insect bite	2 (3.4%)
Conjunctivitis	2 (3.4%)
-Skin infection	2 (3.4%)
Presence of subperiosteal abscess	30 (51.7%)

SD: Standard deviation, IQR: Interquartile range.

A subperiosteal abscess was detected in 30 patients (51.7%). The localization distribution is presented in Table 3. Surgical intervention was performed in 28 of these patients (93.3%). The overall surgical rate in the cohort was 48.3%. The mean age of patients with subperiosteal abscess was 9.0 ± 4.5 years. Abscess localization was the medial wall in 22 patients (73.3%), medial + inferior wall in 4 (13.3%), the superior wall in 2 (6.7%), and the inferior wall in 2 (6.7%).

The mean pre-treatment leukocyte count was 14,350/mm³ in the medical group and 14,800/mm³ in the surgical group. Post-treatment leukocyte counts were 6,440/mm³ and 8,530/mm³, respectively. Although leukocyte levels decreased markedly in both groups, no significant differences were found between groups in baseline or final leukocyte counts ($p > 0.05$), indicating that peripheral leukocyte count was not a discriminative predictor of treatment modality.

Culture positivity was detected in 11 of 28 surgically treated patients (39.3%). Overall culture positivity was 19.0%. The isolated microorganisms were: *Streptococcus pyogenes* (27.3%), *Streptococcus pneumoniae* (27.3%), *S. aureus* (18.2%), Neisseria meningitidis (9.1%), Haemophilus influenzae (9.1%), and *Streptococcus anginosus* (9.1%).

Antibiotic regimens included ampicillin/sulbactam plus clindamycin in 28 patients (48.3%), vancomycin in 8 patients (13.8%), teicoplanin in 6 patients (10.3%), piperacillin/tazobactam in 3 patients (5.2%), ampicillin/sulbactam alone in 6 patients (10.3%), clindamycin in 3 patients (5.2%), ceftriaxone in 3 patients (5.2%), and ceftazidime in 1 patient (1.7%).

Systemic corticosteroids (1–2 mg/kg/day) were initiated in 24 patients (41.4%) at a mean of 4.01 ± 1.2 days after starting antibiotics. No significant difference in hospital stay was observed between patients receiving corticosteroids and those who did not (8.2 ± 4.1 vs. 8.0 ± 3.9 days, $p > 0.05$). Among surgically treated patients, 20 (71.4%) received adjunctive corticosteroids and 8 (28.6%) did not. In

Table 3. Characteristics of subperiosteal abscesses.

Characteristic	n (%)
Presence of subperiosteal abscess	30/58 (51.7%)
Surgically treated	28/30 (93.3%)
Abscess location	
Medial wall	22 (73.3%)
Medial + inferior wall	4 (13.3%)
Medial + superior wall	2 (6.7%)
Inferior wall	2 (6.7%)
Mean age (years)	9.0 ± 4.5

Table 2. Comparison of medical and surgical treatment groups.

Variable	Medical group (n = 30)	Surgical group (n = 28)	p-value
Age (mean ± SD), years	6.5 ± 4.4	9.3 ± 6.0	<0.05
Length of hospital stay (days)	6.3 ± 3.2	8.8 ± 7.1	<0.05
Pre-treatment leukocyte count (/mm³)	14,350	14,800	>0.05
Post-treatment leukocyte count (/mm³)	6,440	8,530	>0.05

SD: Standard deviation.

this subgroup, corticosteroid use did not significantly affect length of hospital stay ($p > 0.05$).

In 18 surgically treated patients, visual acuity was assessed preoperatively and on postoperative day 1 and postoperative week 1. Mean visual acuity (logMAR) improved from 0.42 ± 0.30 preoperatively to 0.30 ± 0.20 on day 1 and 0.18 ± 0.15 at week 1. Visual acuity at week 1 showed a significant improvement compared with preoperative values ($p = 0.01$).

Complications occurred in 3 patients (5.2%): osteomyelitis in 1 (1.7%) and subdural empyema in 2 (3.4%).

DISCUSSION

Orbital cellulitis is a common pediatric condition that may lead to life-threatening complications. This single-center study evaluated the epidemiological, clinical, and laboratory characteristics and the medical and surgical treatment strategies in patients with orbital cellulitis. The mean patient age (7.1 ± 4.52 years) was consistent with previously reported series (4,8,9).

Paranasal sinus infection—particularly ethmoid sinusitis—is the most frequent etiological factor in orbital cellulitis (4,6,10-13). In our study, sinusitis was present in 65.5% of patients, consistent with rates reported in the literature (60–90%) (8,14). Odontogenic, traumatic, and skin-derived infections were less common, showing distributions similar to published data (6,13). The presence of sinusitis in all patients with subperiosteal abscess supports the pivotal role of paranasal sinus infection in orbital complications (7,15,16). The frequent coexistence of maxillary and ethmoid sinusitis (63.2%) aligns with prior studies reporting ethmoid sinus as the most commonly involved sinus due to its anatomical proximity to the orbit (8,14).

The mean age was significantly higher in the surgical group. Previous studies indicate that younger children respond better to medical therapy, whereas older children more frequently require surgery (11,12). Age-related anatomical changes in sinus ostia and increasing anaerobic infection risk may contribute to more complex infections and reduced response to medical therapy (11,12). Consistent with prior reports, our findings confirm increasing surgical requirement with advancing age (11,17).

The hospital stay was significantly longer for surgically treated patients, consistent with prior series (12). However, unlike some studies, we found no significant difference in leukocyte counts between treatment groups, suggesting that leukocyte level alone may not reliably predict disease severity or surgical need (13-16).

Culture positivity was higher in surgically drained cases, supporting the diagnostic value of surgical sampling. Isolated microorganisms predominantly included *Streptococcus* and *Staphylococcus* species, consistent with existing literature (7,8,14). The relatively low overall culture positivity rate may be attributable to antibiotic treatment prior to referral.

The surgical intervention rate (48.3%) was near the upper range reported in the literature, which likely reflects referral of more severe cases to our tertiary center and the availability of multidisciplinary orbital and ear, nose and throat surgical management.

Subperiosteal abscesses were most commonly located along the medial wall, consistent with previous reports. Superior wall abscesses, considered high-risk for intracranial spread, were surgically treated in all affected patients (15).

Adjunctive systemic corticosteroids did not reduce the length of hospital stay in either the overall cohort or the surgically treated subgroup. These findings align with pediatric series evaluating steroid use and outcomes; evidence remains inconsistent across studies (18-22).

Study Limitations

The retrospective design resulted in non-standardized treatment protocols and incomplete data in some records. The single-center tertiary-care setting may have introduced selection bias toward more severe cases, limiting generalizability. Additionally, variability in the timing and indications for corticosteroid initiation hindered the standardized evaluation of steroid efficacy.

CONCLUSION

Subperiosteal abscess and older age predict increased surgical need and longer hospitalization in pediatric orbital cellulitis. Leukocyte count alone is insufficient to guide treatment decisions. Systemic corticosteroids do not significantly shorten hospital stay in surgically treated patients. Multidisciplinary care remains essential.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Gazi University Faculty of Medicine Clinical Research Ethics Committee (approval number: E-77082166-604.01-1284873, date: 10.07.2025). The study was conducted in accordance with the Declaration of Helsinki.

Informed Consent: Due to the retrospective design, informed consent was waived by the ethics committee.

Footnotes

Authorship Contributions

Surgical and Medical Practices: B.T., O.K., Concept: B.T., O.K., Design: B.T., O.K., Data Collection or Processing: B.T., O.K., Analysis or Interpretation: B.T., O.K., Literature Search: B.T., O.K., Writing: B.T., O.K.

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

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

REFERENCES

1. Saltagi MZ, Rabbani CC, Patel KS, Wannemuehler TJ, Chundury RV, Illing EA, et al. Orbital complications of acute sinusitis in pediatric patients: Management of chandler iii patients. Allergy Rhinol (Providence). 2022; 13: 21526575221097311.
2. Wong SJ, Levi J. Management of pediatric orbital cellulitis: A systematic review. Int J Pediatr Otorhinolaryngol. 2018; 110: 123-9.
3. Anosike BI, Ganapathy V, Nakamura MM. Epidemiology and management of orbital cellulitis in children. J Pediatric Infect Dis Soc. 2022; 11: 214-20.

4. Sharma M, Taylor M, Salehian S, Espinel A, Lloyd KM, Ansusinha E, et al. Clinical epidemiology and microbiology of orbital cellulitis in children. J Pediatric Infect Dis Soc. 2026; 15: piaf113.
5. Burek AG, Tregoning G, Pan A, Liegl M, Harris GJ, Havens PL. Pediatric orbital cellulitis/abscess: Microbiology and pattern of antibiotic prescribing. WMJ. 2023; 122: 52-5.
6. Fanella S, Singer A, Embree J. Presentation and management of pediatric orbital cellulitis. Can J Infect Dis Med Microbiol. 2011; 22: 97-100.
7. Harris GJ. Subperiosteal abscess of the orbit. Age as a factor in the bacteriology and response to treatment. Ophthalmology. 1994; 101: 585-95.
8. Georgakopoulos CD, Eliopoulou MI, Stasinos S, Exarchou A, Pharmakakis N, Varvarigou A. Periorbital and orbital cellulitis: A 10-year review of hospitalized children. Eur J Ophthalmol. 2010; 20: 1066-72.
9. Iftikhar M, Junaid N, Lemus M, Mallick ZN, Mina SA, Hannan U, et al. Epidemiology of primary ophthalmic inpatient admissions in the United States. Am J Ophthalmol. 2018; 185: 101-9.
10. Botting AM, McIntosh D, Mahadevan M. Paediatric pre- and post-septal peri-orbital infections are different diseases. A retrospective review of 262 cases. Int J Pediatr Otorhinolaryngol. 2008; 72: 377-83.
11. Garcia GH, Harris GJ. Criteria for nonsurgical management of subperiosteal abscess of the orbit: Analysis of outcomes 1988-1998. Ophthalmology. 2000; 107: 1454-6; discussion 1457-8.
12. Nghiem AZ, Sanz-Magallon Duque de Estrada B, Farwana R, Osborne SF. Pediatric preseptal and orbital cellulitis - a 6 year experience from a London tertiary centre. Orbit. 2024; 43: 301-6.
13. Aygün D, Doğan C, Hepokur M, Arslan OŞ, Çokuğraş H, Camcıoglu Y. Evaluation of patients with orbital infections. Turk Pediatri Ars. 2017; 52: 221-5.
14. Bülbül L, Özkul Sağlam N, Kara Elitok G, Mine Yazici Z, Hatipoglu N, Hatipoğlu S, et al. Preseptal and orbital cellulitis: Analysis of clinical, laboratory and imaging findings of 123 pediatric cases from Turkey. Pediatr Infect Dis J. 2022; 41: 97-101.
15. Wan Y, Shi G, Wang H. Treatment of orbital complications following acute rhinosinusitis in children. Balkan Med J. 2016; 33: 401-6.
16. Aryasit O, Aunruan S, Sanghan N. Predictors of surgical intervention and visual outcome in bacterial orbital cellulitis. Medicine (Baltimore). 2021; 100: e26166.
17. Villwock MR, Villwock JA. Incidence and extent of sinus procedures in treatment of pediatric orbital cellulitis. Int J Pediatr Otorhinolaryngol. 2020; 135: 110086.
18. Chen L, Silverman N, Wu A, Shinder R. Intravenous steroids with antibiotics on admission for children with orbital cellulitis. Ophthalmic Plast Reconstr Surg. 2018; 34: 205-8.
19. Davies BW, Smith JM, Hink EM, Durairaj VD. C-reactive protein as a marker for initiating steroid treatment in children with orbital cellulitis. Ophthalmic Plast Reconstr Surg. 2015; 31: 364-8.
20. Yen MT, Yen KG. Effect of corticosteroids in the acute management of pediatric orbital cellulitis with subperiosteal abscess. Ophthalmic Plast Reconstr Surg. 2005; 21: 363-6; discussion 366-7.
21. Gill PJ, Mahant S, Hall M, Parkin PC, Shah SS, Wolter NE, et al. Association between corticosteroids and outcomes in children hospitalized with orbital cellulitis. Hosp Pediatr. 2022; 12: 70-89.
22. Gonsalves CL, Borkhoff C, Mahant S, Drouin O, Pound C, Quet J, et al. Association of systemic corticosteroids and clinical outcomes in children hospitalised with severe orbital infections. BMJ Paediatr Open. 2025; 9: e004161.

Gazi Med J 2026;37(2):157-165

DOI: http://dx.doi.org/10.12996/gmj.2025.4598

Polyethylenimine-Mediated Delivery of miR-379-5p Suppresses MTDH and FOXP2 in Colorectal Cancer Cells

Polietilenimin Aracılı miR-379-5p Taşınmasının, Kolorektal Kanser Hücrelerinde MTDH ve FOXP2'yi Baskılaması

Ekin Çelik[1], Ertan Kanbur[2], Cem Bayram[3], Çağrı Aydoğan[4], Miraç Berkay Köksal[4], Selahattin Ata Aydın[4], Yiğit Emre Kurt[4]

[1]Department of Medical Biology, Kırşehir Ahi Evran University, Faculty of Medicine, Kırşehir, Türkiye
[2]Department of Immunology, Kırşehir Ahi Evran University, Faculty of Medicine, Kırşehir, Türkiye
[3]Department of Nanotechnology and Nanomedicine, Hacettepe University, Graduate School of Science and Engineering, Ankara, Türkiye
[4]Kırşehir Ahi Evran University, Faculty of Medicine, Kırşehir, Türkiye

ABSTRACT

Objective: To develop an optimized polyethylenimine (PEI)-based nanocarrier for the intracellular delivery of miR-379-5p and to evaluate its efficacy in suppressing the oncogenic targets metadherin (MTDH) and Forkhead box P2 (FOXP2) in KRAS-wild-type colorectal cancer cells.

Methods: PEI-miRNA nanocomplexes were synthesized at various nitrogen-to-phosphate (N:P) ratios and characterized via dynamic light scattering, zeta potential measurements, and scanning electron microscopy (SEM). Cytotoxicity was assessed in Caco-2 cells using MTT assays to determine the optimal therapeutic concentration. Gene silencing efficiency and intracellular uptake were quantified using reverse transcription quantitative polymerase chain reaction.

Results: The formulation prepared at an N:P ratio of 20:1 exhibited optimal physicochemical properties, featuring a mean hydrodynamic diameter of ~254 nm, a compact spherical morphology, and a highly positive zeta potential (+56.9 mV). At the optimized concentration of 50 nM, the nanocomplexes maintained favorable cell viability while facilitating significant intracellular accumulation of miR-379-5p. Consequently, this delivery strategy achieved robust downregulation of MTDH and FOXP2 expression compared to naked miRNA treatment.

Conclusion: The optimized PEI-miRNA nanocomplexes effectively overcome delivery barriers, enabling successful gene silencing in Caco-2 cells. By restoring the miR-379-5p regulatory axis and suppressing

ÖZ

Amaç: Hücre içi miR-379-5p taşınması için optimize edilmiş polietilenimin (PEI) temelli bir nanotaşıyıcı geliştirilmesi ve bu sistemin, KRAS-wild tip kolorektal kanser hücrelerinde onkojenik hedefler olan metadherin (MTDH) ve Forkhead box P2 (FOXP2) üzerindeki baskılayıcı etkinliğinin değerlendirilmesi amaçlanmıştır.

Yöntemler: PEI-miRNA nanokompleksleri, farklı azot/fosfat (N:P) oranlarında sentezlenmiş ve dinamik ışık saçılımı, zeta potansiyel ölçümleri ile taramalı elektron mikroskobu (SEM) kullanılarak karakterize edilmiştir. Sitotoksisite, optimal terapötik konsantrasyonu belirlemek amacıyla Caco-2 hücrelerinde MTT testi ile değerlendirilmiştir. Gen susturma etkinliği ve hücre içi alım, gerçek zamanlı ters transkripsiyon kantitatif polimeraz zincir reaksiyonu ile nicel olarak analiz edilmiştir.

Bulgular: N:P oranı 20:1 olan formülasyon, yaklaşık 254 nm ortalama hidrodinamik çap, kompakt küresel morfoloji ve yüksek pozitif zeta potansiyeli (+56,9 mV) ile en uygun fizikokimyasal özellikleri göstermiştir. Optimize edilen 50 nM konsantrasyonda, nanokompleksler uygun hücre canlılığını korurken miR-379-5p'nin hücre içinde anlamlı düzeyde birikimini sağlamıştır. Buna bağlı olarak, bu taşıma stratejisi çıplak miRNA uygulamasına kıyasla MTDH ve FOXP2 ekspresyonunda güçlü bir azalma sağlamıştır.

Sonuç: Optimize edilmiş PEI-miRNA nanokompleksleri, taşınma engellerini etkili biçimde aşarak Caco-2 hücrelerinde başarılı gen

Cite this article as: Çelik E, Kanbur E, Bayram C, Aydoğan Ç, Köksal MB, Aydın SA, et al. Polyethylenimine-mediated delivery of miR-379-5p suppresses MTDH and FOXP2 in colorectal cancer cells. Gazi Med J. 2026;37(2):157-165

Address for Correspondence/Yazışma Adresi: Ekin Çelik, Department of Medical Biology, Kırşehir Ahi Evran University, Faculty of Medicine, Kırşehir, Türkiye
E-mail / E-posta: ekin.celik@ahievran.edu.tr
ORCID ID: orcid.org/0000-0003-1966-3907

Received/Geliş Tarihi: 04.12.2025
Accepted/Kabul Tarihi: 20.12.2025
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

FOXP2, this system constitutes a promising molecular platform for targeted RNAi-based interventions in colorectal malignancy.

Keywords: Colorectal cancer, polyethylenimine, miR-379-5p, gene delivery, nanoparticles, RNA interference

ÖZ

susturulmasını mümkün kılmıştır. miR-379-5p düzenleyici ekseninin yeniden kurulması ve FOXP2'nin baskılanması yoluyla bu sistem, kolorektal malignitelerde hedefe yönelik RNA girişimi (RNAi) temelli uygulamalar için umut verici bir moleküler platform oluşturmaktadır.

Anahtar Sözcükler: Kolorektal kanser, polietilenimin, miR-379-5p, gen taşınımı, nanopartiküller, RNA girişimi

INTRODUCTION

Colorectal cancer (CRC) is one of the most prevalent malignancies and remains a leading cause of cancer-related mortality worldwide, representing a persistent global health challenge due to its high morbidity and mortality rates (1). Conventional treatment approaches such as surgery, chemotherapy, and radiotherapy are frequently applied in clinical practice. However, their therapeutic efficacy is markedly reduced in advanced or metastatic stages of the disease. These limitations are exacerbated by the severe treatment-associated toxicities and the development of resistance. These factors not only diminish treatment outcomes but also significantly increase the risk of recurrence, necessitating novel therapeutic strategies (2).

Recent advances in molecular oncology have highlighted gene-based therapeutic approaches, particularly RNA interference, as promising alternatives for overcoming these challenges. RNAi represents a highly specific gene-silencing mechanism that can suppress oncogenes or other disease-driving genes, thereby inhibiting tumor initiation, progression, and metastasis (3). In the context of CRC, RNAi-based therapies are attracting increasing attention for their ability to modulate pathways involved in cell proliferation, angiogenesis, invasion, and chemoresistance, offering a powerful tool for personalized and precision-targeted interventions (4-6). However, the clinical translation of RNAi therapeutics is severely hindered by physiological delivery barriers. Naked RNA molecules are inherently unstable; they are subject to rapid degradation by serum nucleases and, due to their hydrophilicity and negative charge, cannot efficiently traverse the cellular membrane (7). These limitations necessitate the development of advanced delivery platforms that can enhance the stability of RNA molecules, protect them against enzymatic degradation, and ensure efficient intracellular delivery (8-10).

Among synthetic carriers, cationic polymers, particularly polyethylenimine (PEI), have emerged as a gold standard for nucleic acid delivery in cancer research (11-14). The high density of positively charged amino groups in PEI enables strong electrostatic interactions with the negatively charged phosphate backbone of RNA, facilitating the formation of stable polyelectrolyte complexes (15,16). These complexes not only shield RNA molecules from enzymatic degradation but also potentially protect them from nonspecific binding to serum proteins, thereby prolonging circulation time and preserving biological activity. Furthermore, PEI uniquely promotes intracellular uptake via endocytosis and facilitates the "proton sponge effect," enabling endosomal escape, a critical step for functional gene silencing. Importantly, the structural diversity of PEI, which is available in both linear and branched configurations, affords considerable flexibility; branched PEI typically achieves higher transfection efficiency, whereas linear PEI is associated with reduced cytotoxicity, allowing optimization of the carrier system according to specific therapeutic requirements (17,18). These unique properties position PEI as a valuable candidate for advancing RNAi-based therapeutic strategies in colorectal cancer, where overcoming delivery challenges remains the key barrier to clinical translation (19-21).

Despite these advantages, the clinical utility of PEI is often restricted by biocompatibility issues. High-molecular-weight PEI, while efficient, is associated with pronounced cytotoxicity (22). To address these limitations, research has increasingly focused on the use of low-molecular-weight PEI derivatives or chemically modified forms to reduce toxicity while preserving or enhancing gene transfer efficiency (23,24). Accordingly, there is a pressing need for innovative delivery systems capable of safely and selectively transporting RNA molecules, particularly in malignancies with limited therapeutic options, such as colorectal cancer.

Among the therapeutic candidates for RNAi-based strategies, microRNAs have attracted considerable attention due to their ability to regulate multiple oncogenic pathways simultaneously. Specifically, miR-379-5p has been identified as a potent tumor suppressor with activity reported in several malignancies. Previous studies have demonstrated that miR-379 exerts anti-tumorigenic effects by inhibiting cell proliferation, inducing cell cycle arrest, and reducing metastatic potential (25-28). Its tumor-suppressive functions have been described in breast, ovarian, and prostate cancers; its function and therapeutic potential in colorectal cancer remain underexplored (25,29,30). One such potential target is the Forkhead box P2 (FOXP2). The role of FOXP2 in CRC is complex and context-dependent. While some studies suggest a tumor-suppressive function in KRAS-mutant contexts, recent evidence highlights that high FOXP2 expression is significantly correlated with poor overall survival in KRAS-wild-type CRC patients (31-33). Given that Caco-2 cells represent a KRAS-wild-type model, we hypothesized that FOXP2 may function as a driver of malignancy in this specific genetic background.

Consequently, this study aims to evaluate the therapeutic relevance of delivering miR-379-5p via PEI nanocomplexes for suppressing FOXP2 and metadherin (MTDH). We used these nanocomplexes to enhance intracellular uptake while minimizing cytotoxicity, thereby providing initial evidence for the efficacy of PEI-mediated delivery of this specific miRNA. Collectively, these findings offer a novel framework for the rational design of targeted RNAi-based interventions in colorectal malignancy.

MATERIALS AND METHODS

Synthesis of PEI-miRNA Nanocomplexes

Polyethylenimine (PEI; Sigma, cat. no. 408719) was selected as the carrier polymer and was complexed with microRNA-379-5p (miRNA mimic; Qiagen) at varying concentrations to obtain nanocomplexes. The complexes were subsequently characterized to determine the optimal N:P ratio.

For synthesis, miRNA solutions (0.1 mM) prepared in NaCl were incubated with PEI solutions at N:P ratios of 5:1, 10:1, 15:1, and 20:1. Incubations were performed at room temperature, with gentle mixing on a magnetic stirrer when required. The N:P ratio was calculated according to the following formula (34):

$$m_{(polymer)} = n_{(miRNA)} \times \text{number of phosphate groups} \times MW_{(protonatable\ units)} \times (N:P)$$

where m represents the required polymer mass, n represents the number of moles of miRNA, N denotes the number of nitrogen groups, and P denotes the number of phosphate groups.

Characterization of PEI-miRNA Complexes

Physicochemical Characterizations

DLS and Zeta Potential

The hydrodynamic diameter and surface charge of the nanocomplexes were determined using a Zetasizer Nano ZSP (Malvern Instruments, UK). Complexes were prepared at the optimized N:P ratio and were diluted 1:10 in nuclease-free water immediately prior to measurement. DLS measurements were performed at 25 °C with a backscattering angle of 173°. Each sample was analyzed in triplicate, and mean ± standard deviation (SD) values were recorded. Zeta potential was measured using disposable folded capillary cells, with three independent runs performed per sample.

Morphological Characterization (SEM)

The surface morphology and size distribution of the optimized PEI-miRNA nanocomplexes (N:P 20:1) were examined using a scanning electron microscope (Tescan GAIA3, USA). Prior to imaging, the nanocomplex dispersion was diluted in nuclease-free water to minimize particle aggregation. A droplet of the suspension was deposited onto a clean silicon wafer and allowed to air-dry at room temperature for 24 hours. To ensure electrical conductivity and prevent surface charging under the electron beam, the dried samples were sputter-coated with a thin layer of gold/palladium. Imaging was performed under high-vacuum conditions at an accelerating voltage of 5-10 kV.

Gel Retardation Assay

Complex formation between PEI and miRNA was assessed by agarose gel electrophoresis, as described previously (34). Nanocomplexes prepared at different N:P ratios were loaded onto 2% agarose gels and electrophoresed at constant voltage for 15 min. Gels were visualized using a ChemiDoc-It Imager (UVP, USA). A decrease in miRNA fluorescence intensity in the miRNA-PEI complex was expected, reflecting complexation with PEI and a reduction in free phosphate groups.

Biological Evaluation

The biological performance of the PEI-miRNA nanocomplexes was assessed through cytotoxicity assays, transfection studies, and gene expression analysis in colorectal cancer cells.

Cytotoxicity

Cell viability was determined using the MTT assay (211091, Abcam, UK). Caco-2 cells were seeded in 96-well plates and incubated with PEI-miRNA nanocomplexes at the selected concentrations for 48 h. Following treatment, MTT solution was added to each well, and cells were incubated to allow formazan crystal formation. Crystals were solubilized in an appropriate solvent, and the absorbance was measured spectrophotometrically at 570 nm. Each experimental and control group was analyzed in eight replicates.

Transfection

Caco-2 cells were cultured in DMEM (Life Sciences, USA) supplemented with 10% fetal bovine serum (FBS, Gibco, USA) at 37 °C in a humidified incubator with 5% CO_2. For transfection, 3×10^6 cells were seeded in 6-well plates and allowed to adhere for 12 h. The medium was then replaced with serum-free DMEM containing either naked miRNA or PEI-miRNA complexes, which were synthesized in this study, at a final concentration of 50 nM miRNA. After 4 h of incubation, the medium was replaced with DMEM containing 10% FBS, and the cells were cultured for an additional 48 h before harvest.

Gene Expression

Real-time PCR was performed to confirm miR-379-5p internalization and evaluate its effect on target gene expression. Total RNA, including miRNAs, was isolated from transfected and control cells. Separate cDNA synthesis protocols were employed for the determination of miRNA and mRNA levels; miRNA analyses were performed using stem-loop RT primers (Integrated DNA Technologies, USA), whereas mRNA analyses were conducted using a standard oligo-dT/reverse transcriptase-based kit protocol. Real-time experiments were conducted using the PowerUp SYBR Green Master Mix (Thermo Scientific, USA) on a LightCycler 480 system (Roche, Switzerland). The comparative Ct method was applied to determine relative expression levels of miR-379-5p and its target genes FOXP2 and MTDH (27,35). β-Actin (ACTB) and U6 snRNA were used as internal reference genes. Primers for U6 snRNA were purchased from ABM (Canada), and all primer sequences are listed in Table 1.

Statistical Analysis

All experiments were performed with at least three independent biological replicates. For the MTT cytotoxicity assay, each condition was tested in eight technical replicates (n = 8). For transfection and real-time PCR analyses, three independent experiments were conducted with three technical replicates per sample (n = 3). For physicochemical characterization (DLS, zeta potential), three independent preparations were analyzed (n = 3). Data were expressed as mean ± SD. Statistical analyses were performed using GraphPad Prism 9.0 (GraphPad Software, USA). Comparisons between two groups were assessed using unpaired Student's t-test, while one-way analysis of variance (ANOVA) with Tukey's post-hoc test was applied for multiple group comparisons. A p-value of <0.05 was considered statistically significant.

RESULTS

DLS and Zeta Potential

Dynamic light scattering (DLS) analysis of the PEI formulations revealed a dominant sub-micron population in all samples. Intensity-weighted main peaks ranged from ~215 to 865 nm. While a minor fraction of micron-sized aggregates (~5-5.5 µm) was detected, it represented a negligible portion of the total intensity distribution (<3% total area) and did not dominate the sample properties. Therefore, only the dominant sub-micron peaks are reported in Table 2 because they represent the functionally relevant nanoparticle population that drives cellular uptake.

Zeta potential measurements in water at 25 °C confirmed that all three NP concentrations were strongly cationic. The principal peaks ranged from approximately +10 to +56 mV, representing >90% of the particle population. Overall, the cationic polymer NPs form positively charged sub-micron dispersions; the high positive zeta potentials are consistent with efficient coating by the polymer and provide sufficient electrostatic stabilization to prevent functional aggregation under measurement conditions.

Among the tested formulations, the 20:1 PEI:NP ratio was selected for further studies based on its physicochemical profile. This formulation yielded nanoparticles with a mean hydrodynamic diameter of 254.6 ± 48.9 nm and a markedly high positive zeta potential of +56.9 ± 5.72 mV. Across the series (5:1 to 20:1), increasing the proportion of cationic polymer resulted in a gradual decrease in particle size and a concomitant increase in surface charge, indicating more efficient complexation/coating and the formation of more compact, densely charged particles. The particle size of approximately 200-300 nm remains suitable for biomedical applications, while the high positive zeta potential is expected to provide strong electrostatic stabilization and reduced aggregation. Taken together, these characteristics identify the 20:1 formulation as the most favorable candidate for subsequent biological evaluation; therefore, all following experiments were performed using this formulation.

Morphological Characterization

Scanning electron microscopy (SEM) was employed to validate the hydrodynamic size measurements and visualize the topography of the nanocomplexes. The SEM micrographs revealed that the optimized formulation (N:P 20:1) formed nanoparticles with a distinct, spherical morphology and a relatively smooth surface texture (Figure 1). The particles appeared compact, with most displaying diameters in the range of ~180 to 220 nm. Notably, the size observed via SEM was slightly smaller than the hydrodynamic diameter obtained by DLS (~254 nm). This difference is expected and can be attributed to the dehydration of the nanocomplexes during SEM sample preparation, which results in the collapse of the hydration shell and a slight shrinkage of the polymeric matrix, in contrast to the swollen state measured in the aqueous phase by DLS. Consistent with the DLS polydispersity data, occasional aggregates were also observed in the dry state.

Gel Retardation Assay

The complexation efficiency of PEI with miRNA was assessed via agarose gel electrophoresis. As expected, naked miRNA (Lane A) migrated freely through the gel matrix, appearing as a distinct band, consistent with its low molecular weight and negative charge. In contrast, at the optimized N:P ratio of 20:1 (Lane B), the miRNA was completely retained in the loading well. The complete absence of a migrating band in this lane confirms that the cationic PEI successfully neutralized the anionic phosphate backbone of the miRNA, resulting in the formation of stable, large-molecular-weight nanocomplexes with no electrophoretic mobility (Figure 2).

Cell Viability

Cell viability assays were performed to evaluate the cytotoxicity of the 20:1 PEI-miRNA nanoparticles at increasing miRNA concentrations (10, 50, and 100 nM). Treatment with 10 nM miRNA-PEI complexes did not appreciably affect cell viability compared with the untreated control, which remained close to 100%. Similarly, exposure to 50-nM complexes preserved high viability, with values remaining above ~90%.

Table 1. Primer sequences used in the study.

	F primer (5'-3')	R primer (5'-3')
miRNA-379-5p	GCGCGTGGTAGACTATGGAA	AGTGCAGGGTCCGAGGTATT
ACTB	CACCATTGGCAATGAGCGGTT	AGGTCTTTGCGGATGTCCACGT
MTDH	GGAGTCAAGACACTGGAGATG	GGGTTGATTACGGCTAACATCC
FOXP2	CAACAGCAGCAGCCAGGA	GAGGCCCCAGTCTCCCTA

FOXP2: Forkhead box P2, MTDH: Metadherin, ACTB: β-actin.

Table 2. Physicochemical characterization of the nanoparticle formulations.

Formulation	Hydrodynamic diameter (nm)	Zeta potential (mV)	PDI
5:1	789.7 ± 75.2	10.1 ± 6.99	0.706
10:1	464 ± 45.4	26.03 ± 7.60	0.444
15:1	302.2 ± 80.23	34.5 ± 11.40	0.518
20:1	254.6 ± 48.88	56.9 ± 5.72	0.386

PDI: Polydispersity index.

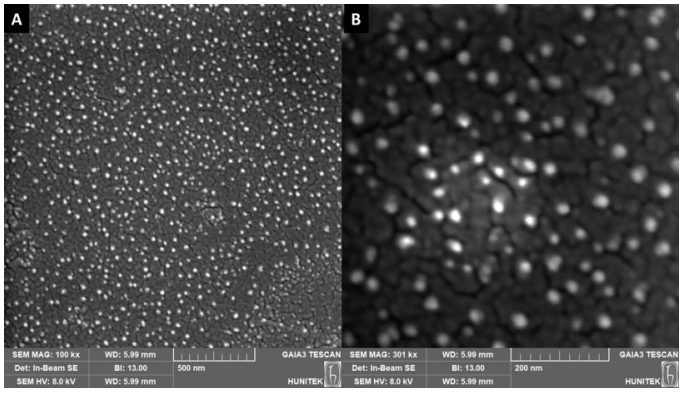

Figure 1. Morphological characterization of PEI-miRNA nanocomplexes via scanning electron microscopy (A) Representative micrograph acquired at 100 kx magnification, illustrating the general distribution and dispersion of the optimized nanocomplexes (N:P 20:1). The scale bar represents 500 nm. (B) High-magnification view (301 kx) detailing the distinct spherical morphology and compact structure of the nanoparticles. The scale bar represents 200 nm. Images are representative of three independent preparations (n = 3).

PEI: Polyethylenimine.

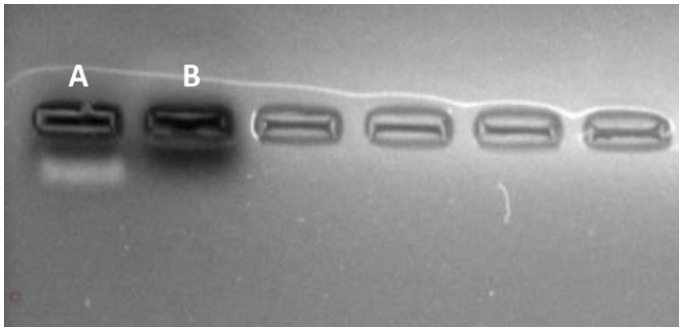

Figure 2. Assessment of miRNA complexation efficiency by agarose gel electrophoresis. The electrophoretic mobility of naked miR-379-5p and PEI-miRNA nanocomplexes was analyzed on a 2% agarose gel. (A) Naked miRNA, exhibits free migration corresponding to its negative charge (B) PEI-miRNA nanocomplexes prepared at a 20:1 ratio show complete retardation, confirming successful complexation. The image is representative of three independent experiments (n = 3).

PEI: Polyethylenimine.

In contrast, treatment with 100 nM miRNA-PEI nanoparticles resulted in a marked reduction in cell viability to approximately 65-70%. These findings indicate that the 20:1 PEI-miRNA formulation is well tolerated at lower concentrations (≤50 nM), whereas higher doses induce a clear, concentration-dependent cytotoxic effect (Figure 3). Based on these findings, 50 nM miRNA in the 20:1 PEI-miRNA formulation was selected for subsequent experiments, as it ensured efficient delivery while maintaining high cell viability.

Gene Expression

To validate the functional intracellular delivery of the cargo, the expression levels of downstream target genes MTDH and FOXP2 were quantified via reverse transcription quantitative polymerase chain reaction (RT-qPCR). As anticipated, treatment with naked miR-379-5p had a negligible impact on gene expression relative to the

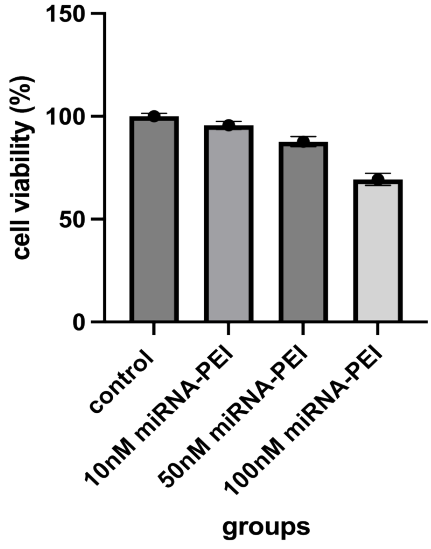

Figure 3. Cytotoxicity profile of PEI-miRNA nanocomplexes in colorectal cancer cells. Cell viability of Caco-2 cells was evaluated using the MTT assay after 48 h of incubation with PEI-miRNA nanocomplexes (N:P ratio 20:1) at miRNA concentrations of 10, 50, and 100 nM. Data are presented as the percentage of viable cells relative to the untreated control group (set to 100%). Results represent the mean ± SD of three independent experiments (n=3). Statistical significance was assessed using one-way ANOVA.

SD: Standard deviation, ANOVA: Analysis of variance, PEI: Polyethylenimine.

control, an effect attributed to the poor cellular uptake of free RNA. In contrast, the optimized PEI-miRNA nanocomplexes facilitated a significant downregulation of both targets. Specifically, MTDH mRNA levels were reduced by approximately 20% compared with the control group ($p < 0.001$); this suppression was significantly more potent than that of the naked miRNA treatment ($p < 0.01$) (Figure 4A). Similarly, FOXP2 expression exhibited a significant decrease (~16%) relative to control ($p < 0.01$), further confirming the superior transfection efficiency of the PEI-complexed formulation compared with naked miRNA ($p < 0.05$) (Figure 4B).

To verify that the observed downregulation of MTDH and FOXP2 was driven by the successful cellular internalization of the therapeutic cargo, the intracellular levels of miR-379-5p were quantified via RT-qPCR. Naked RNA molecules are known to possess poor membrane permeability due to their negative charge and hydrophilicity. Consistent with this, cells treated with naked miR-379-5p exhibited only a marginal, non-significant increase in intracellular miRNA levels compared to the untreated control (Figure 5). In contrast, the delivery of miR-379-5p via the optimized PEI nanocarrier (N:P 20:1) resulted in a statistically significant increase in intracellular miRNA abundance relative to both the control ($p < 0.01$) and the naked miRNA group ($p < 0.05$). These results confirm that the PEI nanocarrier effectively facilitates the transmembrane transport and intracellular accumulation of miR-379-5p in Caco-2 cells, validating the transfection efficiency required for functional gene silencing of the target oncogenes.

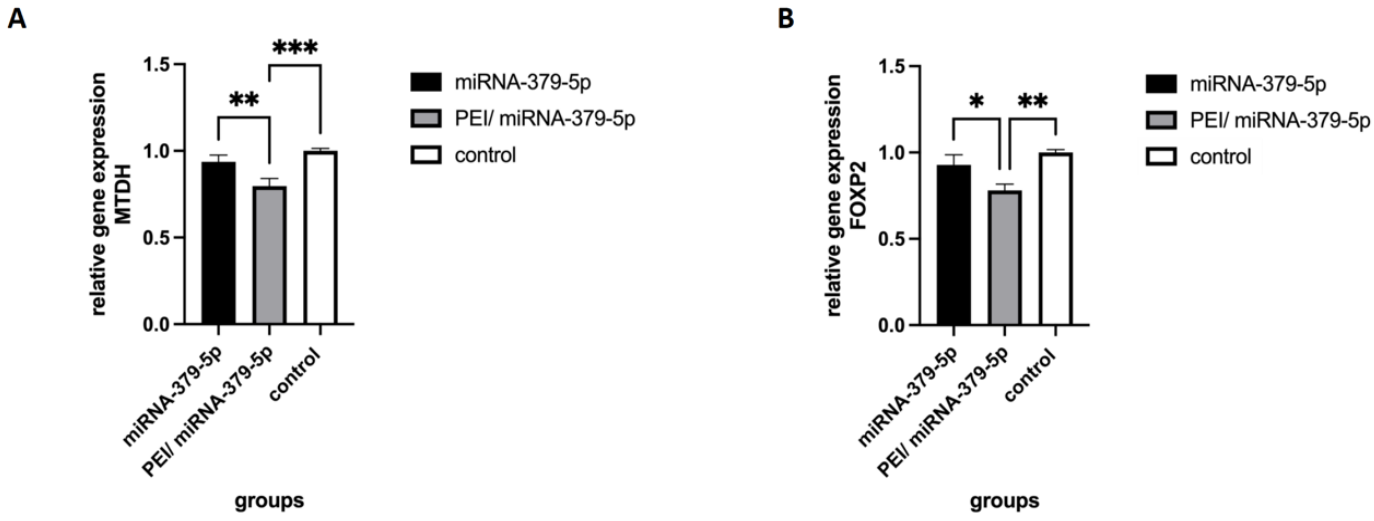

Figure 4. Relative mRNA expression levels of target genes MTDH (A) and FOXP2 (B) in Caco-2 cells. Cells were treated with naked miR-379-5p or PEI/miR-379-5p nanocomplexes (N:P 20:1, 50 nM) for 48 h. Gene expression was quantified using RT-qPCR and normalized to the internal reference gene ACTB (β-actin). Data are presented as fold change relative to the control group (set to 1.0). Values represent the mean ± SD of three independent experiments (n = 3). Statistical significance was determined using one-way ANOVA with Tukey's post-hoc test; significance levels were indicated as *p < 0.05, **p < 0.01, and ***p < 0.001 compared to the untreated control group.

SD: Standard deviation, ACTB: β-actin, MTDH: Metadherin, FOXP2: Forkhead box P2, ANOVA: Analysis of variance, RT-qPCR: Reverse transcription quantitative polymerase chain reaction, PEI: Polyethylenimine..

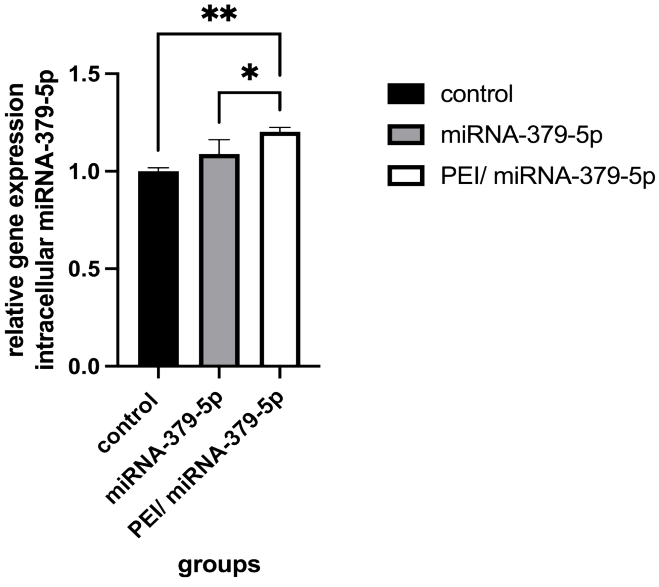

Figure 5. Intracellular accumulation of miR-379-5p in Caco-2 cells. The relative intracellular expression of miR-379-5p was determined by RT-qPCR following 48 h of incubation with either naked miRNA or PEI-miRNA nanocomplexes (N:P 20:1) at a final miRNA concentration of 50 nM. Data are presented as fold changes relative to the control group (set to 1.0) after normalization to the internal reference gene. Values represent the mean ± SD of three independent experiments (n = 3). Statistical significance was analyzed using one-way ANOVA with Tukey's post-hoc test: *p < 0.05 versus naked miRNA; **p < 0.01 versus untreated control.

SD: Standard deviation, ANOVA: Analysis of variance, RT-qPCR: Reverse transcription quantitative polymerase chain reaction, PEI: Polyethylenimine.

DISCUSSION

The development of efficient and safe delivery systems remains the primary bottleneck in the clinical translation of RNAi therapeutics. In this study, we engineered and optimized a PEI-based nanocarrier system for the intracellular delivery of miR-379-5p into CRC cells. Our findings demonstrate that the optimized PEI-miRNA nanocomplexes (N:P ratio 20:1) possess physicochemical characteristics favorable for cellular uptake, exhibit a manageable toxicity profile at therapeutic concentrations, and significantly enhance silencing of the oncogenic targets MTDH and FOXP2 compared with naked miRNA.

The physicochemical properties of nanoparticle systems, specifically hydrodynamic diameter and surface charge, are critical determinants of their biological fate and cellular internalization efficiency (36). Our DLS data revealed that increasing the N:P ratio resulted in more compact nanoparticles, with the 20:1 formulation yielding a mean diameter of approximately 254 nm. While particles in the 200 nm range are classically described as entering cells via clathrin-mediated endocytosis, recent evidence suggests that strict monodispersity may not be the sole prerequisite for effective delivery. González-Domínguez et al. (37) recently demonstrated that micrometric DNA/PEI aggregates (450-650 nm) can actually correlate with higher transient gene expression yields compared to smaller counterparts. This enhanced efficacy is attributed to the uptake of larger complexes via macropinocytosis and their subsequent accumulation near the nuclear envelope, where they function as a sustained "reservoir", progressively releasing the genetic cargo. Consequently, the slight polydispersity observed in our system may functionally contribute to the therapeutic effect by enabling exploitation of multiple entry pathways, including clathrin-mediated uptake of smaller fractions and macropinocytosis of larger aggregates. Furthermore,

morphological analysis via SEM confirmed that the nanocomplexes possess a compact, spherical structure. The particle size observed in SEM micrographs (~180-220 nm) was slightly smaller than the hydrodynamic diameter measured by DLS (~254 nm). This difference is anticipated and characteristic of polymeric systems, as DLS measures the hydrodynamic radius, including the hydration shell, in the aqueous phase, whereas SEM depicts the particles in a dehydrated solid state.

Concomitant with the size reduction, we observed a substantial increase in zeta potential, reaching +56.9 mV at the optimized ratio. This high cationic charge serves a dual purpose: first, it ensures the colloidal stability of the dispersion through electrostatic repulsion, thereby reducing particle aggregation in suspension (38). This high zeta potential was measured in nuclease-free water, reflecting the inherent colloidal stability of the formulation; however, under physiological conditions, the effective surface charge is expected to decrease due to ionic screening and protein corona formation. Secondly, and perhaps more critically, it facilitates the initial electrostatic adsorption of the nanocomplexes onto the negatively charged proteoglycans of the cell membrane, a prerequisite for efficient endocytosis (39). The agarose gel retardation assays corroborated these findings by showing that a 20:1 ratio provided sufficient cationic density to fully neutralize and condense miRNA, thereby protecting it from premature degradation—a crucial feature given the inherent instability of naked RNA in physiological environments.

While the high cationic charge density of PEI is advantageous for cellular uptake and promotes the "proton sponge" effect, which facilitates endosomal escape, it is also the primary source of PEI-mediated cytotoxicity. Our viability assays in Caco-2 cells revealed a concentration-dependent toxicity profile. While the formulation was well tolerated at 10 nM and 50 nM, viability dropped significantly at 100 nM. This observation is consistent with literature attributing PEI toxicity to membrane disruption and mitochondrial damage caused by an excess of free polycations (40). Consequently, the identification of 50 nM as the optimal working concentration represents a critical balance between maximizing therapeutic payload and preserving cellular health.

To explicitly validate the transmembrane transport of the therapeutic cargo, we analyzed intracellular miRNA levels (Figure 4), providing direct evidence of the nanocarrier's transport efficiency. While cells treated with naked miR-379-5p showed negligible intracellular accumulation, likely due to electrostatic repulsion between the negatively charged RNA and the cell membrane, the PEI-complexed group exhibited a statistically significant increase in intracellular miRNA abundance ($p<0.01$). This confirms that the optimized nanocomplexes effectively shielded the miRNA's charge and utilized the cationic PEI surface to facilitate transmembrane crossing, likely via endocytic pathways.

Building on this efficient internalization, we validated the functional capability of the delivered cargo by assessing downstream gene targets. We utilized miR-379-5p, a tumor suppressor often downregulated in various malignancies, to validate our system. The significant downregulation of MTDH and FOXP2 observed in the PEI-complexed group confirms successful gene silencing. Although the formulation was optimized to minimize acute cytotoxicity at this concentration, the robust suppression of these oncogenes suggests significant therapeutic potential through restoration of the miR-379-5p regulatory axis. Although the reduction in mRNA levels appears modest (~16-20%), miRNAs primarily function through translational repression. Therefore, the protein-level suppression might be more profound than what is observed at the transcript level, warranting future proteomic validation. Our observation that FOXP2 silencing promotes cell death stands in contrast to previous reports, such as Liao et al. (33), which characterized FOXP2 as a tumor suppressor that inhibits pyroptosis in CRC. However, this discrepancy can be attributed to the specific genetic context of the cell lines used. A recent study by Liu et al. (31) demonstrated that the function of FOXP2 in colorectal cancer is dictated by KRAS mutation status. While FOXP2 acts as a suppressor in KRAS-mutant cells, its high expression is associated with poor prognosis and adverse clinical outcomes in KRAS-wild-type patients. Since Caco-2 cells are KRAS-wild-type, the high expression of FOXP2 likely supports tumor progression in this model. Therefore, our findings support the therapeutic utility of targeting FOXP2 in KRAS-wild-type CRC subtypes and further emphasize the need for personalized RNAi strategies based on genetic biomarkers.

Study Limitations

Despite the promising results, certain limitations warrant mention. First, the biological evaluation was conducted *in vitro* using a Caco-2 model, which cannot fully replicate systemic physiological barriers or protein-corona effects, necessitating future *in vivo* validation. Second, while transcriptional silencing was confirmed via RT-qPCR, protein-level validation has yet to be performed to fully quantify functional suppression. Given that the therapeutic benefit of targeting FOXP2 is linked to the KRAS-wild-type genotype, the applicability of this strategy to KRAS-mutant subtypes of colorectal cancer requires further comparative investigation.

CONCLUSION

In this study, we engineered a cationic polymer-based nanocarrier system for effective delivery of the tumor suppressor miR-379-5p to colorectal cancer cells. Through systematic physicochemical optimization, the 20:1 PEI-miRNA formulation was identified as the ideal candidate, exhibiting a high cationic charge (+56.9 mV) that promotes stability and cellular uptake, while maintaining a manageable toxicity profile at therapeutic concentrations (50 nM). Although the particles displayed some polydispersity, this heterogeneity likely facilitates multiple cellular entry pathways, including macropinocytosis, thereby enhancing the functional delivery of the genetic cargo. Biological evaluation confirmed that these nanocomplexes significantly downregulated the oncogenic targets MTDH and FOXP2, restoring the tumor-suppressive molecular profile of miR-379-5p, thereby establishing a foundation for inhibiting tumor progression. Collectively, these findings validate the 20:1 PEI-miRNA system as a promising platform for RNAi-based precision oncology and warrant further investigation in colorectal cancer models *in vivo*.

Ethics

Ethics Committee Approval: Ethical approval was not required for this study, as no human participants or animal subjects were involved.

Informed Consent: Informed consent was not required because the study did not require ethics committee approval.

Footnotes

Authorship Contributions

Surgical and Medical Practices: E.Ç., E.K., C.B., Ç.A., M.B.K., S.A.A., Y.E.K., Concept: E.Ç., E.K., C.B., Ç.A., M.B.K., S.A.A., Y.E.K., Design: E.Ç., E.K., C.B., Ç.A., M.B.K., S.A.A., Y.E.K., Data Collection or Processing: E.Ç., E.K., C.B., Ç.A., M.B.K., S.A.A., Y.E.K., Analysis or Interpretation: E.Ç., E.K., C.B., Literature Search: E.Ç., E.K., C.B., Writing: E.Ç., E.K., C.B., Ç.A., M.B.K., S.A.A., Y.E.K.

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

Financial Disclosure: This study was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK-BIDEB) under Project No. 1919B012433754.

Acknowledgment

The authors would like to thank all colleagues who provided valuable input during the development of this study. In particular, we are grateful to Beyza Şahin, Nisa Nur Topaloğlu, Ebrar Dilanur Özben, and Yasemin Nişancı for their assistance and support.

REFERENCES

1. Siegel RL, Jakubowski CD, Fedewa SA, Davis A, Azad NS. Colorectal cancer in the young: epidemiology, prevention, management. Am Soc Clin Oncol Educ Book. 2020; 40: 1-14.
2. Tsai KY, Huang PS, Chu PY, Nguyen TNA, Hung HY, Hsieh CH, et al. Current applications and future directions of circulating tumor cells in colorectal cancer recurrence. Cancers (Basel). 2024; 16: 2316.
3. Goel A, Rastogi A, Jain M, Niveriya K. RNA-based therapeutics: past, present and future prospects, challenges in cancer treatment. Curr Pharm Biotechnol. 2024; 25: 2125-37.
4. Pallathadka H, Hsu CY, Obaid Saleh R, Renuka Jyothi S, Kumar A, Yumashev A, et al. Specific small interfering RNAs (siRNAs) for targeting the metastasis, immune responses, and drug resistance of colorectal cancer cells (CRC). Int Immunopharmacol. 2024; 140: 112730.
5. Xu R, Du A, Deng X, Du W, Zhang K, Li J, et al. tsRNA-GlyGCC promotes colorectal cancer progression and 5-FU resistance by regulating SPIB. J Exp Clin Cancer Res. 2024; 43: 230.
6. Al-Jumaili MMO. A potential interpretation of RNA interference and GC island modification in progressive colorectal cancer suppression. Biomed Biotechnol Res J. 2023; 7: 425-31. Available from: https://journals.lww.com/bbrj/fulltext/2023/07030/a_potential_interpretation_of_rna_interference_and.16.aspx.
7. Castellano M, Blanco V, Calzi ML, Costa B, Witwer K, Hill M, et al. Ribonuclease activity undermines immune sensing of naked extracellular RNA. bioRxiv [Preprint]. 2024: 2024.04.23.590771.
8. Muskan M, Abeysinghe P, Cecchin R, Branscome H, Morris KV, Kashanchi F. Therapeutic potential of RNA-enriched extracellular vesicles: The next generation in RNA delivery via biogenic nanoparticles. Mol Ther. 2024; 32: 2939-49.
9. Zhang H, Vandesompele J, Braeckmans K, De Smedt SC, Remaut K. Nucleic acid degradation as barrier to gene delivery: a guide to understand and overcome nuclease activity. Chem Soc Rev. 2024; 53: 317-60.
10. Wei PS, Thota N, John G, Chang E, Lee S, Wang Y, et al. Enhancing RNA-lipid nanoparticle delivery: Organ- and cell-specificity and barcoding strategies. J Control Release. 2024; 375: 366-88.
11. Domingues C, Jarak I, Coelho J, Carvalho RA, Veiga F, Vitorino C, et al. Optimizing pluronic-PEI nanocarriers for RNAI delivery in oral cancer: from polymer synthesis to functional screening. Biomacromolecules. 2025; 26: 7484-511.
12. Gonzalez-Melero L, Santos-Vizcaino E, Varela-Calvino R, Gomez-Tourino I, Asumendi A, Boyano MD, et al. PLGA-PEI nanoparticle covered with poly(I:C) for personalised cancer immunotherapy. Drug Deliv Transl Res. 2024; 14: 2788-803.
13. Li R, Yu K, Fan G-C, Song Z, Luo X. Polyethyleneimine cross-linked BSA with enhanced antifouling capability for electrochemical detection of cancer antigen 125 in complex biofluids. Sens Actuators B Chem. 2024; 408: 135520.
14. Yang S, Wu Y, Zhong W, Chen R, Wang M, Chen M. GSH/pH dual activatable cross-linked and fluorinated pei for cancer gene therapy through endogenous iron de-hijacking and *in situ* ROS amplification. Adv Mater. 2024; 36: e2304098.
15. Huang P, Deng H, Zhou Y, Chen X. The roles of polymers in mRNA delivery. Matter. 2022; 5: 1670-99.
16. Berger S, Lächelt U, Wagner E. Dynamic carriers for therapeutic RNA delivery. Proc Natl Acad Sci U S A. 2024; 121: e2307799120.
17. Lübbersmeyer TL, Wissel K, Hoffmann A, Kötter L, Lenarz T, Paasche G. Establishing a non-viral transfection system for fibroblasts using branched polyethylenimine. PLoS One. 2025; 20: e0329666.
18. Ismail A, Chou SF. Polyethylenimine Carriers for Drug and Gene Delivery. Polymers (Basel). 2025; 17: 2150.
19. Samadi P, Rahbarizadeh F, Nouri F, Soleimani M, Najafi R, Jalali A. MUC1-driven guanylin gene delivery via succinylated PEI-9 nanocarrier for colorectal cancer treatment: an *in silico* and *in vitro* study. Adv Pharm Bull. 2025; 15: 844-70. Available from: https://apb.tbzmed.ac.ir/Article/apb-45741
20. Lee GY, Lo PY, Cho EC, Zheng JH, Li M, Huang JH, et al. Integration of PEG and PEI with graphene quantum dots to fabricate pH-responsive nanostars for colon cancer suppression *in vitro* and *in vivo*. FlatChem. 2022; 31: 100320.
21. Jiazhen N, Meihui S, De-E L, Na L, Youtao X, Qixian C, et al. Remodeling the inflammatory and immunosuppressive tumor microenvironment for enhancing antiangiogenic gene therapy of colorectal cancer. Adv Healthc Mater. 2025; 14: e2402887.
22. Lin G, Huang J, Zhang M, Chen S, Zhang M. Chitosan-crosslinked low molecular weight pei-conjugated iron oxide nanoparticle for safe and effective DNA delivery to breast cancer cells. Nanomaterials (Basel). 2022; 12: 584.
23. Peng Y, Gao Y, Yang C, Guo R, Shi X, Cao X. Low-molecular-weight poly(ethylenimine) nanogels loaded with ultrasmall iron oxide nanoparticles for T1-weighted MR imaging-guided gene therapy of sarcoma. ACS Appl Mater Interfaces. 2021; 13: 27806-13.
24. Wang C, Pan C, Yong H, Wang F, Bo T, Zhao Y, et al. Emerging non-viral vectors for gene delivery. J Nanobiotechnology. 2023; 21: 272.
25. Cassidy JR, Persson M, Voss G, Underbjerg KR, Ivkovic TC, Bjartell A, Edsjö A, Lilja H, Ceder Y. MicroRNA-379 modulates prostate-specific antigen expression through targeting the androgen receptor in prostate cancer. Cancers (Basel). 2025; 17: 3245.
26. Mosaad H, Abdelrahman AE, Abdullatif A, Lashin ME, Ramadan MSH, El-Azony A, et al. MIR-379-5p Expression in Endometrial

Cancer and Its Correlation with ROR1 Expression. Asian Pac J Cancer Prev. 2023; 24: 239-48.

27. Hu W, Cao W, Liu J. RETRACTED: LncRNA-NEAT1 facilitates autophagy to boost pemetrexed resistance in lung adenocarcinoma via the mir-379-3p/HIF1A pathway. Human & Experimental Toxicology. 2025; 43.
28. Srinath S, Jishnu PV, Varghese VK, Shukla V, Adiga D, Mallya S, et al. Regulation and tumor-suppressive function of the miR-379/miR-656 (C14MC) cluster in cervical cancer. Mol Oncol. 2024; 18: 1608-30.
29. Yang K, Li D, Jia W, Song Y, Sun N, Wang J, et al. MiR-379-5p inhibits the proliferation, migration, and invasion of breast cancer by targeting KIF4A. Thorac Cancer. 2022; 13: 1916-24.
30. Shukla D, Mishra S, Mandal T, Charan M, Verma AK, Khan MMA, et al. MicroRNA-379-5p attenuates cancer stem cells and reduces cisplatin resistance in ovarian cancer by regulating RAD18/Polη axis. Cell Death Dis. 2025; 16: 140.
31. Liu J, Wang Y, Liu Y, Mei H, Zhang Q, Gao Z, et al. FOXP2/SOS1/AKT negative feedback loop inhibits cell proliferation in KRAS-mutant colorectal cancer. Apoptosis. 2025; 30: 2455-65.
32. Huang R, Xiang G, Duan X, Wang H, He K, Xiao J. MiR-132-3p inhibits proliferation, invasion and migration of colorectal cancer cells via down-regulating FOXP2 expression. Acta Biochim Pol. 2022; 69: 371-7.
33. Liao P, Huang WH, Cao L, Wang T, Chen LM. Low expression of FOXP2 predicts poor survival and targets caspase-1 to inhibit cell pyroptosis in colorectal cancer. J Cancer. 2022; 13: 1181-92.
34. Çelik E, Bayram C, Denkbaş EB. Chondrogenesis of human mesenchymal stem cells by microRNA loaded triple polysaccharide nanoparticle system. Mater Sci Eng C Mater Biol Appl. 2019; 102: 756-63.
35. He Q, Zhao L, Liu Y, Liu X, Zheng J, Yu H, et al. circ-SHKBP1 regulates the angiogenesis of u87 glioma-exposed endothelial cells through miR-544a/FOXP1 and miR-379/FOXP2 pathways. Mol Ther Nucleic Acids. 2018; 10: 331-48.
36. Prabha S, Arya G, Chandra R, Ahmed B, Nimesh S. Effect of size on biological properties of nanoparticles employed in gene delivery. Artif Cells Nanomed Biotechnol. 2016; 44: 83-91.
37. González-Domínguez I, Puente-Massaguer E, Lavado-García J, Cervera L, Gòdia F. Micrometric DNA/PEI polyplexes correlate with higher transient gene expression yields in HEK 293 cells. N Biotechnol. 2022; 68: 87-96.
38. Wang Y, Ye M, Xie R, Gong S. Enhancing the *in vitro* and *in vivo* stabilities of polymeric nucleic acid delivery nanosystems. Bioconjug Chem. 2019; 30: 325-37.
39. Fröhlich E. The role of surface charge in cellular uptake and cytotoxicity of medical nanoparticles. Int J Nanomedicine. 2012; 7: 5577-91.
40. Grandinetti G, Ingle NP, Reineke TM. Interaction of poly(ethylenimine)-DNA polyplexes with mitochondria: implications for a mechanism of cytotoxicity. Mol Pharm. 2011; 8: 1709-19.

Gazi Med J 2026;37(2):166-171

DOI: http://dx.doi.org/10.12996/gmj.2025.4489

Awareness, Knowledge, Attitudes, and Behaviors of Pediatric Hematology and Oncology Specialists on Preserving Reproductive Health in Children with Cancer: Barriers and Suggestions

Çocuk Hematoloji Onkoloji Uzmanlarının Kanser Tanısı Alan Çocuklarda Üreme Sağlığını Koruma Konusundaki Farkındalık, Bilgi, Tutum ve Davranışları: Engeller ve Öneriler

Melek Yaman Ortaköylü[1], Sonay İncesoy Özdemir[2], Nurdan Taçyıldız[2], Handan Dinçaslan[2], Emel Cabi Ünal[2]

[1]Department of Pediatric Hematology and Oncology, Malatya Training and Research Hospital, Malatya, Türkiye
[2]Department of Pediatric Oncology, Ankara University Faculty, of Medicine, Ankara, Türkiye

ABSTRACT

Objective: To evaluate the awareness, knowledge, attitude, and current practices of pediatric hematology oncology specialists about protecting reproductive health that can be offered to children diagnosed with cancer; to identify the obstacles they encounter in this regard and to make suggestions for the current situation.

Methods: From January 15, 2022, to June 15, 2022, we conducted a cross-sectional survey involving 23 participants recruited via social media who completed electronic surveys.

Results: The majority (78.26%) of participating physicians were female. Most of the participants (95.65%) worked in a metropolitan area. Nearly all (95.65%) of the physicians reported discussing long-term fertility issues with the family of a newly diagnosed cancer patient, with 82.60% always or routinely and 17.39% sometimes recommending fertility preservation methods. The main issues reported in applying fertility-preservation methods were: urgency of treatment (78.26%), financial difficulties and insurance barriers for patients (60.86%), absence of institutional or national standards (60.86%), physicians' inadequate knowledge of surgical methods applicable to pre-adolescent patients (39.13%), and high physician workload (21.73%).

Conclusion: Although 86.95% of surveyed pediatric oncologists agreed to refer all adolescent males to a reproductive health specialist prior

ÖZ

Amaç: Çocuk hematoloji onkoloji uzmanlarının, kanser tanısı alan çocuklara sunulabilecek üreme sağlığını koruma konusundaki farkındalık, bilgi, tutum ve mevcut uygulamalarını değerlendirmek, bu konuda karşılaştıkları engelleri saptamak ve mevcut duruma yönelik önerilerde bulunmaktır.

Yöntemler: Kesitsel anket çalışması olan araştırma, 15 Ocak 2022 ve 15 Haziran 2022 tarihleri arasında, elektronik anketlerin sosyal medya aracılığı ile ulaştırılması sonucunda, araştırmaya katılmayı kabul eden 23 kişi ile gerçekleştirilmiştir.

Bulgular: Araştırmaya katılan hekimlerin %78,3'ü kadın, %21,7'si erkekti. Katılımcıların %95,7'si büyükşehirde çalışmakta, %52,2'si üniversite hastanesinde hizmet vermekteydi. Hekimlerin %95,7'si yeni tanı alan kanserli çocuğun ailesine uzun dönemde yaşanabilecek fertilite ile ilgili sorunları anlattığını, %82,6'sı her zaman/rutin olarak, %17,4'ü bazen fertilite koruma yöntemlerini önerdiğini bildirmiştir. Hekimler en fazla post-pubertal çocuklarda sperm ve oosit kriyoprezervasyonu yöntemlerini kullandıklarını bildirmişlerdir. Fertilite koruma yöntemlerini uygulama konusunda yaşanılan en sık problemler olarak; tedavinin aciliyeti (%76,2), hastaya ait maddi zorluk/sigorta engelleri (%61,9), kurumsal veya ulusal bir standardın olmaması (%61,9), hekimlerin ergenlik öncesi hastalar için uygulanabilecek

Cite this article as: Yaman Ortaköylü M, İncesoy Özdemir S, Taçyıldız N, Dinçaslan H, Cabi Ünal E. Awareness, knowledge, attitudes, and behaviors of pediatric hematology and oncology specialists on preserving reproductive health in children with cancer: barriers and suggestions. Gazi Med J. 2026;37(2):166-171

Address for Correspondence/Yazışma Adresi: Sonay İncesoy Özdemir, Department of Pediatric Oncology, Ankara University Faculty, of Medicine, Ankara, Türkiye
E-mail / E-posta: sincesoy@yahoo.co.uk
ORCID ID: orcid.org/0000-0003-2863-901X

Received/Geliş Tarihi: 08.07.2025
Accepted/Kabul Tarihi: 15.12.2025
Epub: 26.02.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

This study was presented as a poster presentation at the 22nd National Pediatric Cancer Congress, held on March 8–12, 2023, in Antalya, Türkiye.

ABSTRACT

to cancer treatment, only 47.82% reported implementing this practice. Only 43.47% of physicians followed the American Society of Clinical Oncology's 2006 guidelines for preserving fertility in patients with cancer. Identifying barriers in this regard is expected to improve future efforts to preserve reproductive health in our country.

Keywords: Knowledge, behavior, cancer, pediatric oncology, fertility preservation, barriers

ÖZ

cerrahi yöntemler konusunda yeterli bilgiye sahip olmaması (%38,1) ve hekimin iş yükü fazlalığı (%23,8) bildirilmiştir.

Sonuç: Çalışma, ankete cevap veren pediatrik onkologların %86'sının tüm ergenlik çağındaki erkekleri onkolojik tedaviden önce üreme sağlığı konusunda uzmanlaşmış bir doktora sevk etme önerisine katılmalarına rağmen, yalnızca %46'sının bunu uyguladığını bildirdi. Hekimlerin sadece %44'ü, Amerikan Klinik Onkoloji Derneği tarafından 2006 yılında kanser hastalarında doğurganlığın korunması için yayınlanan önerilere uyum bildirmiştir. Ülkemizde çocuk hematoloji ve onkoloji uzmanlarının büyük çoğunluğunun bu konuda bilgi, tutum ve mevcut uygulamalarının literatüre göre olumlu olduğu görülmektedir. Bu konuda yaşanan engellerin saptanmasının, ülkemizde gelecekteki üreme sağlığının korunması çabalarını geliştireceği öngörülmektedir.

Anahtar Sözcükler: Bilgi, davranış, kanser, pediatrik onkoloji, doğurganlığın korunması, engeller

INTRODUCTION

As a result of advances in modern multidisciplinary treatment approaches and supportive care, the cure rate for childhood cancer has exceeded 80% (1). As the number of survivors of childhood cancer continues to rise, there is increasing emphasis on the future quality of life of these individuals in adulthood. One of the primary challenges faced by childhood cancer survivors in adulthood is the risk of infertility (2). Compromised reproductive capacity is a profound consequence of life-preserving interventions for pediatric malignancies (3).

This study aims to assess the knowledge, attitudes, and current practices of pediatric hematology-oncology specialists in our country regarding methods to preserve reproductive health in patients with cancer. The study also aims to identify the obstacles they encounter in this context. Increasing survival rates among childhood cancer survivors highlight the need for a proactive approach to addressing the long-term consequences of cancer treatments, particularly regarding reproductive health. Understanding the perspectives and practices of healthcare professionals in our country is crucial to developing strategies to mitigate the impact of potential infertility on the quality of life of childhood cancer survivors in their adult years.

MATERIALS AND METHODS

The cross-sectional survey, conducted between January 15 and June 15, 2022, was distributed via social media involved 23 participants who consented. The survey instrument, developed by the researchers, was distributed to 152 physicians through social media networks (e-mail, WhatsApp). During the study period, 23 responses were collected, resulting in a response rate of 15%. Electronically administered surveys, completed on a voluntary basis, were automatically recorded in the database and analyzed. This study was conducted in accordance with the Declaration of Helsinki. Ethical approval for this study was obtained from the Ethics Committee of Ankara University (approval number: 2022/24, date: 12.01.2022). The medical doctors who agreed to participate were informed about the study.

Statistical Analysis

Statistical analyses were performed using SPSS 22.0 (SPSS Inc., Chicago IL, USA). Descriptive statistics were used to summarize the data. Continuous variables were presented as mean ± standard deviation or median (minimum–maximum) according to data distribution, while categorical variables were expressed as numbers and percentages. The normality of continuous variables was assessed using visual methods (histograms and probability plots) and the Shapiro–Wilk test.

RESULTS

Of the participating physicians, 18 (78.3%) were female and 5 (21.7%) were male. All but one participant worked in metropolitan areas. Twelve (52.1%) of the participating physicians provided services in university hospitals. Twenty-two physicians (95.7%) reported discussing long-term fertility-related issues with the families of children newly diagnosed with cancer; 19 (82.6%) reported routinely or always recommending fertility preservation (FP) methods, while 4 (17.4%) reported suggesting these methods occasionally (Table 1). Among the 22 physicians recommending FP, 18 (81.8%) stated that they always applied these methods before starting treatment. Two physicians reported applying these methods until the end of the first course of treatment. Physicians reported using sperm and oocyte cryopreservation methods most frequently for post-pubertal children. 20 (90.9%) of the physicians mentioned recommending sperm cryopreservation, 7 (31.8%) recommended testicular tissue cryopreservation, 15 (68.1%) recommended oocyte cryopreservation, 9 (40.9%) recommended ovarian tissue cryopreservation, and 2 (9%) recommended ovarian suppression with gonadotropin-releasing hormone agonists. No physician reported recommending embryo cryopreservation. The most common problems encountered in implementing FP methods were urgency of treatment (76.2%), financial difficulties/insurance barriers for the patient (61.9%), absence of institutional or national standards (61.9%), inadequate physician knowledge about surgical methods applicable to pre-adolescent patients (38.1%), and physician workload (23.8%) (Table 2).

To improve services provided to preserve reproductive health, physicians recommend establishing an official protocol approved by the Ministry of Health and creating public oncofertility centers.

DISCUSSION

This study provides valuable insight into the knowledge, attitudes, and practices of pediatric hematology and oncology specialists regarding the preservation of reproductive health among children with cancer in Türkiye, a country where data on this subject remain limited. The survey reveals both encouraging results and barriers to practice. These findings underscore the need for institutional support and centralized oncofertility units.

The awareness, knowledge, attitudes, and behaviors of pediatric hematology and oncology specialists regarding the preservation of reproductive health in children with cancer are crucial factors that significantly affect patient care and outcomes. Understanding the barriers and recommendations related to this issue can improve the quality of care for pediatric cancer patients.

Pediatric hematology and oncology specialists are typically well aware of the potential impact of cancer treatments on children's reproductive health. They possess knowledge about the mechanisms through which chemotherapy, radiation therapy, and other cancer treatments can affect fertility. However, the depth of awareness and knowledge may vary among practitioners, influenced by factors such as training, experience, and exposure to recent research findings (4,5). Attitudes toward discussing reproductive health preservation in pediatric cancer patients may vary among specialists. While some practitioners may prioritize these discussions and actively engage with patients and their families, others may perceive FP as a less immediate concern than cancer treatment. Attitudes may also be influenced by cultural factors, personal beliefs, and perceptions of the child's and the family's capacity to comprehend reproductive health issues (6). Discussing FP with pediatric patients and their families can be challenging due to the topic's sensitive nature, the need for age-appropriate communication, and varying levels of understanding (7). Pediatric oncologists often face time constraints

Table 1. General characteristics of the physicians participating in the study.

Characteristics	Number (%)
Median age of physicians (year) (least-most)	45 (32-63)
Gender distribution of physicians (male/female)	5/18
Professional title/duty of physicians	
Specialist	8 (34.8)
Dr. faculty member	4 (17.4)
Associate professor	5 (21.7)
Professor	6 (26.1)
Institution where physicians work	
Training and research hospital	11 (47.9)
University hospital	9 (39.1)
Private/foundation university hospital	3 (13)
Annual number of newly diagnosed patients by physicians	
20-40	4 (17.4)
40-60	7 (30.4)
60-80	5 (21.7)
80-100	3 (13)
Over 100	4 (17.4)
Admission age of newly diagnosed patients	
0-18 years	22 (95.7)
0-23 years	1 (4.3)
Is there a reproductive health clinic as a separate unit in the center where you work?	
Yes	15 (65.2)
No	8 (34.8)
Are you informing the family, whose child recently diagnosed with cancer, about potential long-term fertility issues?	
Always/routinely	22 (95.7)
If there are questions from the patient and his/her family	1 (4.3)
Do you recommend fertility preservation methods to the family of your newly diagnosed child with cancer?	
Yes	19 (82.6)
No	0 (0.0)
Sometimes	4 (17.4)

Table 2. Distribution of problems experienced by physicians in applying fertility preservation methods (One physician reported more than one problem).

Problems	Number (%)
Urgency of treatment	16 (76.2)
The lack of infrastructure (reproductive health center) and shortage of physicians in the institution and/or region being worked in	13 (61.9)
The absence of a corporate or national standard	13 (61.9)
Patient age	13 (61.9)
Financial difficulties/insurance obstacles of the patient	13 (61.9)
I do not have sufficient knowledge about surgical methods that can be applied to pre-adolescent patients	8 (38.1)
Psychosocial difficulties	5 (23.8)
Insurance problem for refugee patients	5 (23.8)
Physician's high workload	5 (23.8)
Cultural and religious challenges	3 (14.3)
Lack of patient awareness	2 (9.5)
I do not have sufficient information about the use and effectiveness of fertility preservation methods for female patients	2 (9.5)
Other	1 (4.8)
Legal obstacles	0 (0.0)

during clinic visits, which may limit their ability to thoroughly discuss reproductive health preservation and explore patients' concerns or preferences (8,9). Time constraints during initial consultations, often dominated by urgent treatment decisions, may preclude in-depth discussions concerning future reproductive plans. This underscores the need for institutional support, such as referral pathways, dedicated fertility liaisons, or integrated oncofertility clinics, to streamline the process.

In 2006, the American Society of Clinical Oncology (ASCO) published its recommendations for the preservation of fertility in cancer patients. These recommendations state that oncologists should discuss FP with patients shortly after the initial cancer diagnosis and, if appropriate, refer them to an FP specialist as soon as possible (10,11). In 2011, a study was conducted to determine pediatric oncologists' attitudes and practice models following the publication of ASCO guidelines. The study found that although 86% of pediatric oncologists who responded to the survey agreed with the recommendation to refer all adolescent males to a specialist in reproductive health before oncological treatment, only 46% reported implementing this practice. Additionally, only 44% reported adherence to ASCO guidelines (12). It's necessary to address the areas in which adherence to guidelines is challenging. Effective FP care requires structured protocols that ensure consistent screening for FP needs, facilitate seamless referrals between oncology and fertility services, provide timely access to fertility consultations and preservation options, and offer guidance on navigating financial aspects (13).

In our country, most pediatric hematology and oncology specialists appear to have knowledge, attitudes, and practices consistent with the literature on this issue. This positive outcome could create a hopeful foundation for strengthening efforts to preserve reproductive health. In 2015, the Society of Reproductive Health and Infertility aimed to increase awareness of FP among healthcare professionals by preparing a guide on preserving fertility in Türkiye. The aim of this project is to inform and guide patients based on current scientific evidence and to ensure access to centers implementing this practice. FP in patients diagnosed with cancer requires a multidisciplinary approach. In addition to fertility specialists involved in the care of patients requiring FP, awareness and knowledge should be increased among healthcare professionals across specialties who care for patients diagnosed with cancer (14,15).

In our study, the majority of respondents were female physicians (78.3%). Although our study did not perform statistical analyses of gender-based differences in practice, the physician's gender may influence the likelihood of initiating discussions on sensitive issues such as reproductive health. Future research could investigate whether this factor affects the frequency or depth of FP counseling in pediatric oncology.

The integration of reproductive health preservation into pediatric cancer care may face various obstacles, with resource limitations constituting a primary challenge. Access to FP services, such as sperm or egg banking facilities, may be limited in certain geographic areas or healthcare settings, posing logistical challenges for patients and providers. In the United Kingdom (UK), the Children's Cancer and Leukemia Group's Late Effects Working Group examined disparities in the provision of FP options to young patients with cancer. They found variability in the provision of FP for children with cancer across the country. The absence of dedicated government funding to support adherence to global standards has led to inconsistent care depending on location within the UK (16).

To overcome these barriers and improve the integration of reproductive health preservation into pediatric cancer care, several suggestions are proposed. Providing pediatric hematology and oncology specialists with ongoing education and training on FP techniques, guidelines, and communication strategies can improve their confidence and competence in addressing reproductive health concerns (17). Children's Oncology Group (COG) has developed a stratification system for gonadal dysfunction and infertility based on leukemia and lymphoma phase 3 protocols conducted between

2000 and 2022, to provide a standardized guide for assigning gonadotoxic risk. This comprehensive guide serves as a valuable tool to enhance and standardize reproductive health counseling for patients undergoing COG-based leukemia/lymphoma care, both at diagnosis and during survivorship (18). The PanCareLIFE Consortium, in collaboration with the International Late Effects of Childhood Cancer Guideline Harmonization Group, developed a clinical practice guideline tailored for young female patients with cancer diagnosed during childhood, adolescence, or young adulthood (up to age 25). This guideline offers comprehensive advice on evaluating fertility risks and outlines various options for preserving fertility (19). Numerous guidelines have been issued regarding this matter (20-22). Each country should develop similar guidelines tailored to their own protocols. To improve the integration of FP into clinical care, participants suggested national protocols and centralized oncofertility units. Facilitating collaboration between pediatric oncology teams, fertility specialists, and other relevant healthcare professionals can enhance the comprehensive care of pediatric cancer patients, ensuring that their reproductive health needs are addressed holistically. The creation of referral algorithms endorsed by Turkish pediatric hematology and oncology groups could also foster uniformity in care. Patient education and family education also play a crucial role in this process. Developing culturally sensitive, age-appropriate educational materials and resources for pediatric cancer patients and their families can empower those families to make informed decisions about FP options and to advocate for their reproductive health needs. Embedding fertility education into survivorship care plans may further reinforce its importance. Healthcare institutions can support the integration of the preservation of reproductive health into pediatric cancer care by allocating resources to FP services, establishing referral pathways to fertility specialists, and fostering a culture of open communication regarding reproductive health. By addressing these suggestions and overcoming barriers, pediatric hematology and oncology specialists can enhance their ability to provide comprehensive, patient-centered care that considers the long-term reproductive health outcomes of children with cancer. Ultimately, prioritizing the preservation of reproductive health in pediatric cancer care can improve quality of life and survivorship outcomes for pediatric cancer patients.

Study Limitations

Our study has some limitations. One of these is the need to reach a greater number of pediatric oncologists nationwide to comprehensively assess their knowledge, attitudes, and practices regarding this matter. Additionally, a prospective study on this subject that would prevent recall bias will provide more reliable data. (23,24).

CONCLUSION

Identifying existing barriers and strengthening related efforts can help to develop more effective strategies for preserving the reproductive health of individuals undergoing cancer treatment in our country in the future. In this context, measures such as education and awareness-raising efforts among healthcare professionals, policy changes, and resource allocation can be implemented.

Ethics

Ethics Committee Approval: This study was conducted in accordance with the Declaration of Helsinki. Ethical approval for this study was obtained from the Ethics Committee of Ankara University (approval number: 2022/24, date: 12.01.2022).

Informed Consent: The medical doctors who agreed to participate were informed about the study.

Footnotes

Authorship Contributions

Concept: S.İ.Ö., Design: S.İ.Ö., Data Collection or Processing: S.İ.Ö., Analysis or Interpretation: S.İ.Ö., N.T., H.D., E.C.Ü. Literature Search: M.Y.O., S.İ.Ö., Writing: M.Y.O., S.İ.Ö.

REFERENCES

1. Hilgendorf I, Bergelt C, Bokemeyer C, Kaatsch P, Seifart U, Stein A, et al. Long-term follow-up of children, adolescents, and young adult cancer survivors. Res Treat. 2021; 44: 184-9.
2. Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, et al. Chronic health conditions in adult survivors of childhood cancer. N Engl J Med. 2006; 355: 1572-82.
3. Vadaparampil S, Quinn G, King L, Wilson C, Nieder M. Barriers to fertility preservation among pediatric oncologists. Patient Educ Couns. 2008; 72: 402-10.
4. Zhu LX, Jin L, Jiang JH, Yang L, Fang ZS, Wang M, et al. Update knowledge assessment and influencing predictor of female fertility preservation in oncologists. Curr Med Sci. 2022; 42: 824-31.
5. Mailankody S, Bajpai J, Arora PR, Sreedharan R, Chitalkar P, Kurkure P, et al. Oncofertility and pregnancy in adolescent and young adult cancers: physicians' knowledge and preferences in India. JCO Glob Oncol. 2024; 10: e2300205.
6. Al Ghaithi A, Al Rashdi E, Al Shukri M, Al Ghabshi R, Albalushi H. Oncologists' knowledge, practice and attitude toward fertility preservation: A National Survey. Life (Basel). 2023; 13: 801.
7. Vesali S, Navid B, Mohammadi M, Karimi E, Omani-Samani R. Little information about fertility preservation is provided for cancer patients: A survey of oncologists' knowledge, attitude and current practice. Eur J Cancer Care (Engl). 2019; 28: e12947.
8. Boone AN, Arbuckle JL, Ye Y, Wolfson JA. How accurate Is oncologist knowledge of fertility preservation options, cost, and time in female adolescents and young adults? J Adolesc Young Adult Oncol. 2023; 12: 110-7.
9. Melo C, Fonseca A, Silva C, Almeida-Santos T, Canavarro MC. Portuguese oncologists' practices regarding female fertility preservation: Which barriers most relate to these practices? Eur J Cancer Care (Engl). 2018; 27: e12812.
10. Lee SJ, Schover LR, Partridge AH, Patrizio P, Wallace WH, Hagerty K, et al. American Society of Clinical Oncology recommendations on fertility preservation in cancer patients. J Clin Oncol. 2006; 24: 2917-31.
11. Loren AW, Mangu PB, Beck LN, Brennan L, Magdalinski AJ, Partridge AH, et al. Fertility preservation for patients with cancer: American Society of Clinical Oncology Clinical Practice guideline update. J Clin Oncol. 2013; 31: 2500-10.
12. Köhler TS, Kondapalli LA, Shah A, Chan S, Woodruff TK, Brannigan RE. Results from the survey for preservation of adolescent reproduction (SPARE) study: gender disparity in delivery of fertility preservation

message to adolescents with cancer. J Assist Reprod Genet. 2011; 28: 269-77.

13. Yang EH, Strohl HB, Su HI. Fertility preservation before and after cancer treatment in children, adolescents, and young adults. Cancer. 2024; 130: 344-55.
14. Yağız Altıntaş R, Kavlak O. Kanser Tanısı Alan Kadınlarda Fertilitenin Korunması ve Hemşirenin Rolü. İKÇÜSBFD. 2023; 8: 551-5.
15. Üreme Sağlığı ve İnfertilite Derneği. TSRM Klavuzu 4-Üremenin Korunması [Internet]. 2015 Nov [cited 2026 Jan 27]. Available from: https://www.tsrm.org.tr/
16. Newton HL, Picton HM, Friend AJ, Hayden CM, Brougham M, Cox R, et al. Inconsistencies in fertility preservation for young people with cancer in the UK. Arch Dis Child. 2022; 107: 265-70.
17. Lampic C, Wettergren L. Oncologists' and pediatric oncologists' perspectives and challenges for fertility preservation. Acta Obstet Gynecol Scand. 2019; 98: 598-603.
18. Close A, Burns K, Bjornard K, Webb M, Chavez J, Chow EJ, et al. Fertility preservation in pediatric leukemia and lymphoma: a report from the Children's Oncology Group. Pediatr Blood Cancer. 2023; 70: e30407.
19. Mulder RL, Font-Gonzalez A, Hudson MM, van Santen HM, Loeffen EAH, Burns KC, et al. Fertility preservation for female patients with childhood, adolescent, and young adult cancer: recommendations from the PanCareLIFE Consortium and the International Late Effects of Childhood Cancer Guideline Harmonization Group. Lancet Oncol. 2021; 22: e45-56.
20. Halpern JA, Hill R, Brannigan RE. Guideline based approach to male fertility preservation. Urol Oncol. 2020; 38: 31-5.
21. ESHRE Guideline Group on Female Fertility Preservation; Anderson RA, Amant F, Braat D, D'Angelo A, Chuva de Sousa Lopes SM, et al. ESHRE guideline: female fertility preservation. Hum Reprod Open. 2020; 2020: hoaa052.
22. Roberts JE, Benoit J, Foong S, Saumet J, Korkidakis A, Marr K, et al. Fertility preservation in patients undergoing gonadotoxic treatments: a Canadian Fertility and Andrology Society Clinical Practice guideline. Reprod Biomed Online. 2024; 48: 103767.
23. Coker Appiah L, Fei YF, Olsen M, Lindheim SR, Puccetti DM. Disparities in female pediatric, adolescent and young adult oncofertility: a needs assessment. Cancers (Basel). 2021; 13: 5419.
24. Trost LW, Brannigan RE. Oncofertility and the male cancer patient. Curr Treat Options Oncol. 2012; 13: 146-60.

Gazi Med J 2026;37(2):172-176

DOI: http://dx.doi.org/10.12996/gmj.2026.4392

The Investigation of NT-proBNP Values in Third-Trimester Gestational Diabetes Women Without Cardiovascular Disease

Kardiyovasküler Hastalığı Olmayan Üçüncü Trimesterdeki Gestasyonel Diabetes Mellitus Olan Kadınlarda NT-Probnp Değerlerinin Araştırılması

Meltem Altınsoy[1], Ezgi Polat Ocaklı[1], Sadun Sucu[2], Burak Menteş[1], Belma Kalaycı[1], Hüseyin Demirci[3]

[1]Department of Cardiology, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Ankara, Türkiye
[2]Department of Perinatology, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Ankara, Türkiye
[3]Department of Endocrinology, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Ankara, Türkiye

ABSTRACT

Objective: The clinical utility and importance of natriuretic peptides during pregnancy in women with gestational diabetes mellitus (GDM) and no known cardiovascular disease are not well-known. We aimed to investigate N-terminal pro-B-type natriuretic peptide (NT-proBNP) values in women with gestational diabetes in the third trimester without cardiovascular disease.

Methods: This study included pregnant women in the third trimester who were referred to our outpatient cardiology clinic for dyspnea assessment between August and September 2024. Patients were classified into two groups: the GDM group (n = 31) and the control group (n = 20). NT-proBNP values and clinical and laboratory parameters were compared between the two groups.

Results: No differences were observed between the two groups with respect to age, height, weight changes during pregnancy, or biochemical and echocardiographic parameters. Weight and body mass index were significantly elevated in the GDM group. NT-proBNP values were higher in the GDM group than in the control group. However, there were no statistically significant differences in NT-proBNP values between the two groups [28.9 (4–258) pg/mL vs. 18 (4–230) pg/mL, p = 0.331].

Conclusion: We found that NT-proBNP values were similar in gestational and non-gestational pregnant women in the third trimester without cardiovascular disease.

Keywords: Gestational diabetes mellitus, NT-proBNP, body mass index, third trimester

ÖZ

Amaç: Kardiyovasküler hastalığı olmayan gestasyonel diyabeti olan gebelerde natriüretik peptitlerin klinik kullanımı ve önemi iyi bilinmemektedir. Biz bu çalışmada üçüncü trimesterdeki gestasyonel diyabetes mellitusu (GDM) olan ve kardiyovasküler hastalığı olmayan kadınlarda N-terminal pro-B tipi natriüretik peptid (NT-proBNP) değerlerini ile ilgili boşluğu doldurmayı amaçladık. Bu konu kapsamlı bir şekilde araştırılmamıştır.

Yöntemler: Nefes darlığı ile ayaktan kardiyoloji kliniğine Ağustos-Eylül 2024 tarihlerinde başvuran üçüncü trimesterdeki gebe kadınlar çalışmaya dahil edildi. Hastalar iki gruba ayırıldı: GDM grup (n = 31) ve kontrol grubu (n = 20).

Bulgular: Yaş, boy, gebelikteki kilo değişimi, biyokimyasal ve ekokardiyografik parametreler bakımından iki grup arasında fark saptanmadı. Kilo ve vücut kitle indeksi GDM grubunda anlamlı olarak yüksek saptandı. Numerik olarak NT-proBNP değerleri GDM grubunda daha yüksek saptandı. Ancak istatistiksel olarak iki grup arasında NT-proBNP değerleri açsından anlamlı bir fark yoktu [28.9 (4-258) pg/mL vs. 18 (4-230) pg/mL, p = 0.331].

Sonuç: Çalışmamızda NT-proBNP değerlerini kardiyovasküler hastalığı olmayan üçüncü trimester gebelerde gestasyonel diyabet olsun olmasın benzerdi.

Anahtar Sözcükler: Gestasyonel diabetes mellitus, NT-proBNP, vücut kitle indeksi, üçüncü trimester

Cite this article as: Altınsoy M, Polat Ocaklı E, Sucu S, Menteş B, Kalaycı B. The investigation of NT-proBNP values in third-trimester gestational diabetes women without cardiovascular disease. Gazi Med J. 2026;37(2):172-176

Address for Correspondence/Yazışma Adresi: Meltem Altınsoy, Department of Cardiology, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Ankara, Türkiye
E-mail / E-posta: meltemaltinsoy@hotmail.com
ORCID ID: orcid.org/0000-0002-3494-5999

Received/Geliş Tarihi: 06.02.2025
Accepted/Kabul Tarihi: 26.01.2026
Epub: 12.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

INTRODUCTION

Pregnancy is generally a physiological state characterized by enhanced beta-cell function and insulin resistance, mediated primarily by placental secretion of diabetogenic hormones (1). These changes develop and are most prominent during the third trimester. Gestational diabetes mellitus (GDM) develops in pregnant people whose pancreatic beta-cell function is insufficient to overcome the insulin resistance associated with pregnancy (2).

Type 2 diabetes mellitus is a chronic disease and one of the main risk factors for heart failure (3). Despite normal heart function and structure, diabetes mellitus increases the risk of heart failure. However, unlike type 2 diabetes mellitus, GDM is generally a temporary condition. GDM is also associated with increased risks of adverse pregnancy outcomes and their morbidities and metabolic effects. Furthermore, there is no well-established relationship between GDM and natriuretic peptides that indicates elevated ventricular filling pressure.

Despite the wide use of natriuretic peptides in the non-pregnant population for the diagnosis and prognosis of cardiovascular disease, the clinical utility of natriuretic peptides in pregnant women remains less well understood. There are no recognized reference values for natriuretic peptides in pregnant women. Some physiologic adaptations during pregnancy may impact natriuretic peptide levels. N-terminal pro-B-type natriuretic peptide (NT-proBNP) may be produced by fetal membranes and may generate myometrial tissue during pregnancy (4). Moreover, pregnant women undergo several hemodynamic and cardiac structural changes, including increases in plasma volume and cardiac output and enlargement of the left and right ventricles (5). These changes may impact NT-proBNP levels during pregnancy. Furthermore, the clinical utility of natriuretic peptides in pregnant women with GDM is not well known. In contrast to prior studies, which primarily evaluated unselected or asymptomatic pregnant women, the present study focuses specifically on pregnant women in the third trimester referred for cardiology evaluation because of dyspnea and compares NT-proBNP levels between women with and without GDM, all without known cardiovascular disease.For this reason, we aimed to investigate the effect of NT-proBNP values in third-trimester GDM patients.

MATERIALS AND METHODS

Study Population

Pregnant women in the third trimester who were referred from the obstetrics department to the cardiology outpatient clinic were enrolled in the study between August and September 2024. We enrolled 51 pregnant women in the third trimester who were aged and gt; 18 years. Both the GDM and control groups comprised pregnant women in the third trimester who were referred to the cardiology outpatient clinic for evaluation of dyspnea. Participants with a history of cardiovascular and systemic diseases such as hypertension, coronary artery disease, valvular heart disease (including more than mild mitral and/or tricuspid regurgitation), chronic renal disease, and chronic inflammatory disease were excluded. Additionally, twin pregnancies were excluded. Women were excluded if they had preexisting diabetes, an abnormal glucose screening result before 24 weeks' gestation, prior gestational diabetes, a history of stillbirth, or multifetal gestation. All women with GDM received medical nutrition therapy. All the women who participated in the study provided written informed consent. Approval was obtained from the Ethics Board of the University of Health Sciences Türkiye, Ankara Etlik City Hospital (approval number: AESH-BADEK-2024-788, date: 30.10.2024).

Measurement of NTproBNP, laboratory, and echocardiographic parameters NT-proBNP was measured in a blood sample (Roche Diagnostics) at our outpatient cardiology clinic. Additionally, hemograms, hemoglobin A1c (HbA1c), and routine biochemical tests, including fasting glucose, urea, creatinine, uric acid, and electrolytes, were recorded.

GDM was diagnosed using the Carpenter–Coustan criteria at the 24th week of gestation (6). According to these criteria, GDM was defined as a fasting glucose level of greater than or equal to 95 mg per deciliter (5.3 mmol per liter) and two or more timed glucose measurements exceeding the following thresholds: 1-hour: 180 mg per deciliter (10.0 mmol per liter); 2-hour: 155 mg per deciliter (8.6 mmol per liter); and 3-hour: 140 mg per deciliter (7.8 mmol per liter). Pregnant women with a negative 50-g glucose challenge screening test formed the control group.

Transthoracic echocardiography was performed in all study patients using a GE S60 ultrasound system. The Investigated echocardiographic parameters were the dimensions of the left atrium; the end-diastolic and end-systolic dimensions; the interventricular septum and posterior wall; and the ejection fraction.

Statistical Analysis

SPSS 19.0 software was used for statistical analyses. Shapiro-Wilk test was used to determine whether the data were normally distributed. For continuous variables, mean ± standard deviation or median (minimum–maximum) values were used; for categorical variables, frequencies or percentages were used where appropriate. An independent-samples t-test or the Mann-Whitney U test was used to compare continuous variables between two groups. Spearman's correlation analysis was performed to determine the relationship between continuous variables. $p < 0.05$ was considered statistically significant for all tests.

RESULTS

Thirty-one patients in the GDM group and twenty in the control group were enrolled in the study. Baseline characteristics and laboratory parameters are shown in Table 1. Age was similar between the two groups (32 ± 5.8 vs. 31.6 ± 4.6; p = 0.834). Similarly, no differences were found in current weight, weight changes during pregnancy, and weeks of pregnancy between the two groups. Weight and body mass index (BMI) were significantly high in the GDM group. Kidney function, electrolytes, and hemogram parameters were similar between the two groups. Although NT-proBNP levels were numerically higher in the GDM group, the difference was not statistically significant [28.9 (4–258) vs. 18 (4–230) pg/mL; p = 0.331]. As mentioned before, all women with GDM received medical nutrition therapy. Only five women with GDM were on insulin therapy, which indicates insulin-requiring GDM (irGDM). Due to the small number of irGDM cases, we did not separate the GDM group into "received medical nutrition" and "irGDM" subgroups.

Fasting glucose was higher in the GDM group, as expected (97 ± 19 vs. 84 ± 15; p = 0.050). The HbA1c level was slightly higher in the GDM group than in the control group (5.0% ± 0.39 vs. 4.5%, p < 0.001). However, both HbA1c values were within normal limits.

All echocardiographic parameters considered similar between the two groups are shown in Table 2. Correlation analyses were performed to assess the relationships between NT-proBNP values and BMI, HbA1c, and fasting glucose (Table 3). We did not observe any correlation between NT-proBNP values and the variables considered.

DISCUSSION

In our study, NT-proBNP values did not differ between gestational and non-gestational diabetes in the third trimester. Normal levels of NT-proBNP are not well established, and there are no recognized reference values in pregnancy. However, NT-proBNP levels in pregnant women were approximately twice those in non-pregnant women, particularly in the first trimester (7). It may be related to the volume status and to placental production of NT-proBNP. However, the clinical utility of natriuretic peptides during pregnancy in women with or without GDM is not well known. In addition, whether a relationship exists between NT-proBNP and glucose metabolism merits consideration.

BNP and its inactive N-terminal fragment (NT-proBNP) are released in response to cardiac strain and are commonly used to evaluate suspected heart failure. The systemic effects of natriuretic peptides are well described in the literature, including diuresis, natriuresis, vasodilatation, and regulation of salt-water balance and maintenance of volume status (8). It has been suggested that the natriuretic peptide system is involved in glucose metabolism and plays a role in diabetes. Experimental data suggest that low levels of atrial natriuretic peptide (ANP) promote the development of insulin resistance and diabetes through activation of the renin-angiotensin system, leading to increased oxidative stress and inflammatory responses, disrupting signaling pathways between insulin and angiotensin and impairing glucose transport (9).

Chang et al. (10) found a strong association between NT-proBNP levels and both heart failure and preeclampsia. However, they did not find any significant differences in NT-proBNP levels in pregnancies complicated by arrhythmia, pregnancy-induced hypertension, postpartum hemorrhage, or adverse fetal/neonatal events.

It is well-established in the literature that obesity is inversely associated with NT-proBNP levels, effectively suppressing them (11). In our study, despite the GDM group having a significantly higher BMI, they exhibited numerically higher, though not statistically significant, NT-proBNP levels. This suggests that the volume overload or subclinical effects induced by GDM may have attenuated the

Table 1. Baseline characteristics of all study patients.

Variables	GDM group (n = 31)	Control group (n = 20)	p-value
Age, years	32 ± 5.8	31.6 ± 4.6	0.834
Weight, kg	83 ± 10	72 ± 13	0.004
Height, cm	161.2 ± 7	161 ± 6	0.958
BMI, kg/cm^2	32 ± 4	27 ± 5.3	0.005
HbA1c, %	5.0 ± 0.39	4.5 ± 0.36	<0.001
Weight changes during pregnancy, kg	10.2 ± 5.9	8 ± 5.7	0.229
Week of pregnancy	31.0 ± 3.2	30.2 ± 3.5	0.224
Fasting glucose, mg/dL	97 ± 19	84 ± 15	0.050
NT-proBNP, pg/mL	28.9 (4–258)	18 (4–230)	0.331
Urea, mg/dL	13 (7–28)	15 (6–20)	0.117
Creatinin, mg/dL	0.5 ± 0.1	0.5 ± 0.08	0.958
Sodium, mEq/L	137 ±1.7	136 ± 1.5	0.134
Potassium, mEq/L	4.1 ± 0.30	4.1 ± 0.37	0.822
TSH	1.98 ± 0.8	2.3 ± 1	0.250
Uric acide, mg/dL	3.9 ± 0.8	3.5 ± 1.3	0.315
White blood count, 10^3/mcL	10.2 ± 2.9	10.6 ± 2.6	0.639
Hemoglobin, g/dL	11.4 ± 2.1	11.7 ± 1.2	0.601
MCV, fL	89 (80–97)	88 (60–98)	0.815
RDW, fL	42 ± 10	44 ± 5	0.476
Platelet count, 10^3/mcL	228 ± 54	231 ± 47	0.882
MPV, fL	10 (9–12)	11 (9–13)	0.103
PCT, %	0.27 ± 0.11	0.24 ± 0.05	0.359

BMI: Body mass index, GDM: Gestational diabetes mellitus, MCV: Mean corpuscular volume, PCT: Plateletcrit, RDW: Redcell distrubation width, TSH: Thyroid stimulating hormone, HbA1c: Hemoglobin A1c.

Table 2. Echocardiographic parameters among two groups.

Variables	Gestational diabetes mellitus group (n = 31)	Control group (n = 20)	p-value
Left atrial diamater, mm	34 ± 2.8	33 ± 2.7	0.361
End diastolic diamater, mm	45 (40–50)	44 (37–49)	0.442
End sistolic diamater, mm	25 (24–30)	27 (24–30)	0.162
Interventricular septum, mm	9 (8–11)	9 (7–11)	0.351
Posterior wall, mm	9 (8–10)	9 (7–11)	0.398
Left ventricular ejection fraction %	60.7 ± 1.8	60	0.110

Table 3. Bivariate correlation analysis with NT-proBNP value.

	r	p
HgA1c	0.043	0.786
BMI	0.027	0.071
Fasting glucose	0.075	0.231

r: correlation coefficient, BMI: Body mass index, HbA1c: Hemoglobin A1c, NT-proBNP: N-terminal pro-B-type natriuretic peptide.

known suppressive effect of obesity. Minhas et al. (11) investigated NT-proBNP levels in pregnant women without known cardiovascular disease, stratified by pregnancy status and trimester. According to their findings, the prevalence of elevated NT-proBNP (>125 pg/mL) was 20.0% among pregnant women in the first trimester, 2.4% among women in the third trimester, and 8.0% among non-pregnant women. Additionally, NT-proBNP was 44% higher in the first trimester of pregnancy than in non-pregnant women [absolute difference, 26.4 pg/mL; 95% confidence interval (CI): 11.2–41.6]. Further, among pregnant women, adjusted NT-proBNP was 46% lower [absolute difference 22.2 pg/mL (95% CI: 36.9 to 7.5)] in women in the third trimester compared with women in the first trimester. According to the study researcher's explanation that NT-proBNP decreases to lower levels by the third trimester, increased ventricular size and remodeling may contribute to the resolution of the initial rise in natriuretic peptide levels. They also found that NT-proBNP was inversely associated with BMI and systolic blood pressure.

According to the current literature, a study similar to ours exists. In this study, NT-proBNP was significantly lower in patients with irGDM than in controls in the subgroup analysis (35 ± 25 pg/mL vs. 53 ± 43 pg/mL; p = 0.012). However, they found that NT-proBNP levels were similar between women with GDM receiving medical nutrition therapy and the control group, as our study findings indicated. If we compare our study with the mentioned study, they did not exclude cardiovascular disease or report echocardiographic parameters. The mean gestational age was higher than in the other study (30 vs. 28 weeks). Due to the small number of women with irGDM, we did not perform any statistical analyses comparing women with irGDM and those with medical nutrition therapy GDM.

From a cardiology perspective, dyspnea in late pregnancy is common but non-specific and often prompts referral to exclude underlying cardiac disease. In this symptomatic third-trimester cohort, NT-proBNP values did not differ significantly between women with and without GDM, suggesting that GDM itself does not materially confound NT-proBNP levels in this setting. Accordingly, when evaluating pregnant women in the third trimester who present with dyspnea and have no known cardiovascular disease, NT-proBNP may retain its supportive value in identifying those who warrant closer cardiac evaluation, rather than reflecting glycemic status alone. These findings may help clinicians interpret NT-proBNP more confidently in pregnant women with dyspnea and may underscore that biomarker results should be integrated with clinical assessment and echocardiography.

The present study has several strengths. First, its design enabled standardized and systematic collection of clinical, laboratory, and echocardiographic data. However, several limitations should be acknowledged. Given the small number of women with irGDM, we could not perform a subgroup analysis. Another limitation was that we did not evaluate hemodynamic parameters, such as pulse and systolic and diastolic blood pressure, which may affect NT-proBNP levels.

Additionally, we did not evaluate the concentrations of other enzymes or components of the natriuretic pathway that might help explain the mechanisms underlying our findings, such as BNP, ANP, and cyclic guanosine monophosphate (cGMP).

These findings may serve as a basis for future studies involving larger cohorts, including women in the final month of pregnancy, when the risk of pregnancy-associated cardiomyopathy is particularly high.

Study Limitations

This study has certain limitations. First, the fact that the research was conducted at a single center and had a relatively small sample size (n = 51) may limit the generalizability of the findigs. Due to the cross-sectional design of the study, it is not possible to assess the cause-effect relationship and the temporal course of changes in NT-proBNP levels. Second, cardiac hemodynamic indicators such as heart rate, diastolic and systolic blood pressure were not examined in the pregnant women included in the study. Investigating these parameters could have helped to better understand the relationship between NT-proBNP levels and cardiac load. Finally, the number of GDM cases receiving insulin therapy was low (n = 5); therefore, subgroup analyses to evaluate the effect of insulin therapy on NT-proBNP could not be performed. Additionally, physical factors that could affect NT-proBNP levels, such as electrolyte balance, obesity

level, and changes in plasma volume, were not controlled. Cardiac hemodynamic parameters (heart rate, systolic/diastolic blood pressure) were not evaluated. Other biomarkers in the natriuretic pathway, such as BNP, ANP, and cGMP, were not analyzed in the study.

CONCLUSION

In conclusion, NT-proBNP values did not differ between gestational and non-gestational pregnant women in the third trimester without cardiovascular disease. For this reason, our findings suggest that third-trimester NT-proBNP levels in women with uncomplicated GDM are not significantly different from those in healthy pregnant controls, indicating that GDM does not introduce additional bias in the cardiac evaluation of these patients.

Ethics

Ethics Committee Approval: Approval was obtained from the Ethics Board of the University of Health Sciences Türkiye, Ankara Etlik City Hospital (approval number: AESH-BADEK-2024-788, date: 30.10.2024).

Informed Consent: All the women who participated in the study provided written informed consent.

Footnotes

Authorship Contributions

Surgical and Medical Practices: S.S., Concept: B.K., E.P.O., Design: B.K., E.P.O., Data Collection or Processing: M.A., B.M., Analysis or Interpretation: B.K., Literature Search: M.A., B.K., E.P.O., Writing: M.A., B.K.

REFERENCES

1. Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The pathophysiology of gestational diabetes mellitus. Int J Mol Sci. 2018; 19: 3342.
2. Proceedings of the 4th International Workshop-Conference on Gestational Diabetes Mellitus. Chicago, Illinois, USA. 14-16 March 1997. Diabetes Care. 1998; 21 Suppl 2: B1-167.
3. He J, Ogden LG, Bazzano LA, Vupputuri S, Loria C, Whelton PK. Risk factors for congestive heart failure in US men and women: NHANES I epidemiologic follow-up study. Arch Intern Med. 2001; 161: 996-1002.
4. Carvajal JA, Delpiano AM, Cuello MA, Poblete JA, Casanello PC, Sobrevia LA, et al. Brain natriuretic peptide (BNP) produced by the human chorioamnion may mediate pregnancy myometrial quiescence. Reprod Sci. 2009; 16: 32-42
5. Burlingame JM, Yamasato K, Ahn HJ, Seto T, Tang WHW. B-type natriuretic peptide and echocardiography reflect volume changes during pregnancy. J Perinat Med. 2017; 45: 577-83.
6. Carpenter MW, Coustan DR. Criteria for screening tests for gestational diabetes. Am J Obstet Gynecol. 1982; 144: 768-73.
7. Hameed AB, Chan K, Ghamsary M, Elkayam U. Longitudinal changes in the B-type natriuretic peptide levels in normal pregnancy and postpartum. Clin Cardiol. 2009; 32: E60-2.
8. Cauliez B, Berthe MC, Lavoinne A. Le BNP: aspects physiologiques, biologiques et cliniques [Brain natriuretic peptide: physiological, biological and clinical aspects]. Ann Biol Clin (Paris). 2005; 63: 15-25.
9. Velloso LA, Folli F, Perego L, Saad MJ. The multi-faceted cross-talk between the insulin and angiotensin II signaling systems. Diabetes Metab Res Rev. 2006; 22: 98-107.
10. Chang SA, Khakh P, Janzen M, Lee T, Kiess M, Rychel V, et al. Trending cardiac biomarkers during pregnancy in women with cardiovascular disease. Circ Heart Fail. 2022; 15: e009018.
11. Minhas AS, Rooney MR, Fang M, Zhang S, Ndumele CE, Tang O, et al. Prevalence and correlates of elevated NT-proBNP in pregnant women in the general U.S. population. JACC Adv. 2023; 2: 100265.

Gazi Med J 2026;37(2):177-182

DOI: http://dx.doi.org/10.12996/gmj.2026.4561

Comparison of Two and Three-Dimensional Ultrasound in the Diagnosis of Polycystic Ovary Morphology: A Cross Sectional Study in Unselected Women

Polikistik Over Morfolojisinin Tanısında İki ve Üç Boyutlu Ultrasonografinin Karşılaştırılması: Seçilmemiş Kadınlarda Kesitsel Bir Çalışma

Zuhal Yapıcı Coşkun[1], Gürkan Bozdağ[2], Bülent Okan Yıldız[3], İbrahim Esinler[4], Hakan Yaralı[4,5]

[1]Department of Obstetrics and Gynecology, Ankara Bilkent City Hospital, Ankara, Türkiye
[2]Bahçeci Health Group IVF Center, İstanbul, Türkiye
[3]Division of Endocrinology and Metabolism, Department of Internal Medicine, Hacettepe University Faculty of Medicine, Ankara, Türkiye
[4]Department of Obstetrics and Gynecology, Hacettepe University Faculty of Medicine, Ankara, Türkiye
[5]Anatolia IVF Center, Ankara, Türkiye

ABSTRACT

Objective: To compare the visual appearance of the ovaries with two dimensional and three dimensional (3D) ultrasound (US) modalities in unselected women of reproductive age and to evaluate whether 3D changes the likelihood of polycystic ovary syndrome (PCOS) diagnosis.

Methods: This cross-sectional study was conducted in patients diagnosed within an unselected population of women (n = 115). The primary outcome measures were antral follicle count (AFC) and ovarian volume (OV), assessed using two different sonographic modalities. The agreement of these two sonographic modalities was tested using a Bland–Altman plot.

Results: Whereas the mean AFC with both sonographic modalities was 10.3 ± 4.8 vs. 11.2 ± 5.5 (p = 0.61), the OV estimated were 7.5 ± 3.0 vs. 9.6 ± 4.0 (p < 0.001) mL respectively. The mean bias, representing the upper and lower limits of agreement between the modalities for estimation of AFC, was – 0.17 (-8.45 to 8.11). The respective figure for OV was -1.99 mL (-7.72 to 3.75). Therefore, 3D revealed OV values that were 22.5% larger and identified 2 additional cases of PCOS according to the Rotterdam criteria.

Conclusion: 3D US yields larger OV values and classifies slightly more women as having PCOS. Although concerns related to the diagnosis of PCOM mainly focus on AFC, those findings also underscore the reliability of OV as a diagnostic parameter for the syndrome. These

ÖZ

Amaç: Bu çalışmanın amacı fertil çağda olan seçilmemiş kadın populasyonunda iki boyutlu (2D) ve üç boyutlu (3D) ultrasonografi (US) ile yumurtalıkların görünümünü karşılaştırmak ve 3D'nin polikistik over sendromu (PKOS) tanı olasılığını değiştirip değiştirmediğini değerlendirmektir.

Yöntemler: Bu kesitsel çalışma seçilmemiş bir kadın popülasyonunda PKOS tanısı alan hastalar arasında (n = 115) yapılmıştır. Birincil sonuç ölçümleri 2 boyutlu ve 3 boyutlu US ile antral folikül sayısı (AFC) ve over hacmi (OV) idi. 2D ve 3D US'nin uyumu Bland-Altman grafiği ile test edilmiştir.

Bulgular: 2D ve 3D ile ortalama AFC 10,3 ± 4,8 ve 11,2 ± 5,5 (p = 0,61) iken, 2D ve 3D ile tahmin edilen OV 7,5 ± 3,0 ve 9,6 ± 4,0 (p < 0,001) mL idi. AFC tahmini için 2D ve 3D arasındaki ortalama sapma (uyumun üst ve alt sınırları) -0,17 (-8,45 ila 8,11) idi. OV için ilgili rakam -1,99 (-7,72 ila 3,75) mL idi. Dolayısıyla, 3D OV için %22,5 daha büyük değerler ortaya koymuş ve Rotterdam kriterlerine göre ek olarak 2 PKOS vakası daha tespit etmiştir.

Sonuç: 3D US daha büyük OV değerleri vermekte ve hafifçe daha fazla sayıda kadına PKOS tanısı konmasına yardımcı olmaktadır. PKOM tanısı ile ilgili endişeler esas olarak antral folikül sayımına odaklansa da, bu bulgular sendromun tanısında OV'nin güvenilirliğini de vurgulamaktadır. Bu bulgular ovaryan ölçümlerin değerlendirilmesi ve tanısal kriterlerin

Cite this article as: Yapıcı Coşkun Z, Bozdağ G, Yıldız BO, Esinler İ, Yaralı H. Comparison of two and three-dimensional ultrasound in the diagnosis of polycystic ovary morphology: A cross sectional study in unselected women. Gazi Med J. 2026;37(2):177-182

Address for Correspondence/Yazışma Adresi: Zuhal Yapıcı Coşkun, Department of Obstetrics and Gynecology, Ankara Bilkent City Hospital, Ankara, Türkiye
E-mail / E-posta: zuhalyapici@yahoo.com
ORCID ID: orcid.org/0000-0002-9338-4669

Received/Geliş Tarihi: 02.10.2025
Accepted/Kabul Tarihi: 17.02.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

findings suggest that the imaging modality should be considered when interpreting ovarian measurements and applying diagnostic criteria.

Keywords: Ultrasonography, polycystic ovary syndrome, antral follicle count, three dimensional ultrasonography

ÖZ

uygulanmasında görüntüleme tekniği ve modalitesinin göz önünde bulundurulmasını önermektedir.

Anahtar Sözcükler: Ultrason, polikistik over sendromu, antral folikül sayımı, üç boyutlu ultrasonografi

INTRODUCTION

The role of assessment of ovarian morphology for the diagnostic value in polycystic ovary syndrome (PCOS) has been a debate (1). Following the consensus meeting held by the National Institutes of Health (NIH), it has been stated that clinical and/or biochemical hyperandrogenism accompanied by chronic anovulation are required for the diagnosis (2). PCOS Consensus Workshop Group sponsored by European Society of Human Reproduction and Embryology /American Society for Reproductive Medicine (ESHRE/ASRM) suggested that polycystic ovary morphology (PCOM) should be one of the three criteria for the confirmation of PCOS diagnosis (3). Alternatively, Androgen Excess-PCOS Society (AE–PCOS Society) recommended that PCOM would not have diagnostic value for PCOS without any finding of hyperandrogenism accompanying it (4).

As of today, not only the importance but also the definition of PCOM remains controversial (1). Initially, it was reported that antral follicle count (AFC) of more than or equal to 12 or ovarian volume (OV) of ≥10 cm^3 is essential for PCOM in Rotterdam (3). However, a revision of the morphological criteria for PCOM is required after this initial definition. The first reason is that PCOM is observed among otherwise healthy women whose AFCs overlap with those of women with PCOS. Therefore, it might be sensible to implement a higher cut-off, primarily for AFC, which would classify fewer otherwise healthy women as PCOM. On the other hand, the metabolic/ovarian states of women with PCOM-only are unclear, and it may not be straightforward to determine their health risk, if any, when they are classified as "normal". Secondly, due to technological improvement of ultrasound (US) evaluation and its resolution, we might be assigning more ovaries as "polycystic" than early ages. As noted previously (5), follicle enumeration with advanced US and grid-based methods necessitates a higher threshold across the ovary. And lastly, data about the validity of three dimensional (3D) features of US while counting antral follicles and OV in the diagnosis of PCOM is not clear. Although 3D US is increasingly used in gynecology, its concordance with two dimensional (2D) US for PCOM warrants investigation and clarification.

Our study aims to compare the visual properties of the ovaries using 2D and 3D US in unselected women of reproductive age, and to evaluate whether 3D changes the likelihood of a PCOS diagnosis based on either AFC or OV.

MATERIALS AND METHODS

Ethical Approval

This study was approved by the Institutional Ethics Review Board of Hacettepe University Faculty of Medicine on 19 November 2009 (approval number: 2545, date: 19.11.2009) and conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants and/or their legal guardians.

Participants

This study was conducted on an unselected population of women, evaluating them with both 2D and 3D US at the Institute of Mineral Research and Exploration (6). Unselected women describes volunteers from the general women population who responded to an open invitation. No inclusion criteria were based on clinical symptoms or suspicion of PCOS. All participants provided informed consent before participating in the study.

The female subjects were aged 18-45 year. Menopausal status, a history of hysterectomy or bilateral oophorectomy, pregnancy, and use of oral contraceptive pills for any reason were exclusion criteria (n = 43). A total of 115 women undergoing both 3D and 2D US are enrolled in the current study.

Study Protocol

For initial evaluation, a standardized medical form was used to obtain interview-based information on women's age, obstetric history, medical conditions, medications, menstrual regularity, gynecological and family history. Menstrual cycles with an interval ≥35 or ≤23 days were supposed to define ovulatory dysfunction.

The amount and distribution of terminal hair on the designated body areas were evaluated using a modified Ferriman Gallwey (mFG) scoring system, including 9 areas of the body, as described previously (7,8), to describe hirsutism. In case of mFG score being ≥6 participant was accepted to have hirsutism regardless of alopecia or acne existance. Biochemical hyperandrogenism was evaluated by analysing blood samples an overnight fast on the second to fifth day of menstruation between 8:00 and 10:30 AM. For participants using oral contraceptives, samples were collected during the interval. To define biochemical hyperandrogenism (hyperandrogenemia) in this study, otherwise healthy, non-hirsute women without PCOM and with regular menstrual cycles were considered the reference group. This group corresponds to 216 of the 392 women. Biochemical hyperandrogenemia is defined as increased levels of at least one of the androgens over 95th percentile such as total testosterone (tT), androstenedione, (A4), dehydroepiandrosterone sulfate and/or free androgen index (FAI).

If any of the androgen levels exceeds 95th percentile of healthy, non-hirsute women having regular menstrual cycles with no PCOM features (n = 216), then the patient assigned as having hyperandrogenemia.

Hormonal and Biochemical Analyses

After blood sampling, the specimen was transferred to a central laboratory by 11:00 AM. Following 20 minutes of centrifugation, specimen has been stored in polypropylene tubes at -70 °C until

final analysis. The hormonal analyses included TSH, prolactin, 17-OH progesterone, sex hormone binding globulin (SHBG), and previously defined androgens. The FAI was calculated using tT and SHBG levels as follows: (FAI = tT x 100/SHBG).

Ultrasonography

All examinations with 2 and 3D US were performed on second to seventh day of participant's menstrual bleeding with Voluson *e* (GE Healthcare, İstanbul, Türkiye) were performed by a single physician. Based on marital status and patient preference, either an abdominal (2–7 MHz) or a transvaginal (5–9 MHz) probe was used.

PCOM was defined as an AFC ≥12 follicles measuring 2–9 mm and/or an OV ≥10 cm^3 in either ovary. If one ovary could not be assessed, PCOM classification was based on the measurable ovary. When a persistent cyst or dominant follicle prevented accurate volume measurement on one ovary, AFC from both ovaries and OV from the contralateral ovary were used to determine PCOM status.

AFC and OV were determined using both US modalities.Under 2D US, antral follicles that have a diameter between 2 and 9 mm were counted in the transverse section at each site. OV was calculated from three diameters –anterior-posterior (a), maximum longitudinal (b), and transverse (c) –using the formula (a x b x c x 0.5). Under 3D US, the antral follicles were counted at the same time. The OV was processed using the Virtual Organ Computer-aided Analysis imaging program, employing Plane A and 60-degree rotational steps.

Definiton of PCOS

Depending on NIH criteria, PCOS had been defined as biochemical and/or clinical hyperandrogenism accompanied by ovulatory dysfunction as recommended (2). Rotterdam criteria revised the definition of PCOS by the presence of at least two of the findings/symptoms: 1) clinical and/or biochemical hyperandrogenism, 2) ovulatory dysfunction and, 3) PCOM (9). Regarding the AE-PCOS Society criteria, PCOS was diagnosed as biochemical and/or clinical hyperandrogenism associated with ovulatory dysfunction or PCOM (4,10). For the definitions of PCOS, only 2D evaluations were considered.

Statistical Analysis

The independent numerical parameters were analyzed with paired-samples t-tests. Reliability and consistency between parameters were assessed using Spearman correlation analysis. The term "mean bias" is defined as the average difference between the true OV and the estimated volume obtained using either 2D or 3D US. Bland–Altman plot of differences was drawn as defined previously (11). Parameters were noted as mean ± SD, unless stated otherwise. The SPSS 13.0 package (SPSS Inc., Chicago, IL) was used for statistical analysis. The figures of agreement were generated with GraphPad Prism 6.0 (trial version).

RESULTS

In this study, we compared 2D and 3D US measurements of AFC and OV in 115 women to determine whether the imaging method affects the classification of PCOM and PCOS. As depicted in the Figure 1, when 392 women were referred (80.2% of whole population), the prevalence of PCOS due to subsets of NIH (6.1%), Rotterdam ESHRE/ASRM (19.9%) and AE-PCOS Society (15.3%) criteria were calculated (6). According to 2D US, the rate of PCOM in women was 36.5% (143/392). Among 143 women with PCOM, 59 had bilateral and 84 had unilateral PCOM. Notably, PCOM was diagnosed in 95 (66.4%) based on AFC, 6 (4.2%) based on OV and 42 (29.4%) women based on the presence of both.

Of the 392 participants, we evaluated 115 with both 2D and 3D sonographic modalities in at least one ovary (Figure 1). Whereas the mean AFC with these modalities were 10.3 ± 4.8 vs. 11.2 ± 5.5 ($p = 0.61$), the OV estimated were 7.5 ± 3.0 vs. 9.6 ± 4.0 ($p < 0.001$) mL respectively.

Sixty-six ovaries were classified as PCOM when US was performed using 2D imaging (Table 1). Sixty-five (93.4%) of them were confirmed as PCOM under 3D US. However, of the 49 ovaries who were not noted to be PCOM on 2D, 34 were diagnosed as PCOM with 3D US. That revision revealed 2 more cases of PCOS according to the Rotterdam criteria, when OV and hyperandrogenism were considered. However, this numerical increase is small, has not been statistically tested, and is unlikely to represent a statistically significant difference. The findings suggest that 3D US may shift the classification in borderline cases, but may not substantially alter PCOS prevalence in a population of this size.

The coefficients between these sonographic modalities, analysed for AFC and OV calculations, were 0.783 and 0.626, respectively. When the limit of agreement was tested, the mean bias between 2D and 3D for the estimation of AFC was –0.17 (-8.45 to 8.11). The respective figure for OV was -1.99 (-7.72 to 3.75) mL. Therefore, 3D yielded values that were 22.5% larger than the OV estimated by 2D (Figures 2 and 3).

DISCUSSION

Our main finding is that 3D US yields consistently larger ovarian-volume measurements than 2D US, resulting in more women being classified with PCOM and, in some cases, PCOS. For the AFC, both the agreement and the consistency of 2D and 3D US were superior to those of OV. However, despite good agreement in OV between the two methods, 3D estimated OV was 22.5% larger, which might be clinically important in the diagnosis of PCOM. That finding once again supports the type of statistical analysis used in comparing modalities. This highlights that OV may be less stable and more sensitive to technical differences than AFC; therefore, diagnostic thresholds may need to be adjusted for different US technologies. The major strength of this study is that it reports US findings using sonographic modalities of different dimensionalities in a subset of a large unselected population.

In the available literature, the agreement among sonographic modalities of different dimensionalities regarding AFC and OV is inconclusive. According to our data, in spite of good agreement between 2D and 3D US regarding AFC and OV across the whole population, 3D US detected a slightly higher number of cases among patients with PCOS according to the Rotterdam criteria. The technical aspects of transabdominal US with poor spatial resolution and low-frequency probes limit clarity compared to transvaginal probe that provides 3D evaluation more clearly (12).

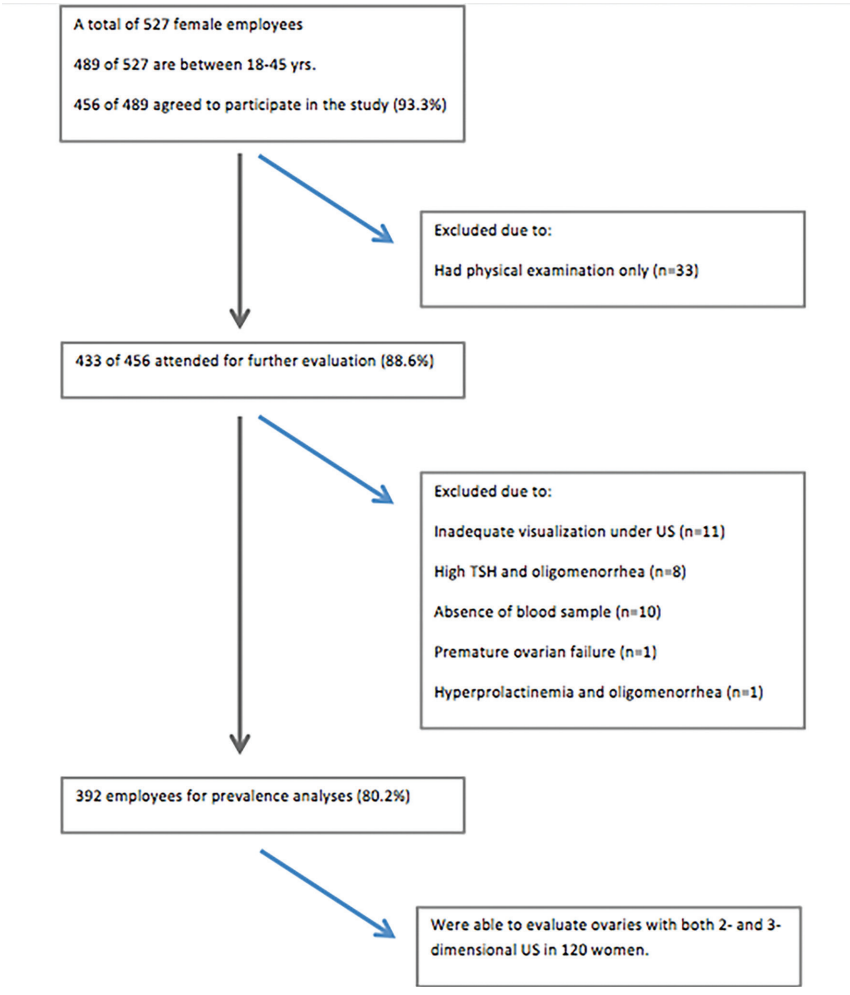

Figure 1. Study population.

TSH: Thyroid-stimulating hormone, US: Ultrasound.

Table 1. The relation of 2D and 3D US regarding polycystic ovarian morphology.

	2D, PCOM (+)	2D, PCOM (-)
3D, PCOM (+)	65	34
3D, PCOM (-)	5	16

PCOM: Polycystic ovarian morphology, 2D: Two dimensional, 3D: Three dimensional, US: Ultrasound.

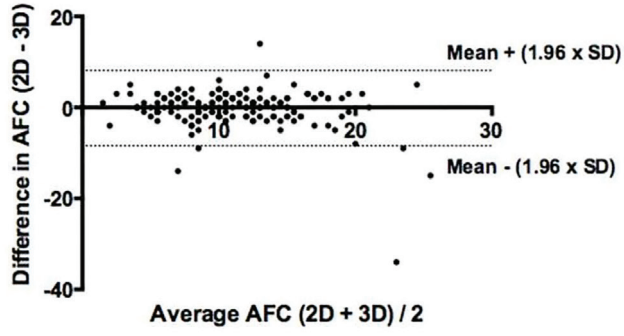

Figure 2. Bland-Altman plot of 2D and 3D US for AFC.

AFC: Antral follicle count, 2D: Two dimensional, 3D: Three dimensional, US: Ultrasound, SD: Standard deviation.

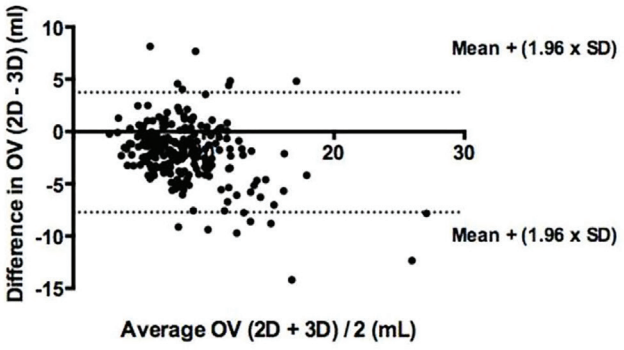

Figure 3. Bland-Altman plot of 2D and 3D US for OV.

SD: Standard deviation, 2D: Two dimensional, 3D: Three dimensional, US: Ultrasound, OV: Ovarian volume, SD: Standard deviation.

Although a significant correlation is observed between 2D and 3D US in our study similar to previous literature (13), the level of agreement might present some diversions. According to Scheffer et al (14), the coverage intervals for the difference between two methods were -5.3 to 8.3 follicles which indicate a moderate agreement. Of note, the agreement gets worse as the number of antral follicles increase (14). In another study that assesses the reliability of measurements made by manual 3D or with sono-automatic volume count software, the respective mean ± standard deviation of AFCs were 6.5 ± 4.8 vs. 19.4 ± 10.9 (15). We observed better agreement between the two methods for AFC, but agreement worsened for OV estimation. Nevertheless, previously we had reported that whereas 2D US brought 18% larger, 3D US revealed 11% smaller values compared to the absolute OV that was calculated according to Archimedes' principles following oophorectomy in patients undergoing surgery (16). Therefore, one may hypothesize that different thresholds might be required depending on US specifications and the availability of software to define PCOM, which would not be reasonable to apply in clinical practice.

Since defining universally accepted morphological criteria for the appearance of polycystic ovaries is unlikely, two main strategies might be pursued. Those are revising the threshold levels with parallel to the improvement in imaging or secondly, replacing it with a biochemical marker that is less physician and cycle dependent. However, the former strategy requires revisiting the thresholds regularly as technology improves (1). Nevertheless, a recent task force by AE-PCOS Society (17) recommended using higher threshold levels of AFC for the definition of PCOM, particularly when using newer technology (i.e. transducer frequency ≥8 MHz) which provides better resolution for ovarian follicles. Therefore, the technological evolution of US does not allow the use of the same threshold for AFC or OV over time. In this context the diagnostic tresholds for US evaluation of follicle number per ovary has been revised and increased in the International Evidence-based Guideline for the assessment and management of PCOS 2023 (18).

A recent study has shown that artificial intelligence, using the backpropagation algorithm, accurately identifies the three-dimensional ovarian structure and measures both OV and the AFC. US parameters, in addition to endocrine and metabolic parameters, represent objective diagnostic tools and have clinical importance. Comprehensive studies on this issue with larger participant population needs to be performed to better diagnose and evaluate PCOS and direct clinical management (19).

Study Limitations

As highlighted earlier, the sampling methodology represents one of the principal limitations of this study. Although women working at the institution participated in the study with a high response rate, potential selection bias due to undetermined differences between the study sample and the background community cannot be excluded. The second limitation is the low proportion of women who had at least one ovary eligible for evaluation by both modalities. The unusually high rate of missing data may be attributable to certain issues. Initially, all patients declined the endocavitary probe because of virginal status and discomfort with the vaginal probe (n = 174). Nevertheless, the technical aspects of transabdominal US with poor spatial resolution and low-frequency probes limit clarity (12). Potentially, image quality may be reduced secondary to central obesity, which is commonly seen in women with PCOS (12). Among the whole group, the percentages of overweight women [body mass index (BMI): 25.0–29.9 kg/m^2] and obese women (BMI ≥30 kg/m^2) were 24.0 and 10.2%, respectively. We did not include subjects in the final analysis unless one ovary was properly visualized. The physician's meticulous policy might also have caused the lower number of cases included. The effect of these missing data on the outcome and conclusions of the study is unknown, but the interpretation of the results using the Bland-Altman plot might be more valuable than comparing the mean values. These factors should be considered when interpreting the results.

CONCLUSION

We suggest that 3D US estimates larger OV when compared with 2D, resulting in a slightly higher number of women being diagnosed with PCOS. A difference of 22.5% might be clinically important for women whose values are close to current cut-offs for the definition of PCOM and, in turn, for assigning them to the syndrome. Although previous debates have focused on detecting antral follicles, those findings also underscore the reliability of OV as a diagnostic parameter for PCOS. These differences suggest that diagnostic thresholds for PCOM may need to be adjusted for newer imaging technologies**.** In the following studies, the concordance and consistency of 2D and 3D sonographic modalities for sonographic parameters such as OV and AFC needs to be investigated with larger study population.

Ethics

Ethics Committee Approval: The study was approved by the Hacettepe University Faculty of Medicine Institutional Review Board on 19 November 2009 (approval number: 2545, date: 19.11.2009).

Informed Consent: Prior to inclusion in the study, informed consent was obtained from the participants.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Concept: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Design: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Data Collection or Processing: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Analysis or Interpretation: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Literature Search: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Writing: Z.Y.C., G.B., B.O.Y., İ.E., H.Y.

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

Financial Disclosure: This study was partially sponsored by Merck Serono.

REFERENCES

1. Dewailly D, Gronier H, Poncelet E, Robin G, Leroy M, Pigny P, et al. Diagnosis of polycystic ovary syndrome (PCOS): revisiting the threshold values of follicle count on ultrasound and of the serum AMH level for the definition of polycystic ovaries. Hum Reprod. 2011; 26: 3123-6.
2. Zawadzki JK, Dunaif A. Diagnostic criteria for polycystic ovary syndrome. In: Dunaif A, Givens JR, Haseltine F, Merriam GR, editors. Polycystic ovary syndrome. Boston: Blackwell Scientific Publications; 1992. p. 377-84.
3. Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil Steril. 2004; 81: 19-6.
4. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, et al. Positions statement: criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an Androgen Excess Society guideline. J Clin Endocrinol Metab. 2006; 91: 4237-28.
5. Lujan ME, Jarrett BY, Brooks ED, Reines JK, Peppin AK, Muhn N, et al. Updated ultrasound criteria for polycystic ovary syndrome: reliable thresholds for elevated follicle population and ovarian volume. Hum Reprod. 2013; 28: 1361-7.
6. Yildiz BO, Bozdag G, Yapici Z, Esinler I, Yarali H. Prevalence, phenotype and cardiometabolic risk of polycystic ovary syndrome under different diagnostic criteria. Hum Reprod. 2012; 27: 3067- 6.
7. Ferriman D, Gallwey JD. Clinical assessment of body hair growth in women. J Clin Endocrinol Metab. 1961; 21: 1440-7.
8. Yildiz BO, Bolour S, Woods K, Moore A, Azziz R. Visually scoring hirsutism. Hum Reprod Update. 2010; 16: 51-64.
9. Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). Hum Reprod. 2004; 19: 41-7.
10. Azziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The prevalence and features of the polycystic ovary syndrome in an unselected population. J Clin Endocrinol Metab. 2004; 89: 2745-9.
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1: 307-10.
12. Yoo RY, Sirlin CB, Gottschalk M, Chang RJ. Ovarian imaging by magnetic resonance in obese adolescent girls with polycystic ovary syndrome: a pilot study. Fertil Steril. 2005; 84: 985-10.
13. Nardo LG, Buckett WM, Khullar V. Determination of the best-fitting ultrasound formulaic method for ovarian volume measurement in women with polycystic ovary syndrome. Fertil Steril. 2003; 79: 632-1.
14. Scheffer GJ, Broekmans FJ, Bancsi LF, Habbema JD, Looman CW, Te Velde ER. Quantitative transvaginal two- and three- dimensional sonography of the ovaries: reproducibility of antral follicle counts. Ultrasound Obstet Gynecol. 2002; 20: 270-5.
15. Deb S, Jayaprakasan K, Campbell BK, Clewes JS, Johnson IR, Raine-Fenning NJ. Intraobserver and interobserver reliability of automated antral follicle counts made using three-dimensional ultrasound and SonoAVC. Ultrasound Obstet Gynecol. 2009; 33: 477-83.
16. Bozdag G, Salman MC, Mumusoglu S, Yapici Z, Gunalp S. Is ovarian volume estimation reliable when compared with true volume? Am J Obstet Gynecol. 2012; 206: 44.e1-4.
17. Dewailly D, Lujan ME, Carmina E, Cedars MI, Laven J, Norman RJ, et al. Definition and significance of polycystic ovarian morphology: a task force report from the Androgen Excess and Polycystic Ovary Syndrome Society. Hum Reprod Update. 2014; 20: 334-52.
18. Teede HJ, Tay CT, Laven J, Dokras A, Moran LJ, Piltonen TT, et al. Recommendations from the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome†. Hum Reprod. 2023; 38: 1655-79.
19. Wei L, Wu F, Zhang J, Li J, Yang D, Wen G. Evaluation of endocrine and metabolic changes in polycystic ovary syndrome by ultrasonic ımaging features under an ıntelligent algorithm. Comput Math Methods Med. 2022; 2022: 1411943.

Gazi Med J 2026;37(2):183-188

DOI: http://dx.doi.org/10.12996/gmj.2026.4460

Impact of Baseline Characteristics and Parental Risk Factors on CHDs: A Comparative Analysis Study

Temel Özelliklerin ve Ebeveyn Risk Faktörlerinin Doğuştan Kalp Hastalıkları Üzerindeki Etkisi

Sana Ashiq[1,2], Syed Najam Hyder[3], Muhammad Farooq Sabar[4]

[1]Department of Molecular Genetics, Centre for Applied Molecular Biology, University of the Punjab, Lahore, Pakistan
[2]Department of Medical Laboratory Technology, Shalamar School of Allied Health Sciences, Affiliated with University of Health Sciences, Lahore, Pakistan
[3]Department of Paediatric Cardiology, University of Child Health Sciences, The Children's Hospital, Lahore, Pakistan
[4]Department of School of Biochemistry and Biotechnology, University of the Punjab, Lahore, Pakistan

ABSTRACT

Objective: To identify the potential role of the patient's baseline characteristics and assess the association of congenital heart defects (CHDs) with parental risk factors, particularly maternal chronic disease and socioeconomic status.

Methods: This case-control study included 376 subjects. Extensive patient histories, including the subjects' anthropometric parameters and paternal risk factors, were collected from multiple hospitals in Lahore between March 2021 and April 2022. Children's physical parameters, including body mass index, were measured according to the Centre's for Disease Control and Prevention guidelines. Statistical analyses were conducted using R-manager and GraphPad Prism.

Results: In the current study, 65.8% of CHD subjects and 64.9% of healthy subjects were male. The comparative assessment of the patient's anthropometric parameters suggested no significant association with the heart defect. However, compared to healthy subjects, CHD patients were significantly underweight ($p < 0.0001$). However, the difference was not significant for the comparison between cyanotic and acyanotic CHD groups. The assessment of maternal risk factors showed significant associations for maternal hypertension [3.09 95% confidence interval (CI): 1.64-5.79] and maternal diabetes [2.92 (95% CI: 1.24-6.88)]. In addition, the impact of parental socioeconomic status was substantial: 25.7% and 46.6% of patients were from poor and middle-income families, respectively.

ÖZ

Amaç: Hastanın temel özelliklerinin potansiyel rolünü belirlemek ve doğuştan kalp kusurlarının (DKK) ebeveyn risk faktörleriyle, özellikle annenin kronik hastalığı ve sosyoekonomik durumuyla ilişkisini değerlendirmek.

Yöntemler: Bu olgu-kontrol çalışmasına 376 kişi dahil edildi. Mart 2021 ile Nisan 2022 tarihleri arasında Lahor'daki birçok hastaneden, deneklerin antropometrik parametreleri ve baba risk faktörleri de dahil olmak üzere kapsamlı hasta öyküleri toplandı. Çocukların vücut kitle indeksi de dahil olmak üzere fiziksel parametreleri, Hastalık Kontrol ve Önleme Merkezi kılavuzlarına göre ölçüldü. İstatistiksel analizler R-manager ve GraphPad Prism kullanılarak yapıldı.

Bulgular: Mevcut çalışmada, DKK'lı deneklerin %65,8'i ve sağlıklı deneklerin %64,9'u erkekti. Hastaların antropometrik parametrelerinin karşılaştırmalı değerlendirmesi, kalp kusuru ile anlamlı bir ilişki göstermedi. Bununla birlikte, sağlıklı deneklere kıyasla, DKK hastaları anlamlı derecede düşük kiloluydu ($p < 0,0001$). Ancak, siyanotik ve asiyanotik DKK grupları arasındaki karşılaştırmada fark anlamlı değildi. Anne risk faktörlerinin değerlendirilmesi, maternal hipertansiyon için anlamlı ilişkiler gösterdi [3,09 (%95 güven aralığı (GA): 1,64-5,79)] ve maternal diyabet için anlamlı ilişkiler gösterdi [2,92 (%95 GA: 1,24-6,88)]. Ayrıca, ebeveynlerin sosyoekonomik durumunun etkisi önemliydi: Hastaların %25,7'si ve %46,6'sı sırasıyla yoksul ve orta gelirli ailelerden geliyordu.

Cite this article as: Ashiq S, Hyder SN, Sabar MF, et al. Impact of baseline characteristics and parental risk factors on CHDs: a comparative analysis study. Gazi Med J. 2026;37(2):183-188

Address for Correspondence/Yazışma Adresi: Sana Ashiq, Department of Molecular Genetics, Centre for Applied Molecular Biology, University of the Punjab; Department of Medical Laboratory Technology, Shalamar School of Allied Health Sciences, Affiliated with University of Health Sciences, Lahore, Pakistan
E-mail / E-posta: sanaashiq72@gmail.com
ORCID ID: orcid.org/0000-0003-0418-4022

Received/Geliş Tarihi: 20.05.2025
Accepted/Kabul Tarihi: 23.02.2026
Epub: 12.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

Conclusion: CHD in children was significantly associated with patients' health status, including maternal hypertension and diabetes. However, this relationship was not found between cyanotic and acyanotic patients. In addition, parental socioeconomic status patients' poses a significant burden on patients' families and the healthcare system.

Keywords: CHD, BMI, hypertension, socioeconomic status, congenital heart defect

ÖZ

Sonuç: Çocuklarda DKK, maternal hipertansiyon ve diyabet de dahil olmak üzere hastaların sağlık durumuyla anlamlı derecede ilişkiliydi. Bununla birlikte, bu ilişki siyanotik ve asiyanotik hastalar arasında bulunmadı. Ayrıca, ebeveynlerin sosyoekonomik durumu, hastaların aileleri ve sağlık sistemi üzerinde önemli bir yük oluşturmaktadır.

Anahtar Sözcükler: DKK, BMI, hipertansiyon, sosyoekonomik durum, konjenital kalp kusuru

INTRODUCTION

Globally, congenital heart defects (CHDs) remained the top reason of infant morbidity and mortality. Thus, included as one of the major agendas in the 2015 sustainable development goals of the United States (1). Worldwide, 10% of all births are affected by this disease; however, due to recent advancements in pediatric cardiology treatment, overall mortality has declined. While in comparison to high-income countries a higher mortality rate was observed in lower-middle-income countries and lower-income countries with an average of 1.2 deaths per 100,000 cases and 4.9 deaths per 100,000 cases respectively (2). A meta-analysis of 260 studies suggested that overall, after every five years the prevalence of congenital heart malformation increased by 10% due to milder lesions with the highest prevalence in Asian regions as compared to Africa (3). It is a complex birth defect that occurs during cardiogenesis and may involve a complex interaction between genetic and environmental risk factors, particularly maternal factors (4). Depending on the clinical severity of the disease, CHDs can be further classified as mild, moderate, and severe. Moreover, the clinical presentation of patients is further divided into two major categories cyanotic or a-cyanotic (5-7). Although several studies have been conducted to analyze the underlying cause of the disease, the exact mechanism of pathogenesis remains unclear. One poorly understood mechanism in pathophysiology is the role played by modifiable risk factors, including maternal health (particularly during pregnancy), use of medications during pregnancy, a history of stillbirths, smoking, and diabetes mellitus. In addition, these risk factors may vary across different populations and cultures (8,9). Over the past few decades, several studies identified genetic variants and chromosomal abnormalities in syndromic CHD but the etiology of complex non-syndromic CHD needs further studies that can provide us better insights into disease pathogenesis and may provide us a new direction for future prevention of disease (10). To date, the majority of research studies have focused on the genetics of CHD compared with other extrinsic factors, while few extrinsic factors have been studied. This study comprehensively analyzed all the crucial paternal and anthropometric parameters in non-syndromic CHD, which may help us strengthen management guidelines and improve genetic counseling, thereby reducing future disease risk.

MATERIALS AND METHODS

This study was designed as a case-control study to compare the baseline characteristics and maternal risk factors. Ethical approval was obtained from the Punjab University Ethical Committee, Faculty of Economics and Management Sciences, University of the Punjab (approval number: 84/DFEMS, date: 05.04.2021) and Institutional Review Board of The Children's Hospital & The Institute of Child Health, Lahore (approval number: 2021-282-CHICH, date: 18.05.2021). This study includes paediatric cardiologist-confirmed cases of both cyanotic and acyanotic CHD. After obtaining informed consent, all critical information was collected from the study subjects. The information collected from participants was divided into two phases: first, the patient's baseline parameters were studied; second, parental risk-factor analysis was performed and is described in the detailed sections below.

Patient's and Healthy Subjects Characteristics

The baseline parameters from subjects with cyanotic or acyanotic CHDs and control subjects were recorded for comparative evaluation. The following information was collected: age, gender, weight in kg, height in cm, and body mass index (BMI) in kg/m^2. The BMI results were further used to classify the study subjects as underweight, healthy weight, overweight, and obese. BMI was calculated according to the Centers for Disease Control and Prevention guidelines. The following cut-off points were used for classification: less than the 5^{th} percentile for underweight, 5^{th} percentile to up to the 85^{th} percentile for healthy subjects, 85^{th} to less than the 95^{th} percentile for overweight, and equal to, or greater than the 95^{th} percentile for obese subjects (11).

Parental Risk Factors

The paternal risk-factor analysis was divided into four major sections: demographics, family history, chronic disease evaluation, and maternal pregnancy complications. The demographic data further include the parental age, and socioeconomic status evaluation according to the Kuppuswamy scale, and data were divided into upper-class, middle-class, poor class, and very-poor class (12,13). Maternal risk factors recorded were smoking, of medicines, a history of chronic disease or complicated pregnancy, and the number of children aborted.

Statistical Analysis

Categorical data were represented as percentages or frequencies, and continuous data as means with standard deviation, respectively. The t-test and chi-square test were used to compare continuous and categorical variables, respectively. The associations were further expressed as odds ratios (ORs) and 95% confidence intervals (CIs). A p-value <0.05 was considered statistically significant. The statistical analyses were performed using R-Manager, the Statistical Package for the Social Sciences 22.0, and GraphPad Prism version 5.

RESULTS

A total of 376 subjects were recruited, including 225 cases and 151 controls. Among CHDs, 101 were cyanotic and 124 were acyanotic. The cases had 148 males and 77 females, while the controls had 98 males and 53 females (p = 0.95). The comparative assessment of baseline characteristics between patients with CHDs and healthy subjects revealed no significant differences except for BMI (p < 0.0001, Table 1).

The comparison of health status suggested a significant difference between the two categories: the percentages of healthy, underweight, overweight, and obese individuals were 92.54%, 0%, 4.48%, and 2.98% in controls, and 58.59%, 37.37%, 4.04%, and 0% in cases (Figures 1 and 2).

There was no statistically significant difference in anthropometric parameters between cyanotic and acyanotic CHDs (p > 0.05, Table 1). The percentages of healthy, underweight, overweight, and obese in cyanotic and a-cyanotic groups were 60%, 37.78%, 2.22%, and 0%, and 57.41%, 37.04%, 5.55%, and 0%, respectively (Figure 3). The impact of paternal socioeconomic status was significant in the Pakistani population, as the majority of patients (46.67%) belonged to the middle class and 25.78% to the poor class. Moreover, high-cost treatment strategies were a key reason for the burden on patients' families and the healthcare system. However, in 79.56% of cases, clinical outcomes improved after open-heart, closed-heart, or interventional treatments.

The maternal age was 27.64 ± 5.64 and 26.88 ± 4.51 (p = 0.16), and the paternal age was 30.79 ± 5.35 and 30.37 ± 4.43 (p = 0.42) in congenital heart disease and healthy subjects, respectively. Similarly, there was no statistically significant difference between cyanotic and a-cyanotic CHD for maternal and paternal age (p = 0.51, p = 0.41) (Table 2).

Assessment of underlying maternal chronic disease conditions suggested a significant association between maternal hypertension and CHDs when patients were compared with controls (OR: 3.09, 95% CI: 1.64-5.79, p = 0.0003). Similarly, maternal diabetes was significantly associated with CHD in children (OR: 2.92, CI: 1.24-6.88, p = 0.01). Child mortality due to maternal pregnancy complications and the use of medicines was non-significant (p = 0.78 and 0.58), respectively. The maternal and paternal histories also showed no significant association with congenital heart disease (p > 0.05) in the Pakistani population (Tables 3 and 4).

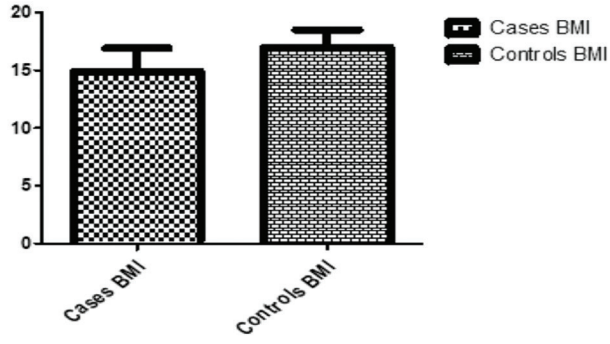

Figure 1. Evaluation of BMI between CHD and control subjects.
CHDs: Congenital heart defects, BMI: Body mass index.

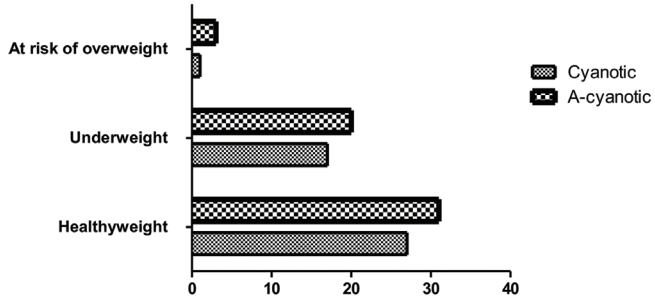

Figure 2. Comparison of health status between cases and controls.

Table 1. Role of baseline characteristics in congenital heart defect and comparative assessment between cyanotic and a-cyanotic CHDs.

Characteristics	CHDs	Controls	p-value	Cyanotic CHDs	A-cyanotic CHDs	p-value
Gender						
Male	148 (65.8%)	98 (64.9%)	0.95	63 (62.4%)	85 (68.5%)	0.33
Female	77 (34.2%)	53 (35.1%)		38 (37.6%)	39 (31.5%)	
Age:						
<1 year (months)	5.27 ± 2.94	5.72 ± 2.66	0.28	5.24 ± 3.25	5.28 ± 2.80	0.95
≥1 year (years)	4.41 ± 3.68	4.06 ± 3.58	0.47	4.24 ± 3.85	4.58 ± 3.51	0.58
Weight (kg)	8.73 ± 5.60	9.31 ± 8.01	0.40	8.89 ± 5.64	8.60 ± 5.58	0.70
<2 year weight	4.92 ± 1.70	4.68 ± 1.25	0.26	5.15 ± 1.81	4.75 ± 1.60	0.19
≥2 year weight	13.58 ± 5.06	15.12 ± 9.07	0.16	13.5 ± 5.33	13.6 ± 4.89	0.92
Height (cm)	72.83 ± 23.39	70.33 ± 24.73	0.32	73.07 ± 23.47	72.21 ± 23.97	0.78
<2 year height	55.83 ± 7.72	54.33 ± 7.71	0.16	56.27 ± 8.05	55.47 ± 7.48	0.56
≥2 year height	94.46 ± 18.18	90.39 ± 24.10	0.21	93.98 ± 19.11	93.91 ± 20.26	0.98
BMI (kg/m²)	14.89 ± 2.03	16.97 ± 1.51	<0.0001*	14.88 ± 1.97	14.90 ± 2.09	0.96
≥2 year (years)						

*: p < 0.05.
CHDs: Congenital heart defects, BMI: Body mass index, Kg: Kilogram, cm: Centimeter, Kg/m²: Kilogram/meter².

DISCUSSION

The current study was the first comprehensive report from the Pakistani cohort that comparatively analyzed the role of patients' baseline characteristics and parental risk factors. The results of this study suggested an association between underlying maternal chronic disease conditions, including hypertension and diabetes, and congenital heart disease in children. In addition, the findings of this study indicated that CHD patients had a compromised health status compared with healthy subjects. Furthermore, it posed a substantial burden on patients' families and the healthcare system, as the majority of patients in Pakistan are from middle- or low-income families. The expensive surgical treatments are considered a major challenge for cardiac surgeons, patients' families, and Pakistan's healthcare system.

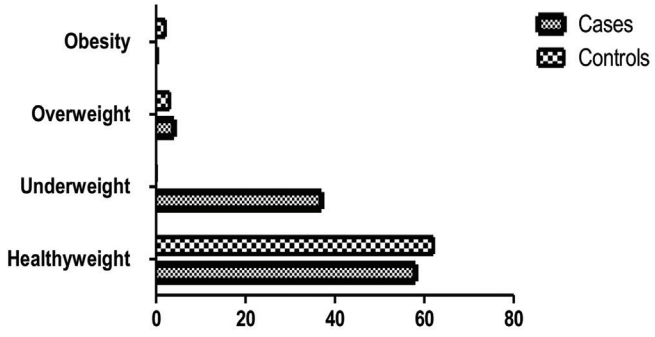

Figure 3. Comparative assessment of health status between cyanotic and a-cyanotic CHD.

CHD: Congenital heart defect.

The low BMI of patients indicates poorer nutritional status compared with healthy subjects. Okoromah et al. (14) also reported severe malnutrition and underweight in congenital heart disease patients as compared to controls.

This study's findings were similar to results from another cohort that reported 21% of CHD patients were underweight ($p < 0.001$). While cyanotic vs. a-cyanotic CHD analysis showed that cyanotic were more underweight (15).

Xiang et al. (16) reported a 97% survival rate after surgeries of pediatric patients and suggested that middle and low-income families were at high risk of poor prognosis after cardiac surgery. The overall percentage of low and middle-income status patients was 69% (16). A population-based study from California showed a significant association between patient socioeconomic status and environmental triggers with high CHD incidence (17). Maternal occupation and socioeconomic disparities also showed a significant association ($p < 0.001$) with the disease in Iran (18). Paternal low socioeconomic status and remoteness of residence were associated with high patient mortality and adverse disease outcomes (19).

Table 2. Impact of maternal and parental age.

Parental age	CHDs	Controls	p-value	CHDs		p-value
				Cyanotic	A-cyanotic	
Maternal age	27.64 ± 5.64	26.88 ± 4.51	0.16	27.92 ± 5.57	27.43 ± 5.71	0.51
Paternal age	30.79 ± 5.35	30.37 ± 4.43	0.42	31.11 ± 5.26	30.53 ± 5.43	0.41

CHDs: Congenital heart defects.

Table 3. Comparative assessment of parental risk factors.

Risk factors	CHDs n (%)	Controls n (%)	OR (95% CI)	p-value
Maternal hypertension				
Yes	54 (24%)	14 (9.27%)	3.09 (1.64-5.79)	0.0003*
No	171 (76%)	137 (90.73%)		
Maternal diabetes				
Yes	28 (12.4%)	7 (4.6%)	2.92 (1.24-6.88)	0.01*
No	197 (87.6%)	144 (95.4%)		
Children deceased due to maternal pregnancy complications				
Yes	19 (8.44%)	14 (9.3%)	0.90 (0.43-1.86)	0.78
No	206 (91.6%)	137 (90.7%)		
Medicines used during pregnancy				
Yes	20 (8.89%)	16 (10.6%)	0.82 (0.41-1.64)	0.58
No	205 (91.11%)	135(89.4%)		
Paternal smoking history				
Yes	12 (5.33%)	11 (7.3%)	0.71 (0.30-1.67)	0.43
No	213 (94.67%)	140 (92.7%)		

*: $p < 0.05$.
CHDs: Congenital heart defects, CI: Confidence interval, OR: Odds ratio.

Table 4. Comparative study of parental risk factors.

Characteristics	CHDs (n = 225)	Controls (n = 151)	Cyanotic CHDs (n = 101)	A-cyanotic CHDs (n = 124)
Maternal hypertension	54 (24%)	14 (9.27%)	26 (25.7%)	28 (22.6%)
Maternal diabetes	28 (12.4%)	7 (4.6%)	14 (13.9%)	14 (11.3%)
Children deceased due to maternal pregnancy complications	19 (8.44%)	14 (9.3%)	11 (10.89%)	8 (6.45%)
Number of children aborted				
One abortion	10 (4.45%)	14 (9.3%)	8 (7.92%)	2 (1.61%)
Two abortions	3 (1.33%)	0 (0%)	0 (0%)	3 (2.42%)
Three abortions	1 (0.44%)	0 (0%)	1 (0.99%)	0 (0%)
No abortions	211 (93.78%)	137 (90.7%)	92 (91.09%)	119 (95.97%)
Medicines used during pregnancy	20 (8.89%)	16 (10.6%)	9 (8.91%)	11 (8.87%)
Maternal smoking history	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Paternal smoking history	12 (5.33%)	11 (7.3%)	4 (3.96%)	8 (6.45%)
Family history	25 (11.11%)	7 (4.64%)	13 (12.87%)	12 (9.68%)

CHDs: Congenital heart defects.

A cohort study reported no significant association between paternal age and CHDs in the Danish population (20). However, findings from the Indian cohort showed an association between maternal and paternal age and heart defects in children (21). This study's results were in accordance with the study of Taylor et al. (22) as they reported a 0.96 OR and 0.85-1.07, 95% CI for paternal smoking. Other case-control results suggested a relationship between maternal smoking and the high risk of CHD in children (23). Likewise, maternal hypertension ($p < 0.01$), maternal smoking ($p < 0.01$), and maternal diabetes ($p < 0.01$) were found as strongly associated risk factors for congenital heart malformation in children (24). Hypertensive disorder of pregnancy has been shown to be associated with a threefold increase in CHD (OR: 2.51, 95% CI: 2.38-2.64, $p \leq 0.001$), and the OR for maternal diabetes and CHD in children was 5.14 (95% CI: 5.04-5.23, $p \leq 0.001$) (25). The OR for maternal drug use was 2.68 ($p < 0.05$), the OR for family history was 4.14, and for maternal abortions was 1.12 (26). Maternal hypertension, diabetes, preeclampsia, and smoking during pregnancy showed a statistically significant association with CHDs in the pediatric population ($p < 0.001$) (27).

CONCLUSION

The findings of this study suggest a potential role for maternal risk factors, including hypertension, diabetes, and socioeconomic status. Furthermore, the analysis revealed that children with CHD were more likely to be underweight compared with healthy subjects. However, in the cyanotic group, anthropometric parameters, including BMI, were not statistically different from those in the acyanotic group. In addition, maternal and paternal ages did not differ between the CHDs and control groups, and between the cyanotic and a-cyanotic groups. Appropriate measures should be taken to screen for parental risk factors and family history, which may help us strengthen future CHD prevention guidelines. It is further recommended to conduct additional studies with larger cohorts across different subgroups to provide a more definitive conclusion.

Ethics

Ethics Committee Approval: This study was designed as a case-control study to compare the baseline characteristics and maternal risk factors. Ethical approval was obtained from the Punjab University Ethical Committee, Faculty of Economics and Management Sciences, University of the Punjab (approval number: 84/DFEMS, date: 05.04.2021) and Institutional Review Board of The Children's Hospital & The Institute of Child Health, Lahore (approval number: 2021-282-CHICH, date: 18.05.2021).

Informed Consent: The written informed consent was obtained from the patients.

Footnotes

Authorship Contributions

Concept: S.A., Design: S.A., S.N.H., M.F.S., Data Collection or Processing: S.A., S.N.H., M.F.S., Analysis or Interpretation: S.A., S.N.H., M.F.S., Literature Search: S.A., Writing: S.A., S.N.H., M.F.S.

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

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

Acknowledgements

We would like to acknowledge CAMB, University of the Punjab, for their kind support. We would also like to acknowledge hospital staff, patients, and their families for providing data.

REFERENCES

1. Su Z, Zou Z, Hay SI, Liu Y, Li S, Chen H, et al. Global, regional, and national time trends in mortality for congenital heart disease, 1990-2019: an age-period-cohort analysis for the Global Burden of Disease 2019 study. EClinicalMedicine. 2022; 43.
2. Varela-Chinchilla CD, Sánchez-Mejía DE, Trinidad-Calderon PA. Congenital heart disease: the state-of-the-art on its pharmacological therapeutics. J Cardiovasc Dev Dis. 2022; 9: 201.

3. Liu Y, Chen S, Zühlke L, Black GC, Choy MK, Li N, et al. Global birth prevalence of congenital heart defects 1970-2017: updated systematic review and meta-analysis of 260 studies. Int J Epidemiol. 2019; 48: 455-63.
4. Ashiq S, Ashiq K, Sabar MF. The role of NKX2-5 gene polymorphisms in congenital heart disease (CHD): a systematic review and meta-analysis. Egypt Heart J. 2021; 73: 72.
5. Jin X, Ni W, Wang G, Wu Q, Zhang J, Li G, et al. Incidence and risk factors of congenital heart disease in Qingdao: a prospective cohort study. BMC Public Health. 2021; 21: 1044.
6. Majiyagbe OO, Akinsete AM, Adeyemo TA, Salako AO, Ekure EN, Okoromah CA. Coagulation abnormalities in children with uncorrected congenital heart defects seen at a teaching hospital in a developing country. PLoS One. 2022; 17: e0263948.
7. Madsen NL, Marino BS, Woo JG, Thomsen RW, Videbæk J, Laursen HB, et al. Congenital heart disease with and without cyanotic potential and the long-term risk of diabetes mellitus: a population-based follow-up study. J Am Heart Assoc. 2016; 5: e003076.
8. Fung A, Manlhiot C, Naik S, Rosenberg H, Smythe J, Lougheed J, et al. Impact of prenatal risk factors on congenital heart disease in the current era. J Am Heart Assoc. 2013; 2: e000064.
9. Smith JM, Andrade JG, Human D, Field TS. Adults with complex congenital heart disease: cerebrovascular considerations for the neurologist. Front Neurol. 2019; 10: 329.
10. Feng Y, Yu D, Yang L, Da M, Wang Z, Lin Y, et al. Maternal lifestyle factors in pregnancy and congenital heart defects in offspring: review of the current evidence. Ital J Pediatr. 2014; 40: 1-7.
11. Andonian C, Langer F, Beckmann J, Bischoff G, Ewert P, Freilinger S, et al. Overweight and obesity: an emerging problem in patients with congenital heart disease. Cardiovasc Diagn Ther. 2019; 9 (Suppl 2): S360.
12. Saif-Ur-Rahman KM, Anwar I, Hasan M, Hossain S, Shafique S, Haseen F, et al. Use of indices to measure socio-economic status (SES) in South-Asian urban health studies: a scoping review. Syst Rev. 2018; 7: 196.
13. Wani RT. Socioeconomic status scales—modified Kuppuswamy and Udai Pareekh's scale updated for 2019. J Family Med Prim Care. 2019; 8: 1846-9.
14. Okoromah CA, Ekure EN, Lesi FE, Okunowo WO, Tijani BO, Okeiyi JC. Prevalence, profile and predictors of malnutrition in children with congenital heart defects: a case-control observational study. Arch Dis Child. 2011; 96: 354-60.
15. Chen CA, Wang JK, Lue HC, Hua YC, Chang MH, Wu MH. A shift from underweight to overweight and obesity in Asian children and adolescents with congenital heart disease. Paediatr Perinat Epidemiol. 2012; 26: 336-43.
16. Xiang L, Su Z, Liu Y, Zhang X, Li S, Hu S, et al. Effect of family socioeconomic status on the prognosis of complex congenital heart disease in children: an observational cohort study from China. Lancet Child Adolesc Health. 2018; 2: 430-9.
17. Peyvandi S, Baer RJ, Chambers CD, Norton ME, Rajagopal S, Ryckman KK, et al. Environmental and socioeconomic factors influence the live-born incidence of congenital heart disease: a population-based study in California. J Am Heart Assoc. 2020; 9: e015255.
18. Amini-Rarani M, Vahedi S, Borjali M, Nosratabadi M. Socioeconomic inequality in congenital heart diseases in Iran. Int J Equity Health. 2021; 20: 1.
19. Olugbuyi O, Smith C, Kaul P, Dover DC, Mackie AS, Islam S, et al. Impact of socioeconomic status and residence distance on infant heart disease outcomes in Canada. J Am Heart Assoc. 2022; 11: e026627.
20. Su XJ, Yuan W, Huang GY, Olsen J, Li J. Paternal age and offspring congenital heart defects: a national cohort study. PLoS One. 2015; 10: e0121030.
21. Abqari S, Gupta A, Shahab T, Rabbani MU, Ali SM, Firdaus U. Profile and risk factors for congenital heart defects: a study in a tertiary care hospital. Ann Pediatr Cardiol. 2016; 9: 216-21.
22. Taylor K, Elhakeem A, Thorbjørnsrud Nader JL, Yang TC, Isaevska E, Richiardi L. Effect of maternal prepregnancy/early-pregnancy body mass index and pregnancy smoking and alcohol on congenital heart diseases: a parental negative control study. J Am Heart Assoc. 2021; 10: e020051.
23. Deng C, Pu J, Deng Y, Xie L, Yu L, Liu L, et al. Association between maternal smoke exposure and congenital heart defects from a case-control study in China. Sci Rep. 2022; 12: 14973.
24. Wu Y, Liu B, Sun Y, Du Y, Santillan MK, Santillan DA, et al. Association of maternal prepregnancy diabetes and gestational diabetes mellitus with congenital anomalies of the newborn. Diabetes Care. 2020; 43: 2983-90.
25. Sanapo L, Donofrio MT, Ahmadzia HK, Gimovsky AC, Mohamed MA. The association of maternal hypertensive disorders with neonatal congenital heart disease: analysis of a United States cohort. J Perinatol. 2020; 40: 1617-24.
26. Saad H, Sinclair M, Bunting B. Maternal sociodemographic characteristics, early pregnancy behaviours, and livebirth outcomes as congenital heart defects risk factors—Northern Ireland 2010-2014. BMC Pregnancy Childbirth. 2021; 21: 1-13.
27. Savla JJ, Putt ME, Huang J, Parry S, Moldenhauer JS, Reilly S, et al. Impact of maternal-fetal environment on mortality in children with single ventricle heart disease. J Am Heart Assoc. 2022; 11: e020299.

Original Investigation | Özgün Araştırma

Gazi Med J 2026;37(2):189-194

DOI: http://dx.doi.org/10.12996/gmj.2026.4555

Impact of First-Trimester Vitamin D Levels on Pregnancy Complications

Birinci Trimester Vitamin D Düzeylerinin Gebelik Komplikasyonları Üzerine Etkisi

Seda Fidan[1], Bengü Mutlu Sütçüoğlu[2], Şevval Didem Yiğit[1], Cemre Karçaaltıncaba[3], Pınar Tokdemir Çalış[4], Deniz Karçaaltıncaba[4]

[1]Gazi University, Faculty of Medicine, Ankara, Türkiye
[2]Clinic of Obstetrics and Gynaecology, University of Health Sciences Türkiye, Ankara Atatürk Sanatoryum Research and Training Hospital, Ankara, Türkiye
[3]Clinic of Obstetrics and Gynaecology, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Ankara, Türkiye
[4]Department of Perinatology, Gazi University, Faculty of Medicine, Ankara, Türkiye

ABSTRACT

Objective: Vitamin D deficiency during pregnancy is a global health issue, affecting a significant portion of pregnant women. It has been linked to various adverse maternal and fetal outcomes, yet studies offer conflicting results on its direct impact. This study aimed to evaluate the prevalence of vitamin D deficiency in the first trimester and its association with pregnancy complications.

Methods: This retrospective cohort study involved 1,999 pregnant women who attended Gazi University's Obstetrics and Gynecology Clinic and had their vitamin D levels measured during the first trimester between January 2018 and August 2023. Exclusions included incomplete data, multiple pregnancies, and significant comorbidities. Participants were categorized into four groups based on their vitamin D levels, and associations with fetal and adverse pregnancy outcomes

Results: When we grouped pregnant women according to their vitamin D levels, 36.3% were severely deficient (<10 ng/mL), 47.5% were deficient (10–19 ng/mL), and 13.1% were insufficient (20–30 ng/mL). Only 63 (3.2%) pregnant women had sufficient vitamin D levels (>30 ng/mL). A significant correlation was found between vitamin D levels and age, with older women showing higher levels ($p < 0.001$). There was no clear relationship between vitamin D levels and other pregnancy complications, such as preeclampsia, gestational diabetes, preterm labor, or adverse neonatal outcomes. The incidence of abortion in women with vitamin D levels >30 ng/mL was 18%, which was significantly higher than in lower vitamin D groups ($p = 0.029$). Vitamin D levels were significantly lower in pregnancies with term birth weight below 2500 grams ($p = 0.024$).

ÖZ

Amaç: Gebelikte vitamin D eksikliği, gebelerin önemli bir kısmını etkileyen küresel bir sağlık sorunudur. Çeşitli olumsuz maternal ve fetal sonuçlarla ilişkilendirilmiş olmakla birlikte, doğrudan etkisine ilişkin çalışmalar çelişkili sonuçlar bildirmektedir. Bu çalışmada, birinci trimesterde vitamin D eksikliği prevalansının ve bunun gebelik komplikasyonları ile ilişkisinin değerlendirilmesi amaçlandı.

Yöntemler: Bu retrospektif kohort çalışmasına, Ocak 2018 ile Ağustos 2023 tarihleri arasında Gazi Üniversitesi Kadın Hastalıkları ve Doğum Kliniğine başvuran ve birinci trimesterde vitamin D düzeyi ölçülen 1999 gebe dahil edildi. Eksik verisi olanlar, çoğul gebeliği bulunanlar ve önemli komorbiditeleri olanlar dışlandı. Katılımcılar vitamin D düzeylerine göre dört gruba ayrıldı ve fetal sonuçlar ile olumsuz gebelik sonuçları arasındaki ilişkiler değerlendirildi.

Bulgular: Gebeler vitamin D düzeylerine göre gruplandırıldığında, %36,3'ünde ciddi eksiklik (<10 ng/mL), %47,5'inde eksiklik (10–19 ng/mL) ve %13,1'inde yetersizlik (20–30 ng/mL) saptandı. Yalnızca 63 (%3,2) gebenin vitamin D düzeyi yeterliydi (>30 ng/mL). Vitamin D düzeyleri ile yaş arasında anlamlı bir ilişki bulundu ve ileri yaştaki kadınlarda vitamin D düzeylerinin daha yüksek olduğu görüldü ($p < 0,001$). Vitamin D düzeyleri ile preeklampsi, gestasyonel diyabet, preterm eylem veya olumsuz neonatal sonuçlar gibi diğer gebelik komplikasyonları arasında belirgin bir ilişki saptanmadı. Vitamin D düzeyi >30 ng/mL olan kadınlarda abortus oranı %18 olup, bu oran daha düşük vitamin D düzeylerine sahip gruplara göre anlamlı olarak daha yüksekti ($p = 0,029$). Termde doğum ağırlığı 2500 gramın altında olan gebeliklerde vitamin D düzeyleri anlamlı olarak daha düşüktü ($p = 0,024$).

Cite this article as: Fidan S, Mutlu Sütçüoğlu B, Yiğit ŞD, Karçaaltıncaba C, Tokdemir Çalış P, Karçaaltıncaba D. Impact of first-trimester vitamin D levels on pregnancy complications. Gazi Med J. 2026;37(2):189-194

Address for Correspondence/Yazışma Adresi: Bengü Mutlu Sütcüoğlu, Clinic of Obstetrics and Gynaecology, University of Health Sciences Türkiye, Ankara Atatürk Sanatoryum Research and Training Hospital, Ankara, Türkiye
E-mail / E-posta: drbengumutlu@gmail.com
ORCID ID: orcid.org/0000-0002-5594-1719

Received/Geliş Tarihi: 24.09.2025
Accepted/Kabul Tarihi: 07.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

Conclusion: Our study confirms a high prevalence of vitamin D deficiency among women in the first trimester of pregnancy and shows that vitamin D levels increase with advancing maternal age. Low vitamin D levels were not associated with adverse pregnancy outcomes.

Keywords: Vitamin D deficiency, pregnancy complications, first trimester, obstetric outcomes

ÖZ

Sonuç: Çalışmamız, gebeliğin birinci trimesterindeki kadınlarda vitamin D eksikliğinin yüksek prevalansta olduğunu ve vitamin D düzeylerinin maternal yaş arttıkça yükseldiğini göstermektedir. Düşük vitamin D düzeyleri, olumsuz gebelik sonuçları ile ilişkili bulunmamıştır.

Anahtar Sözcükler: Vitamin D eksikliği, gebelik komplikasyonları, birinci trimester, obstetrik sonuçlar

INTRODUCTION

Vitamin D deficiency, indicated by circulating low levels of 25-hydroxyvitamin (OH) D, affects approximately 40% of pregnant women (1). In Mediterranean countries, prevalence studies have showed that this rate can reach to 60-80% (2). As pregnancy progresses, the need for vitamin D increases, potentially exacerbating vitamin D deficiency. Severe vitamin D deficiency, defined as 25-OH vitamin D concentrations below 10 ng/mL, is associated with impaired bone homeostasis, rickets, and neonatal fractures (3). However, the effects of milder deficiencies on maternal and fetal health remain unclear.

The vitamin D plays an important role for fertilization, placental development, the pregnancy process, and infant health (4). Deficiency has been linked to increased risks of miscarriage, preeclampsia, preterm birth, gestational diabetes and intrauterine growth restriction (5). However, some observational and randomized studies have not found a consistent association between vitamin D levels and adverse obstetric outcomes (5). These inconsistencies may be due to racial and individual factors or the different gestational weeks when studies were conducted.

Most randomized studies on this topic focus on later gestational weeks or have small sample sizes. To address this gap, our study aims to investigate the relationship between baseline vitamin D levels measured in the first trimester–before vitamin supplementation begins, during the critical period of embryogenesis–and gestational morbidity and mortality outcomes.

MATERIALS AND METHODS

This retrospective cohort study was conducted at the Gazi University Obstetrics and Gynecology Clinic and included pregnant women who visited the clinic between January 2018 and August 2023. During this period, 4314 consecutive intrauterine pregnancies who had their vitamin D levels measured in the first trimester were included in the study. Medical records of pregnant patients were reviewed using the hospital information system. Women with incomplete data, those who were lost to follow-up at our clinic, those who were not examined in the first trimester, and those with comorbidities (history of phosphocalcic disorders, hypercalcemia, malabsorptive diseases, pregestational diabetes, preeclampsia, chronic hypertension, or renal disease) were excluded. Only singleton pregnancies were included in the study. A total of 1999 pregnant women were included in the study (Table 1).

The hospital records were used to collect information on maternal age, previous delivery methods, and Rh incompatibility. Pregnancy outcomes (miscarriage, ectopic pregnancy, stillbirth, preterm birth) and complications during pregnancy (vaginal bleeding, preterm premature rupture of membranes, threatened preterm labor, preeclampsia, gestational diabetes mellitus, cholestasis, placental anomalies, neural tube defects, oligohydramnios, or polyhydramnios) were documented. Delivery records were examined for delivery method, difficult labor, postpartum atony, meconium staining, and fetal distress. Additionally, records of the infants born from these pregnancies were reviewed, noting birth weight and length, 1st- and 5th-minute APGAR scores, and the infants' genders.

Women who participated in the study were divided into four groups based on their vitamin D levels: Group 1 (<10 ng/mL, severely deficient), Group 2 (10–19 ng/mL, deficient), Group 3 (20–30 ng/mL, insufficient), and Group 4 (>30 ng/mL, sufficient).

This study was approved by the Ethics Committee of Gazi University (approval number: 2023-1475, date: 19.12.2023).

Statistical Analysis

Data from the study were analyzed using IBM SPSS Statistics 23. Categorical variables are presented as percentages, and continuous variables are presented as mean (standard deviation) and median (maximum–minimum). Normality was assessed using the Shapiro–Wilk and Kolmogorov–Smirnov tests. Categorical variables were compared using the Pearson chi-square test or Fisher's exact test, as appropriate. For non-normally distributed continuous variables, the Mann–Whitney U test (two groups) and Kruskal–Wallis test (≥3 groups) were used. For normally distributed variables, the student's t-test (two groups) and one-way analysis of variance (≥3 groups) were used. A p value < 0.05 was considered statistically significant.

RESULTS

A total of 1999 pregnant women were included in the study. The study flow diagram of patient selection is presented in Figure 1. The demographic data and general characteristics for the four groups, including patients' age, parity, duration of gestation, and characteristics of infants born from these pregnancies, are presented in Table 1. When we grouped pregnant women according to their, vitamin D levels, 36.3% of pregnant women were severely deficient (<10 ng/mL) 47.5% were deficient (10–19 ng/mL), 13.1% were insufficient (20–30 ng/mL). Only 63 (3.2%) pregnant women were found to have sufficient vitamin D levels (>30 ng/mL). A significant difference in age was observed between the groups, with the vitamin D–sufficient group being older ($p < 0.001$). Gestational age at birth, newborn weight and 1st and 5th minutes Apgar scores were similar between vitamin D groups ($p = 0.44$, $p = 0.99$, $p = 0.15$, $p = 0.46$ respectively)

When the population was categorized into three maternal age groups (<26, 26–35, and >35), a significant difference was observed among the groups (p < 0.001) (Table 2). Vitamin D levels were found to increase with advancing maternal age (Figure 2).

When pregnancy complications were compared among the patients, no statistically significant relationship was observed between vitamin D levels and complications such as preeclampsia, gestational diabetes, and preterm labor (p = 0.07, p = 0.49, p = 0.64, respectively) or adverse fetal outcomes such as stillbirth and fetal distress (p = 0.82, p = 0.91). The incidence of abortion in women with vitamin D levels >30 ng/mL was 18%, significantly higher than in lower vitamin D groups (p = 0.029) (Table 3). Due to the absence of a gradual trend among the groups, a subgroup analysis was performed. For abortion, statistically significant differences were identified between Group 1 and Group 4, as well as between Group 2 and Group 4 (p = 0.008 and p = 0.006, respectively).

The birth weights of the patients whose pregnancies ended at term (>37 gestational weeks) were analyzed separately. A statistically significant difference was found between vitamin D levels and birth weights below 2500 grams small-for-gestational-age (SGA) (p = 0.024) (Table 4). The mean and standard deviation of vitamin D levels in term SGA infants were 12.9 (±4.8), while the mean and standard deviation of the group with birth weights over 2500 grams were 14.2 (±5.9).

DISCUSSION

Our study used data from a tertiary center in Türkiye to evaluate the relationship between first-trimester vitamin D levels and adverse pregnancy outcomes and to describe vitamin D status in the reproductive population. The screening of women at low risk for pregnancy complications and the large patient population distinguish our study from other published studies. In our study, vitamin D levels increased with age, but no clear relationship was found between these levels and pregnancy-related complications.

Our study showed that vitamin D deficiency is a common health problem among women of reproductive age in Türkiye. Our prevalence of vitamin D insufficiency was 96.8%. The prevalence of vitamin D insufficiency ranges from 22% to 77% in northern countries (such as Germany and The Netherlands) and from 46% to 97% in Asian countries (6-9). When comparing prevalence studies conducted in different countries, researchers should consider racial and geographical differences. Nevertheless, the consistently high prevalence reported across various regions, irrespective of ethnic and geographical variations, underscores that vitamin D insufficiency is a global issue. Prevalence studies use varying cut-off values for vitamin D levels. For instance, some studies define insufficiency at thresholds of 15 ng/mL or 20 ng/mL, leading to differences in reported prevalence (6,10-12). The findings revealed a high prevalence of vitamin D deficiency, particularly among younger age groups. In our study, vitamin D deficiency was more severe among younger women, and vitamin D levels tended to normalize with age. This normalization may be associated with better dietary habits, increased sun exposure, or improved awareness of vitamin D supplementation among older age groups.

This study, which involved a large patient population and identified high rates of vitamin D insufficiency and deficiency, found no significant overall increase in pregnancy complications. Although the Endocrine Society recommends routine vitamin D supplementation for all pregnant women to prevent intrauterine mortality, preeclampsia, preterm labor, and SGA complications, our study,

Table 1. Characteristics of patients according to vitamin D levels.

	Vitamin D Level (ng/mL) n (%)				
All	**<10** **725 (36.3%)** **Mean ± SD**	**10-19** **949 (47.5%)** **Mean ± SD**	**20-30** **262 (13.1%)** **Mean ± SD**	**>30** **63 (3.2%)** **Mean ± SD**	**p**
Age	28 ± 5	29 ± 5	30 ± 5	32 ± 5	<0.001
Pregnancy outcome (days)	256 ± 50	256 ± 50	256 ± 49	233 ± 83	0.443
Infant weight (gr)	3169 ± 533	3173 ± 532	3172 ± 497	3160 ± 447	0.994
Infant height (cm)	48 ± 3	48 ± 3	48 ± 3	48 ± 2	0.578
APGAR score at 1 minute	8 ± 1	8 ± 1	9 ± 1	9 ± 1	0.147
APGAR score at 5 minutes	10 ± 4	9 ± 1	9 ± 1	10 ± 1	0.458
	n (%)	**n (%)**	**n (%)**	**n (%)**	**p**
Parity					0.740
Nulliparous	285 (39.3%)	358 (37.7%)	107 (40.8%)	27 (42.9%)	
Multiparous	440 (60.7%)	591 (62.3%)	155 (59.2%)	36 (57.1%)	
Vaginal	191 (43.4%)	221 (37.4%)	47 (30.3%)	13 (36.1%)	
Cesarean	249 (56.6%)	370 (62.6%)	108 (69.7%)	23 (63.9%)	
Infant gender					0.260
Male	362 (49.9%)	502 (52.9%)	145 (55.3%)	40 (63.5%)	
Female	363 (50.1%)	447 (47.1%)	117 (44.7%)	23 (36.5%)	

SD: Standard deviation.

Table 2. Vitamin D levels by age.

	Age (n)			p-value
	≤25 (541)	26-35 (1181)	>35 (277)	
Vitamin D level (ng/mL) mean (±SD)	12.8 ± 4.6	14.6 ± 6.1	15.8 ± 7.3	<0.001

SD: Standard deviation.

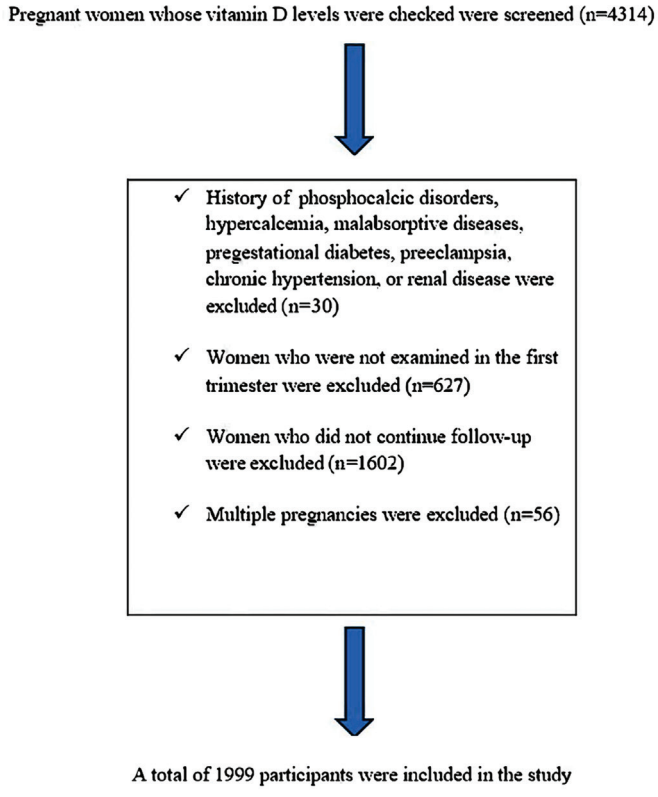

Figure 1. Flow diagram of patient eligibility.

which included a large patient population with a high prevalence of vitamin D deficiency, did not identify a significant increase in these pregnancy complications (13).

The incidence of abortion in women with vitamin D levels >30 ng/mL was significantly higher than in women with lower vitamin D levels. This higher risk may be associated with the older age of this cohort. Contrary to our finding, the literature includes studies reporting that vitamin D deficiency predisposes to miscarriage and threatened miscarriage (4,14,15). Furthermore, women with a history of at least one pregnancy loss have lower vitamin D levels than those without such a history. A recently published meta-analysis on recurrent pregnancy loss reported that preconception vitamin D levels play a role in preventing miscarriage (16). However, this study did not include a subgroup analysis regarding maternal age and supplementation. As our findings highlight, the importance of considering maternal age in studies of miscarriage is another critical factor that needs to be emphasized Considering the tendency for vitamin D levels to be higher in older mothers, increasing vitamin D in this group may not be effective in preventing miscarriages.

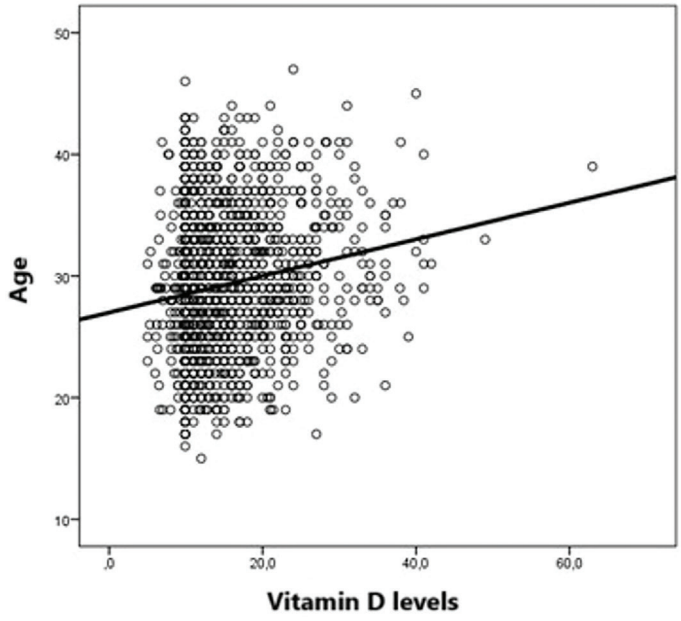

Figure 2. Vitamin D levels by age.

Low birth weight was significantly associated with lower vitamin D levels. This finding suggests that adequate vitamin D levels may be important for optimal fetal growth. A randomized controlled trial conducted in Bangladesh initiated supplementation at three different doses between the 17th and 24th gestational weeks and found no significant correlation with birth weight (17). In contrast, a prospective cohort study published in Sweden identified an association between vitamin D deficiency and SGA (18). A meta-analysis also reported that vitamin D levels below 12 ng/mL increased the risk of SGA by 1.5 times (odds ratio 1.58; 95% confidence interval: 1.14, 2.22) (19). A positive correlation between vitamin D supplementation and infant weight was observed. Infants whose mothers received supplementation had higher weights at 3, 6, and 12 months of age. A six-year follow-up randomized controlled trial conducted in Denmark, starting supplementation from the 24th gestational week, reported no effect on infant weight but found that vitamin D levels below 12 ng/mL were associated with growth retardation (20). These findings from various geographic locations suggest that maintaining sufficient first-trimester serum vitamin D levels is crucial for supporting fetal development.

Our study demonstrated that vitamin D insufficiency and deficiency did not lead to a significant increase in obstetric complications. The relationship between vitamin D levels and adverse obstetric outcomes remains unclear. This uncertainty may stem from conflicting results in the literature, variations in gestational age at screening, differing cut-off levels, small sample sizes, and most studies focusing on a single obstetric outcome. Therefore, our study, with its large sample size and examination of multiple outcomes within the same cohort, will help clarify the relationship between vitamin D levels and adverse obstetric outcomes in the scientific literature. The primary strengths of our study are its large sample size and homogeneous ethnic background of the patient population, which enhance the reliability of the findings. Additionally, measuring vitamin D levels during the first trimester provides critical

Table 3. Pregnancy complications according to vitamin D levels.

	<10 ng/mL 725 (36.3%) n (%)	**10-19 ng/mL 949 (47.5%) n (%)**	**20-30 ng/mL 262 (13.1%) n (%)**	**>30 ng/mL 63 (3.2%) n (%)**	**p**
Vaginal bleeding	41 (5.7%)	65 (6.8%)	14 (5.3%)	4 (6.3%)	0.708
PPROM	19 (2.6%)	35 (3.7%)	4 (1.5%)	0 (0.0%)	0.115
Preterm labor	33 (4.6%)	46 (4.8%)	14 (5.3%)	1 (1.6%)	0.640
Dystocia	48 (6.6%)	58 (6.1%)	18 (6.9%)	4 (6.3%)	0.962
Postpartum hemorrhage	2 (0.3%)	7 (0.7%)	4 (1.5%)	0 (0.0%)	0.157
Atony	1 (0.1%)	2 (0.2%)	0 (0.0%)	0 (0.0%)	0.867
Fetal distress	23 (3.2%)	28 (3.0%)	6 (2.3%)	2 (3.2%)	0.911
Meconium staining	1 (0.1%)	4 (0.4%)	2 (0.8%)	0 (0.0%)	0.462
Preeclampsia	22 (3.0%)	14 (1.5%)	9 (3.4%)	3 (4.7%)	0.070
GDM	47 (6.5%)	70 (7.4%)	15 (5.7%)	2 (3.2%)	0.494
Abortion	58 (8%)	69 (7.3%)	20 (7.7%)	11 (18%)	0.029
Stillbirth	10 (1.4%)	12 (1.3%)	3 (1.1%)	0 (0 %)	0.821
Ectopic pregnancy	1 (0.1%)	0 (0 %)	1 (0.4%)	0 (0 %)	0.363
Cholestasis of pregnancy	7 (1.0%)	11 (1.2%)	4 (1.5%)	0 (0.0%)	0.732
Thrombotic event	20 (2.8%)	30 (3.2%)	8 (3.1%)	2 (3.2%)	0.971
Oligohydramnios	4 (0.6%)	7 (0.7%)	3 (1.1%)	0 (0%)	0.697
Polyhydramnios	7 (1.0%)	5 (0.5%)	1 (0.4%)	0 (0.0%)	0.564
Neural tube defect	2 (0.3%)	2 (0.2%)	0 (0.0%)	1 (1.6%)	0.155
Placental anomaly	10 (1.4%)	10 (1.1%)	2 (0.8%)	0 (0.0%)	0.680
SGA (<2500g)	65 (8.9%)	84 (8.8%)	23 (8.7%)	4 (6.3%)	0.983

SGA: Small-for-gestational-age, PPROM: Preterm premature rupture of membranes, GDM: Gestational diabetes mellitus.

Table 4. Infant weight at term births.

	Infant weight (n)		
	< 2500 (97)	**> 2500 (1555)**	**p-value**
Vitamin D level mean (±SD)	12.9 ± 4.8	14.2 ± 6	0.024

SD: Standard deviation.

baseline data before supplementation begins. However, the lack of information on vitamin D supplementation and the absence of follow-up measurements in later trimesters are notable limitations. This restricts the ability to assess changes in vitamin D status throughout pregnancy. Despite these limitations, our study offers valuable insights into the relationship between early pregnancy vitamin D levels and obstetric outcomes.

CONCLUSION

In our comprehensive study of first-trimester pregnant women, we identified a high prevalence of vitamin D deficiency, and vitamin D levels increased with advancing maternal age. There was no clear relationship between vitamin D levels and most pregnancy-related complications, except for a modest association with term low birth weight and a higher abortion rate in the vitamin D–sufficient group, which may reflect confounding by maternal age. Prospective studies incorporating supplementation data and serial measurements are warranted.

Ethics

Ethics Committee Approval: This study was approved by the Ethics Committee of Gazi University (approval number: 2023-1475, date: 19.12.2023).

Informed Consent: Because of the retrospective design of the study, individual informed consent was waived.

Footnotes

Authorship Contributions

Surgical and Medical Practices: S.F., B.M.S., Ş.D.Y., C.K., P.T.Ç., D.K., Concept: S.F., B.M.S., Ş.D.Y., C.K., P.T.Ç., D.K., Design: S.F., B.M.S., Ş.D.Y., C.K., P.T.Ç., D.K., Data Collection or Processing: S.F., B.M.S., Ş.D.Y., C.K., P.T.Ç., D.K., Analysis or Interpretation: S.F., B.M.S., Ş.D.Y., C.K., P.T.Ç., D.K., Literature Search: S.F., B.M.S., Ş.D.Y., C.K., P.T.Ç., D.K., Writing: S.F., B.M.S., Ş.D.Y., C.K., P.T.Ç., D.K.

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

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

REFERENCES

1. Amrein K, Scherkl M, Hoffmann M, Neuwersch-Sommeregger S, Köstenberger M, Tmava Berisha A, et al. Vitamin D deficiency 2.0: an update on the current status worldwide. Eur J Clin Nutr. 2020; 74: 1498-513.
2. Cui A, Zhang T, Xiao P, Fan Z, Wang H, Zhuang Y. Global and regional prevalence of vitamin D deficiency in population-based studies from 2000 to 2022: a pooled analysis of 7.9 million participants. Front Nutr. 2023; 10: 1070808.
3. Bouillon R, Carmeliet G. Vitamin D insufficiency: Definition, diagnosis and management. Best Pract Res Clin Endocrinol Metab. 2018; 32: 669-84.
4. Heyden EL, Wimalawansa SJ. Vitamin D: effects on human reproduction, pregnancy, and fetal well-being. J Steroid Biochem Mol Biol. 2018; 180: 41-50.
5. van der Pligt P, Willcox J, Szymlek-Gay EA, Murray E, Worsley A, Daly RM. Associations of maternal vitamin d deficiency with pregnancy and neonatal complications in developing countries: a systematic review. Nutrients. 2018; 10: 640.
6. Scharla SH. Prevalence of subclinical vitamin D deficiency in different European countries. Osteoporos Int. 1998; 8 Suppl 2: S7-12.
7. van der Meer IM, Middelkoop BJ, Boeke AJ, Lips P. Prevalence of vitamin D deficiency among Turkish, Moroccan, Indian and sub-Sahara African populations in Europe and their countries of origin: an overview. Osteoporos Int. 2011; 22: 1009-21.
8. Lips P. Vitamin D status and nutrition in Europe and Asia. J Steroid Biochem Mol Biol. 2007; 103: 620-5.
9. Lips P, Cashman KD, Lamberg-Allardt C, Bischoff-Ferrari HA, Obermayer-Pietsch B, Bianchi ML, et al. Current vitamin D status in European and Middle East countries and strategies to prevent vitamin D deficiency: a position statement of the European Calcified Tissue Society. Eur J Endocrinol. 2019; 180: P23-54.
10. Harvey NC, Holroyd C, Ntani G, Javaid K, Cooper P, Moon R, et al. Vitamin D supplementation in pregnancy: a systematic review. Health Technol Assess. 2014; 18: 1-190.
11. Zhu K, Whitehouse AJ, Hart PH, Kusel M, Mountain J, Lye S, et al. Maternal vitamin D status during pregnancy and bone mass in offspring at 20 years of age: a prospective cohort study. J Bone Miner Res. 2014; 29: 1088-95.
12. Yu CK, Sykes L, Sethi M, Teoh TG, Robinson S. Vitamin D deficiency and supplementation during pregnancy. Clin Endocrinol (Oxf). 2009; 70: 685-90.
13. Holick MF, Binkley NC, Bischoff-Ferrari HA, Gordon CM, Hanley DA, Heaney RP, et al. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2011; 96: 1911-30.
14. Hassan SA. Vitamin D deficiency and first trimester miscarriage: a case control study. Indian J Public Health Res Dev. 2018; 9: 1645.
15. Porcaro G, Laganà AS, Neri I, Aragona C. The association of high-molecular-weight hyaluronic acid (HMWHA), alpha lipoic acid (ALA), magnesium, vitamin B6, and vitamin D improves subchorionic hematoma resorption in women with threatened miscarriage: a pilot clinical study. J Clin Med. 2024; 13: 706.
16. Tamblyn JA, Pilarski NSP, Markland AD, Marson EJ, Devall A, Hewison M, et al. Vitamin D and miscarriage: a systematic review and meta-analysis. Fertil Steril. 2022; 118: 111-22.
17. Roth DE, Gernand AD, Morris SK, Pezzack B, Islam MM, Dimitris MC, et al. Maternal vitamin D supplementation during pregnancy and lactation to promote infant growth in Dhaka, Bangladesh (MDIG trial): study protocol for a randomized controlled trial. Trials. 2015; 16: 300.
18. Bärebring L, Bullarbo M, Glantz A, Hulthén L, Ellis J, Jagner Å, et al. Trajectory of vitamin D status during pregnancy in relation to neonatal birth size and fetal survival: a prospective cohort study. BMC Pregnancy Childbirth. 2018; 18: 51.
19. Chen Y, Zhu B, Wu X, Li S, Tao F. Association between maternal vitamin D deficiency and small for gestational age: evidence from a meta-analysis of prospective cohort studies. BMJ Open. 2017; 7: e016404.
20. Nørrisgaard PE, Haubek D, Kühnisch J, Chawes BL, Stokholm J, Bønnelykke K, et al. Association of high-dose Vitamin D supplementation during pregnancy with the risk of enamel defects in offspring: a 6-year follow-up of a randomized clinical trial. JAMA Pediatr. 2019; 173: 924-30.

Gazi Med J 2026;37(2):195-202

DOI: http://dx.doi.org/10.12996/gmj.2026.4487

Evaluation of the Growth of Preterms Born in 2001–2002 During Adolescence

2001–2002 Yılında Doğan Prematüre Bebeklerin Adölesan Dönemde Büyümelerinin Değerlendirilmesi

Kübra Baskın[1,2], Sultan Kavuncuoğlu[2]

[1]Department of Pediatric Immunology and Allergy, Gazi University, Faculty of Medicine, Ankara, Türkiye
[2]Department of Pediatrics, University of Health Sciences Türkiye, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye

ABSTRACT

Objective: In this study, we aimed to evaluate growth characteristics of preterms followed up in a neonatal intensive care unit (NICU) during adolescence.

Methods: The preterms born in 2001–2002 who were followed up in the NICU of our hospital were examined in 2016. Data from the prenatal, natal, and postnatal periods, as well as at the 6–7 years, were obtained from a database with follow–up assessments during adolescence. Measures of growth (height, weight, body mass index), the percentile curves for Turkish children, and the target height formula were used in the assessment. Factors affecting growth, such as maternal problems and neonatal morbidities, were questioned.

Results: Thirty–seven preterm babies (19 females and 18 males) were included in the study. The mean gestational age was 31.5 (28–34) weeks, and the birth weight was 1298 (820–1870) grams. Six newborns (17.6%), and three adolescents (8.1%) were below 3rd percentile for height. All cases had height percentiles within the normal range at 6–7 years of age. Twenty–six adolescents (70.3%) achieved the target height. Eleven premature (29.7%) during the neonatal period, one child (2.7%) during the school period, and one adolescent (3%) were below 3rd percentile for weight. During the neonatal period, one case for height and two cases for weight were above 90th percentiles. While, during school period, 6 cases for height, 2 cases for weight were above 90th percentiles; 1 case for height, 6 cases for weight were above 90th percentiles during adolescence. The frequency of obesity increased from 2.7% to 13.5% during adolescence. 90% of individuals below 3rd percentile of birth weight reached normal weight percentiles during adolescence. Prenatal, natal, and postnatal risk factors, as well as duration of breastfeeding, were not found to be effective in achieving the target height during school–age and adolescent periods.

ÖZ

Amaç: Çalışmamızda yenidoğan yoğun bakım ünitemizde (YYBÜ) izlenen prematürelerin adölesan dönemde büyüme özelliklerini değerlendirmeyi amaçladık.

Yöntemler: 2001–2002 yıllarında doğan ve hastanemiz YYBÜ takip edilen prematüreler 2016 yılında değerlendirildi. Veritabanından hastalara ait prenatal, natal, postnatal, 6–7 yaş ve adolesan dönem takip verileri elde edildi. Değerlendirmede büyüme ölçütleri (boy, kilo, vücut kitle indeksi), Türk çocuklarının büyüme eğrileri ve hedef boy formülü kullanıldı. Büyümeye etki eden faktörler (maternal sorunlar, neonatal dönem morbiditeleri vb.) sorgulandı.

Bulgular: Çalışmaya 37 preterm (19 kız, 18 erkek) dahil edildi. Ortalama gebelik yaşı 31,5 (28–34) hafta, doğum ağırlığı 1298 (820–1870) gramdı. Altı yenidoğan (%17,6) ve 3 adolesan (%8,1) boyca 3 persentil altındaydı. Tüm olgular 6–7 yaşta boyca normal persentillerdeydi. Yirmi altı adolesan (%70,3) hedef boya ulaşmıştı. Yenidoğan döneminde 11 prematüre (%29,7), okul döneminde bir çocuk (%2,7) ve bir adolesan (%3) ağırlıkça 3 persentil altındaydı. Yenidoğan döneminde, bir olgunun boyda ve 2 olgunun ağırlıkta 90 persentil üzerinde olduğu görüldü. Okul döneminde 6 olgu boyda, 2 olgu ağırlıkta 90 persentil üzerinde iken; adolesan dönemde 1 olgu boyda, 6 olgu ağırlıkta 90 persentil üzerindeydi. Obezite sıklığı adolesan dönemde %2,7'den %13,5'e yükselmişti. Doğum ağırlığı 3 persentil altındaki olguların %90'ı adolesan dönemde normal ağırlık persentillerine ulaşmıştı. Okul çağı ve adolesan dönemde hedef boya ulaşmada prenatal, natal ve postnatal risk faktörleri ile anne sütü alma süresinin etkili olmadığı saptandı.

Sonuç: Çalışmamızda, okul döneminde tüm olguların boyca normal persentillerde olmasına rağmen ergenlikte olguların %8,1'nin 3 persentil altında olduğunu saptadık. Ayrıca ağırlıkça 3 persentil altında

Cite this article as: Baskın K, Kavuncuoğlu S. Evaluation of the growth of preterms born in 2001–2002 during adolescence. Gazi Med J. 2026;37(2):195-202

Address for Correspondence/Yazışma Adresi: Kübra Baskın, Department of Pediatric Immunology and Allergy, Gazi University Faculty of Medicine, Ankara, Türkiye
E-mail / E-posta: akbaskin@hotmail.com
ORCID ID: orcid.org/0000–0002–4294–7492

Received/Geliş Tarihi: 06.07.2025
Accepted/Kabul Tarihi: 23.02.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

Conclusion: In this study, we found that all cases had normal percentiles for height in the school period, but on the other hand, 8.1% cases were <3rd percentiles during adolescence. Also, the frequency of cases that were <3rd percentiles for weight was similar during the school period and adolescence. These results reflected that, while catch–up growth could be achieved during the school years, it was not achievable during adolescence. It demonstrates that the growth of individuals born preterm is a dynamic phenomenon.

Keywords: Adolescent, growth, preterm birth, risk factors

ÖZ

olan olguların sıklığı okul dönemi ve ergenlik döneminde benzerdi. Bu sonuçlar, büyüme yakalamasının okul yıllarında sağlanabilmekle birlikte, ergenlik döneminde sağlanamadığını yansıtmaktadır. Bu durum prematüre doğan bireylerin büyümelerinin dinamik bir süreç olduğunu göstermektedir.

Anahtar Sözcükler: Adolesan, büyüme, preterm doğum, risk faktörleri

INTRODUCTION

In parallel with advances in perinatal and neonatal care worldwide, the survival of premature babies has increased, and more are being discharged from the hospital. The follow–up of high–risk preterm infants has highlighted several key questions concerning growth patterns and developmental milestones. Prematurity is associated with a range of short– and long–term complications, including impaired somatic growth, neurodevelopmental delays, and increased risk of chronic health conditions. These outcomes may persist into adolescence and adulthood, underscoring the importance of longitudinal monitoring. A holistic examination of the long–term consequences of premature birth is of great importance for prognosis.

Preterm infants are prone to early postnatal growth restriction; nevertheless, most subsequently achieve somatic catch–up growth. Increases in weight and length typically begin within the first months of life and often normalize by two years of age (1). Some postnatal comorbidities, such as bronchopulmonary dysplasia (BPD), retinopathy of prematurity (ROP), and neurodevelopmental impairment, are also reported to affect catch–up growth (2).

Numerous studies have examined long–term monitoring in very–low–birth–weight premature infants (3-5). Although studies on somatic growth and neurodevelopmental evaluation are available in infancy, early childhood, and school–age, data on adolescent follow–up are limited (6-8).

For this reason, we aimed to evaluate the somatic growth of preterm infants born in 2001–2002 who were followed up in our neonatal intensive care unit (NICU) during adolescence, with a specific focus on their anthropometric outcomes.

MATERIALS AND METHODS

Premature infants born in 2001–2002 at 28–34 weeks of gestational age (GA) according to the Ballard Score and followed up in the NICU of Ministry of Health, Bakırköy Obstetrics and Pediatrics Training and Research Hospital were included in the study. Individuals who are deceased, have metabolic disease, and have anomalies are excluded. Growth curves by Kurtoğlu et al. (9) were used to evaluate somatic growth during the neonatal period. Cases whose birth weight were <3rd percentile (rather than the more commonly used <10th percentile threshold) were recorded as small for GA (SGA) and divided into symmetrical or asymmetrical SGA categories. This stricter definition was chosen to increase specificity, ensuring that infants classified as SGA truly represented those at highest risk of growth failure and adverse outcomes, and to enable a more accurate comparison with cases presenting growth retardation during childhood and adolescence. Cases at the 3rd–90th percentile were considered as appropriate for GA (AGA), and cases >90th percentile were considered as large for GA (LGA). Also, we categorised cases according to GA [extremely preterm (less than 28 weeks), very preterm (28 to less than 32 weeks), moderate to late preterm (32 to 37 weeks)] and birthweight (<1000 g: extremely low birthweight, 1000 to 1499 g: very low birthweight, 1500 to 2500 g: low birthweight).

The following prenatal risk factors were assessed: premature rupture of membranes, chorioamnionitis, maternal diabetes mellitus, hypertensive disorders of pregnancy, presence of chronic disease, and intrauterine growth retardation. Among natal risk factors, the following were investigated: history of fetal distress, placenta previa, and placental detachment. The following postnatal risk factors were investigated and recorded: duration of stay in the NICU, mechanical ventilation (administration and duration), surfactant therapy, sepsis, hypoglycemia, hyperbilirubinemia, respiratory distress syndrome (RDS), BPD, intraventricular hemorrhage, posthemorrhagic hydrocephalus, and ROP. Additionally, breastfeeding duration was ascertained and recorded.

The somatic growth of these children was previously evaluated at school age (6–7 years). Data from the prenatal, natal, and postnatal periods and at 6–7 years of age were obtained from the database. Values >90th percentile during the neonatal, childhood, and adolescent periods were used as a criteria for advanced growth (10).

A single physician (KB) assessed the cohort at 14–15 years of age, following signed parental and child consent. The systemic examination of the patients was performed. Puberty status was assessed according to Tanner staging (11,12). Ages at menarche and thelarche were recorded based on parental recall.

The standing height and weight of the cases and parents were measured by the same physician (using a Sinbo weighing instrument and a Harpenden stadiometer) at least twice, and the averages were used for analysis. Values of the cases were placed in the percentile curves prepared for the Turkish children by Neyzi et al. (13). The target length based on the mother's and father's heights was also calculated and checked after the percentile chart was marked. When calculating the target height, the difference between men and women, according to each society's standards, was considered. For Turkish society, this difference was 13 cm (13,14).

Target height for a girl was calculated as $\frac{(\text{Father's height} - 13\text{ cm}) + (\text{mother's height})}{2}$

Target height for a boy was calculated as $\frac{(\text{Mother's height} + 13\text{ cm}) + (\text{father's height})}{2}$

Body mass index (BMI) was calculated as weight (kg)/height (m^2). Cases were identified according to the World Health Organization BMI classification (15). BMI percentiles were assessed using reference values for Turkish children (16).

This study was approved by the Ethics Committee of Ministry of Health İstanbul Kanuni Sultan Süleyman Training and Research Hospital (approval number: 15929, date: 30.09.2015). Patients and their parents provided the written informed consent.

Statistical Analysis

The findings were analyzed statistically using the Statistical Program in Social Sciences version 24.0 (2016). When evaluating the study data, McNemar test, Fisher exact test, Mann–Whitney U test, and Kruskal–Wallis test were used in addition to descriptive statistical methods (mean, median, standard deviation, frequency). Categorical variables were presented as numbers and percentages. Continuous variables were presented as mean ± standard deviation or median (minimum-maximum), depending on the distribution. The results were evaluated using a 95% confidence interval, with significance set at $p < 0.05$. Missing data were taken into account during analyses.

RESULTS

In this study, 37 adolescents born preterm (19 females, 18 males) were evaluated. Most cases (n = 20, 54%) were SGA; 40.5% (n = 15) were AGA; and most cases (n = 30, 80%) had birth weights of 1000–1500 grams. While the most common prenatal risk factors were intrauterine growth retardation (IUGR) and maternal hypertension, the most common postnatal risk factor was hyperbilirubinemia. Demographic characteristics and prenatal, natal, and postnatal risk factors of the cases are summarized in Table 1.

Growth in the cases during the newborn and school periods (6–7 years) was compared with adolescent growth (Table 2). During adolescence, the mean age of cases was 14.6 ± 0.78 years, the mean height was 162.3 ± 7.1 cm, the mean weight was 56.1 ± 13.8 kg, and the mean BMI was 21 ± 4.04 kg/m^2. Values for height, weight, BMI, and pubarche onset by gender are presented in Table 3. Age at pubarch onset was 12.06 ± 1.2 years, and at menarche onset was 12.2 ± 1.08 years. Only one case was treated for precocious puberty; in the other cases, the onset of pubertal growth and development was prompt.

While 6 cases (17.6%) were found to be <3rd percentiles for height in the newborn period, there were no cases <3rd percentiles for height in the school period. Three adolescents (8.1%) were <3rd percentiles for height (Table 2). Twenty–six cases (70.3%) reached target height by adolescence (Table 4). Of the eleven cases (6 males, 5 females) that did not reach target height, 4 were AGA, 3 were symmetrical SGA, and 4 were asymmetric SGA in the neonatal period (Table 4). Eleven premature (29.7%) during the neonatal period, one child (2.7%) during the school period, and one adolescent (3%) were <3rd percentile for weight.

Table 1. Demographic, perinatal and neonatal characteristics of cases.

	n	%
Gender		
Female	19	51.4
Male	18	48.6
Intrauterine growth		
Symmetric SGA*	8	21.6
Asymmetric SGA*	12	32.4
AGA*	15	40.5
LGA*	2	5.4
GA (week)		
≤28	4	10.8
29–32	20	54
≥33	13	35.1
Birth weight (gram)		
<1000 g	3	8.1
1000–1500 g	30	81
>1500 g	4	10.8
Prenatal risk factors		
IUGR	10	32.3
Hypertensive disorders of pregnancy	10	32.3
Maternal smoking	8	25.8
Mother's age >35 (year)	6	19.4
PROM	4	12.9
Presence of chronic disease	3	9.7
Plesantal dysfunction	2	6.5
Natal risk factors		
Fetal distress	8	100.0
Placenta detachment	4	50.0
Postnatal risk factors		
Hyperbilirubinemia	25	71.4
RDS	13	37.1
Surfactant therapy	13	37.1
Mechanical ventilation	13	37.1
Sepsis	11	31.4
ROP	6	17.1
Apnea	6	17.1
Intracranial hemorrhage	5	14.3
Hypoglycemia	2	5.7
PDA	2	5.7
Hydrocephalus	1	2.9
BPD	1	2.9

*AGA: Appropriate for gestational age, LGA: Large for gestational age, SGA: Small for gestational age, BPD: Bronchopulmonary dysplasia, IUGR: Intrauterine growth retardation, PDA: Patent ductus arteriosus, PROM: Premature rupture of membranes, RDS: Respiratory distress syndrome, ROP: Retinopathy of prematurity.

During the neonatal period, one case for height and two cases for weight were >90th percentiles. While during school period, 6 cases for height and 2 cases for weight were >90th percentiles; 1 case for height and 6 cases for weight were >90th percentiles during adolescence (Table 2). The Frequency of LGA was 5.4% during the neonatal period. Obesity prevalence, when evaluated according to BMI percentiles, was 2.7% during the school period and increased to 13.5% during adolescence.

To evaluate growth trajectories, height and weight percentiles at birth, at age 6–7 years, and during adolescence were examined. No statistically significant difference was observed between height percentiles at birth and in adolescence. Two of the six patients whose birth height was <3rd percentile (33.3%) were <3rd percentile for height currently, and the remainder (n = 4, 66.7%) reached the target height (Table 5). While all individuals with a birth weight <3rd percentiles reached normal weight percentiles at age 6–7 (p = 0.021), 90.9% reached normal weight percentiles during puberty (p = 0.002). Intrauterine growth characteristics (SGA, AGA, LGA), parental heights, and BMI did not significantly affect this group's attainment of the target height (Table 5).

Eight of 11 patients whose birth weight was <3rd percentile, and 18 of remaning 26 patients whose birth weight was >3rd percentile could reach target height. Birth weight was not found to have a statistically significant effect on target height. (p = 1.00)

When the relationship between birth height percentiles and prenatal, natal, and postnatal risk factors were examined, the incidence of hypoglycemia was higher in patients whose height <3rd percentile (p = 0.030). No statistically significant differences were observed between other risk factors and height percentiles at birth (Table 6). There was also no statistically significant association between birth weight percentiles and the aforementioned risk factors (Table 7).

Height and weight percentiles during school age and adolescence were examined according to the intrauterine growth status of the patients (AGA, LGA, SGA), and no statistically significant differences were found. Also, no statistically significant association was found between prenatal, natal, and postnatal risk factors and height and weight percentiles during school age and adolescence.

Breastfeeding duration ranged from 15 days to 2.5 years among the cases. There was no statistically significant relationship between breastfeeding duration and patients' growth (length and weight percentiles) during school–age and adolescence.

DISCUSSION

This study examined the somatic growth features of high–risk preterm infants, most of whom weighed <1500 grams, at approximately 14 years of age. Our results suggested that the growth trajectories of children born preterm differed across follow–up periods. Growth of the cases during the newborn and school periods (6–7 years) was compared with adolescent growth.

While 6 cases (17.6%) for height, 11 cases (29.7%) for weight were found to be <3rd percentiles in the newborn period, all cases caught up in length gain during school age. Only one child was found to be <3rd percentiles for weight. Three adolescents (8.1%) for height, 1 adolescent (3%) for weight were found to be <3rd percentiles. Twenty–six cases (70.3%) could reach target height by adolescence.

Different views exist in the literature regarding catch–up growth. Catch–up growth occurs during the first three years of life in cases without severe neurological impairment, cardiac disease, or genetic disease or syndrome (17). Niklasson et al. (18) showed that normal growth could be achieved at the age of 2 years with optimal postnatal care, even in cases with IUGR, in the absence of perinatal complications and severe congenital anomalies.

Unlike our study, in a study that evaluated neurodevelopmental and physical growth of very low birth weight preterms at an average of 87 months after birth, the rate of cases <3rd percentile was reported as 6% in height and 7% in weight (7). A study of 4423 late preterm children evaluated from birth to 14 years reported that preterm children had lower height and weight than full–term children across all age groups (19). In the Danish birth cohort in which children born between 1996 and 2003 were followed until age 18, GA was associated with height in infancy, but the difference in height between preterm and term children decreased during childhood. It has been reported that many individuals born preterm remain shorter than those born at term during adolescence. It has also been reported that BMI, which was relatively lower in preterm infants, became equal to that of term infants during childhood, with no significant difference by adolescence (20). We found the BMI during adolescence to be similar to normal values according to the WHO classification.

Table 2. Weight and height measurements of the patients at birth, 6–7 years of age and adolescence.

	Birth height		Birth weight		Height (6–7 years)		Weight (6–7 years)		Height (adolescent)		Weight (adolescent)	
	n	%	n	%	n	%	n	%	n	%	n	%
<3rd p	6	17.6	11	29.7	0	0.0	1	3.0	3	8.1	1	2.7
3–10 p	7	20.6	9	24.3	0	0.0	3	9.1	2	5.4	2	5.4
11–25 p	10	29.4	5	13.5	3	9.1	9	27.3	8	21.6	7	18.9
26–50 p	7	20.6	6	16.2	10	30.3	8	24.2	13	35.1	8	21.6
51–75 p	3	8.8	4	10.8	7	21.2	5	15.2	9	24.3	9	24.3
76–90 p	0	0.0	0	0.0	7	21.2	5	15.2	1	2.7	4	10.8
91–97 p	0	0.0	2	5.4	5	15.2	0	0.0	1	2.7	5	13.5
>97th p	1	2.9	0	0.0	1	3.0	2	6.1	0	0.0	1	2.7

p: Percentile.

In the EPICure study, which evaluated extremely premature infants, the authors reported that some catch–up growth in height and weight occurred between 2.5 and 6 years of age, but growth parameters at age 6 remained below population norms. In this study, individuals born preterm were lighter and shorter at 19 years of age than the control group (21).

In a study that examined growth characteristics of very low–weight premature babies at an average postnatal age of 36 months, it was found that gender and AGA or SGA status had no effect on growth. Mechanical ventilation, chronic disease, advanced intracranial hemorrhage, ROP, and lack of breastfeeding were identified as risk factors for growth retardation (6). Akar et al. (8) emphasized in their study evaluating somatic growth in late preterm preschoolers that being AGA or SGA was not a statistically significant predictor of reaching the target height. By contrast, in a Brazilian birth cohort study, preterm children with SGA and low birth weight were shorter and thinner than term children at 5 years (22). Similar to the previous studies (6,8), in our study intrauterine growth characteristics and the aforementioned postnatal risk factors were not associated with growth at school age or during adolescence.

Prenatal risk factors that affect growth include maternal hypertension, chronic disease, gestational diabetes, maternal smoking, and placental dysfunction (23,24). However, our study revealed that prenatal and natal risk factors were not associated with growth at birth, during school age, and during adolescence.

Breastfeeding is also important for growth. In a study, the growth of SGA babies fed breast milk was faster than that of babies fed formula (25). In our study, no statistically significant relationship was observed between breastfeeding duration and patients' growth, measured by height and weight percentiles, during school age and adolescence.

Our study revealed that during the neonatal and school–age periods 2 cases, during adolescence 6 cases for weight were >90th percentiles. Two cases who were >90th percentiles for weight at school age were asymmetrical SGA at birth. Their birth weights were 1100 grams and 1600 grams. While one case had a history of RDS, BPD, PDA, and mechanical ventilation, the other case had only hyperbilirubinemia. Of the six cases who were >90th percentiles for weight during adolescence, 5 were SGA (asymmetric: 3, symmetric: 2) and, one was AGA at birth. Birth weights ranged from 1000 to 1600 grams. In most cases, sepsis, apnea, and hyperbilirubinemia were present as postnatal risk factors. The frequency of obesity increased from 2.7% to 13.5% during adolescence. There was no significant association of intrauterine growth status or prenatal, natal, and postnatal risk factors with weight and height in adolescence.

Few studies have examined the potential link between prematurity and the risk of childhood obesity and have yielded different results. Unlike our study, a study examining extremely preterm infants reported that females with high gains in height/weight or weight during the 2 years after NICU discharge had greater odds of obesity at 10 years, but not at 15 years (26). Another study revealed that premature infants had a greater likelihood of childhood obesity at ages 8 to 11, and no significant difference in childhood obesity risk was found between SGA and AGA preterm infants (27).

Table 3. Height, weight, BMI and pubarche onset distrubition according to gender during adolescence.

	Boys	Girls
Height (cm, median)	166.6	159
(min–max)	(144.5–179.7)	(148–172)
Weight (kg, median)	54.4	51.8
(min–max)	(32.8–93.5)	(41–82.1)
BMI (kg/m^2, median)	19.9	20
(min–max)	(15.7–33.6)	(16,3–27.9)
Pubarche onset (year, median)	13	11.5
(min–max)	(11–14)	(9–14)

BMI: Body mass index, Min: Minimum, Max: Maximum.

Table 4. Characteristics of patients according to reaching target height status in adolescence.

	Reaching target height	
	Yes	No
	n = 26	n = 11
Gender		
Female	15	5
Male	11	6
Intrauterine growth		
Symmetric SGA	6	3
Asymmetric SGA	9	4
AGA	10	4
LGA	1	–

AGA: Appropriate for gestational age, LGA: Large for gestational age, SGA: Small for gestational age.

Table 5. Distribution of clinical features according to height percentiles at birth.

Birth height	<3rd p		3–97th p		p[a]
	n	%	n	%	
Intrauterine growth					
Symmetric SGA	4	66.7	4	14.8	0.052
Asymmetric SGA	2	33.3	10	37.0	
AGA	0	0.0	11	40.7	
LGA	0	0.0	2	7.4	
Reaching target height					
Yes	4	66.7	19	70.4	1.000
No	2	33.3	8	29.6	
Mother's height (mean ± 2 SD) (cm)	160.0 ± 6.2		159.4 ± 5.7		0.800
Father's height (mean ± 2 SD) (cm)	173.7 ± 8.5		170.2 ± 7.2		0.313
BMI (mean ± 2 SD) (kg/m^2)	19.5 ± 3.0		21.7 ± 4.4		0.259

[a]Fisher exact test. Due to the distribution of data, cases were analyzed as <3rd percentile and 3–97th percentile.
p: Percentile, BMI: Body mass index, SD: Standard deviation, SGA: Small for gestational age, AGA: Appropriate for gestational age, LGA: Large for gestational age.

In a French cohort that prospectively evaluated very preterm infants, the obesity rate in adolescence was 13.9% (10.5–18.3%), which is similar to that was observed in our study. This study also revealed that the change in length during the neonatal hospital stay of preterm infants was negatively associated with risk of overweight or obesity at 5 and 15 years, and that the change in BMI between discharge and 2 years was positively associated that risk (28). Unfortunately, we did not have data on patients' growth in length during their NICU stay. Another study evaluating four birth cohorts during adolescence revealed that BMI was similar between preterm and term peers, while very preterm individuals had an increased risk of overweight (29).

Growth in height accelerates during puberty, and the peak of this acceleration occurs in early puberty (Tanner stage 2) in girls and in mid–puberty (Tanner stage 4) in boys. Although the acceleration in height growth occurs earlier in girls, it ends earlier than it does in boys (30). The literature is inconsistent regarding the effect of preterm birth on timing of puberty and pubertal growth patterns. While one study assessing 129 preterms compared to terms found no differences in pubertal growth spurt (31), in another study the onset of puberty (assessed according to Tanner staging) was later for preterm girls than term controls (32). In the ESTER Preterm Birth Study, timing of pubertal growth, age at menarche or voice break were found to be similar in preterm and term individuals (33). When we evaluated our cases with respect to pubertal development, we observed that only 1 case was treated for precocious puberty, whereas the remaining cases had normal pubertal onset and development.

CONCLUSION

Although our study has certain limitations, we believe it is valuable as it includes long-term data from our country. Key limitations of

Table 6. Distribution of prenatal, natal and postnatal risk factors according to height percentiles at birth.

Birth height	<3rd p		3–97th p		p[a]
	n	%	n	%	
Prenatal risk factors					
Hypertensive disorders of pregnancy	3	60.0	6	27.3	0.295
IUGR	2	40.0	6	27.3	0.616
Mother's age >35 (year)	1	20.0	4	18.2	1.000
PROM	1	20.0	3	13.6	1.000
Presence of chronic disease	1	20.0	2	9.1	1.000
Plesantal dysfunction	1	20.0	1	4.5	0.342
Smoking/substance use	0	0.0	7	31.8	0.283
Natal risk factors					
Fetal distress	1	100.0	6	100.0	*
Placenta detachment	0	0.0	4	66.7	0.429
Postnatal risk factors					
Hyperbilirubinemia	4	66.7	18	69.2	1.000
Surfactant therapy	2	33.3	10	38.5	1.000
Hypoglycemia	2	33.3	0	0.0	**0.030**
Sepsis	2	33.3	8	30.8	1.000
Mechanical ventilation	2	33.3	10	38.5	1.000
RDS	1	16.7	11	42.3	0.370
ROP	1	16.7	5	19.2	1.000
Hydrocephalus	0	0.0	1	3.8	1.000
BPD	0	0.0	1	3.8	1.000
Intracranial hemorrhage	0	0.0	5	19.2	0.555
PDA	0	0.0	2	7.7	1.000
Apnea	0	0.0	5	19.2	0.555

[a]Fisher exact test, *p-value could not be calculated due to data distribution. BPD: Bronchopulmonary dysplasia, IUGR: Intrauterine growth retardation, PDA: Patent ductus arteriosus, PROM: Premature rupture of membranes, RDS: Respiratory distress syndrome, ROP: Retinopathy of prematurity.

Tablo 7. Distribution of prenatal, natal and postnatal risk factors according to weight percentiles at birth.

Birth weight	<3rd p		3–97th p		p[a]
	n	%	n	%	
Prenatal risk factors					
IUGR	4	40.0	6	28.6	0.685
Hypertensive disorders of pregnancy	4	40.0	6	28.6	0.685
Mother's age >35 (year)	3	30.0	3	14.3	0.284
Presence of chronic disease	2	20.0	1	4.8	0.237
PROM	2	20.0	2	9.5	0.577
Plesantal dysfunction	1	10.0	1	4.8	1.000
Smoking/substance use	1	10.0	7	33.3	0.222
Natal risk factors					
Fetal distress	2	100.0	6	100.0	*
Placenta detachment	1	50.0	3	50.0	1.000
Postnatal risk factors					
Hyperbilirubinemia	9	90.0	16	64.0	0.218
Sepsis	4	40.0	7	28.0	0.689
Apnea	3	30.0	3	12.0	0.322
Surfactant therapy	3	30.0	10	40.0	0.709
ROP	2	20.0	4	16.0	1.000
RDS	2	20.0	11	44.0	0.259
Mechanical ventilation	2	20.0	11	44.0	0.259
Hypoglycemia	1	10.0	1	4.0	1.000
PDA	1	10.0	1	4.0	1.000
Intracranial hemorrhage	1	10.0	4	16.0	1.000
Hydrocephalus	0	0.0	1	4.0	1.000
BPD	0	0.0	1	4.0	1.000

[a]Fisher exact test, *p-value could not be calculated due to data distribution. BPD: Bronchopulmonary dysplasia, IUGR: Intrauterine growth retardation, PDA: Patent ductus arteriosus, PROM: Premature rupture of membranes, RDS: Respiratory distress syndrome, ROP: Retinopathy of prematurity.

our study include the participation of high-risk preterm infants requiring intensive care and the relatively small sample size, which restrict the generalizability of the findings. Multicenter studies with larger sample sizes will contribute to the literature. A further limitation of our study is the definition of SGA restricted to neonates with birth weight below the 3rd percentile. While this approach identifies the most severely growth-restricted infants, it excludes those between the 3rd and 10th percentiles who are commonly included in the conventional definition. As a result, our findings may not be fully generalizable to the broader SGA population and should be interpreted with caution when compared to studies employing the standard <10th percentile threshold. Moreover, assessing the age of thelarche based solely on parental recall constitutes a limitation in the evaluation of pubertal development.

Our findings indicate that somatic growth in preterm infants is a dynamic, non–linear process with distinct patterns across developmental stages. Although catch–up growth in height and weight was largely achieved by school age, a subset of adolescents exhibited growth failure or excessive weight gain during puberty.

These observations underscore the clinical importance of long–term, structured follow–up for preterm infants that extends beyond early childhood. Regular anthropometric monitoring into adolescence and beyond is essential to detect deviations from expected growth trajectories. Such follow–up enables timely nutritional, endocrinological, and psychosocial interventions, which may improve long–term health outcomes and quality of life.

Ethics

Ethics Committee Approval: This study was approved by the Ethics Committee of Ministry of Health İstanbul Kanuni Sultan Süleyman Training and Research Hospital (approval number: 15929, date: 30.09.2015).

Informed Consent: Patients and their parents provided the written informed consent.

Footnotes

Authorship Contributions

Surgical and Medical Practices: K.B., S.K., Concept: K.B., S.K., Design: K.B., S.K., Data Collection or Processing: K.B., S.K., Analysis or Interpretation: K.B., S.K., Literature Search: K.B., S.K., Writing: K.B., S.K.

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

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

REFERENCES

1. Euser AM, de Wit CC, Finken MJ, Rijken M, Wit JM. Growth of preterm born children. Horm Res. 2008; 70: 319-28.
2. Raaijmakers A, Allegaert K. Catch-up growth in former preterm neonates: no time to waste. Nutrients. 2016; 8: 817.
3. Hack M, Fanaroff AA. Outcomes of children of extremely low birthweight and gestational age in the 1990s. Semin Neonatol. 2000; 5: 89-106.
4. Hack M, Klein NK, Taylor HG. Long-term developmental outcomes of low-birth-weight infants. Future Child. 1995; 5: 176-96.
5. Enhancing the outcomes of low-birth-weight, premature infants. A multisite, randomized trial. The Infant Health and Development Program. JAMA. 1990; 263: 3035-42.
6. Yeşinel S, Aldemir EY, Kavuncuoğlu S, Yeşinel S, Yıldız H. Evaluation of growth in very low birth weight preterm babies. Turk Pediatri Ars. 2014; 49: 289-98.
7. Koç Ö, Kavuncuoğlu S, Ramoğlu MG, Aldemir E, Aktalay A, Eras Z. School performance and neurodevelopment of very low birth weight preterm infants: First report from Turkey. J Child Neurol. 2016; 31: 170-6.
8. Akar S, Kavuncuoğlu S, Akın MA, Aldemir E, Demirhan A. Somatic growth features of moderate-late and ID early preterm infants at early childhood. J Behcet Uz Child Hosp. 2021; 11: 226-32.
9. Kurtoğlu S, Hatipoğlu N, Mazıcıoğlu MM, Akın MA, Çoban D, Gökoğlu S, et al. Body weight, length and head circumference at birth in a cohort of Turkish newborns. J Clin Res Pediatr Endocrinol. 2012; 4: 132-9.
10. Rogol AD. The child with tall stature and/or abnormally rapid growth. UpToDate [Internet]. Waltham (MA): UpToDate Inc.; [cited 2026 Mar 13]. Available from: https://www.uptodate.com/contents/the-child-with-tall-stature-and-or-abnormally-rapid-growth
11. Marshall WA, Tanner JM. Variations in the pattern of pubertal changes in boys. Arch Dis Child. 1970; 45: 13-23.
12. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. Arch Dis Child. 1969; 44: 291-303.
13. Neyzi O, Bundak R, Gökçay G, Günöz H, Furman A, Darendeliler F, et al. Reference values for weight, height, head circumference, and body mass index in Turkish children. J Clin Res Pediatr Endocrinol. 2015; 7: 280-293.
14. Cinaz P, Darendeliler F. Normal büyüme. In: Cinaz P, Darendeliler F, editors. Temel Çocuk Endokrinolojisi. 1st ed. İstanbul: Nobel Tıp Kitabevi; 2013. p. 35. Available from: https://www.nobelkitabevi.com/temel-cocuk-endokrinolojisi.
15. World Health Organization. BMI-for-age (5-19 years) [Internet]. Geneva: WHO; [cited 2026 Mar 13]. Available from: https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators/bmi-for-age
16. Neyzi O, Günöz H, Furman A, Bundak R, Gökçay G, Darendeliler F, et al. Türk çocuklarında vücut ağırlığı, boy uzunluğu, baş çevresi ve vücut kitle indeksi referans değerleri. Çocuk Sağlığı ve Hastalıkları Dergisi. 2008; 51: 1-14.
17. Allen MC. Outcome and follow-up of high-risk infants. In: Taeusch HW, Ballard RA, editors. Avery's diseases of the newborn. 7th ed. Philadelphia: WB Saunders; 1998. p. 413-463. Available from: https://www.sciencedirect.com/book/9780721678244/averys-diseases-of-the-newborn
18. Niklasson A, Engstrom E, Hard AL, Wikland KA, Hellstrom A. Growth in very preterm children: A longitudinal study. Pediatr Res. 2003; 54: 899-905.
19. Yoshida-Montezuma Y, Kirkwood D, Sivapathasundaram B, Keown-Stoneman CDG, de Souza RJ, To T, et al. Late preterm birth and growth trajectories during childhood: A linked retrospective cohort study. BMC Pediatr. 2023; 23: 450.
20. Vinther JL, Ekstrøm CT, Sørensen TIA, Cederkvist L, Lawlor DA, Andersen AN. Gestational age and trajectories of body mass index and height from birth through adolescence in the Danish National Birth Cohort. Sci Rep. 2023; 13: 3298.
21. Jańczewska I, Wierzba J, Jańczewska A, Szczurek-Gierczak M, Domżalska-Popadiuk I. Prematurity and low birth weight and their impact on childhood growth patterns and the risk of long-term cardiovascular sequelae. Children (Basel). 2023; 10: 1599.

22. Rocha AS, Ribeiro-Silva RC, Silva JFM, Pinto EJ, Silva NJ, Paixao ES, et al. Postnatal growth in small vulnerable newborns: A longitudinal study of 2 million Brazilians using routine register-based linked data. Am J Clin Nutr. 2024; 119: 444-55.
23. Mittendorf R, Herschel M, Williams MA, Hibbard JU, Moawad AH, Lee KS. Reducing the frequency of low birth weight in the United States. Obstet Gynecol. 1994; 83: 1056-9.
24. Lin X, Zhou L, Si S, Cheng H, Alifu X, Qiu Y, et al. Association of the comorbidity of gestational diabetes mellitus and hypertension disorders of pregnancy with birth outcomes. Front Endocrinol (Lausanne). 2024; 15: 1468820.
25. Lucas A, Fewtrell MS, Davies PS, Bishop NJ, Clough H, Cole TJ. Breastfeeding and catch-up growth in infants born small for gestational age. Acta Paediatr. 1997; 86: 564-9.
26. O'Shea TM, Register HM, Yi JX, Jensen ET, Joseph RM, Kuban KCK, et al. Growth during infancy after extremely preterm birth: Associations with later neurodevelopmental and health outcomes. J Pediatr. 2023; 252: 40-7.e5.
27. Ou-Yang MC, Sun Y, Liebowitz M, Chen CC, Fang ML, Dai W, et al. Accelerated weight gain, prematurity, and the risk of childhood obesity: A meta-analysis and systematic review. PLoS One. 2020; 15: e0232238.
28. Simon L, Hadchouel A, Arnaud C, Frondas-Chauty A, Marret S, Flamant C, et al. Growth trajectory during the first 1000 days and later overweight in very preterm infants. Arch Dis Child Fetal Neonatal Ed. 2023; 108: 149-55.
29. Vinther JL, Cadman T, Avraam D, Ekstrøm CT, Sørensen TIA, Elhakeem A, et al. Gestational age at birth and body size from infancy through adolescence: An individual participant data meta-analysis on 253,810 singletons in 16 birth cohort studies. PLoS Med. 2023; 20: 1004036.
30. Wood CL, Lane LC, Cheetham T. Puberty: Normal physiology (brief overview). Best Pract Res Clin Endocrinol Metab. 2019; 33: 101265.
31. Persson I, Ahlsson F, Ewald U, Tuvemo T, Qingyuan M, von Rosen D, et al. Influence of perinatal factors on the onset of puberty in boys and girls: Implications for interpretation of link with risk of long-term diseases. Am J Epidemiol. 1999; 150: 747-55.
32. Hui LL, Leung GM, Lam TH, Schooling CM. Premature birth and age at onset of puberty. Epidemiology. 2012; 23: 415-22.
33. Suikkanen J, Nurhonen M, Cole TJ, Paalanne M, Matinolli HM, Tikanmäki M, et al. Preterm birth and subsequent timing of pubertal growth, menarche, and voice break. Pediatr Res. 2022; 92: 199-205.

Gazi Med J 2026;37(2):203-208

DOI: http://dx.doi.org/10.12996/gmj.2026.4606

Diagnostic Comparison of PCR, Chromogenic Agar and Culture for *Vancomycin-Resistant Enterococci* in Intensive Care Units

Yoğun Bakım Ünitelerinde Vankomisine Dirençli Enterokokların Tanısında PCR, Kromojenik Agar ve Kültür Yöntemlerinin Karşılaştırılması

Mustafa Uğuz[1], Gülden Ersöz[2], Berfin Çirkin Doruk[1], Nejla Mendil Erdoğan[3]

[1]Department of Infectious Diseases and Clinical Microbiology, University of Health Sciences Türkiye, Mersin City Hospital, Mersin, Türkiye
[2]Department of Infectious Diseases and Clinical Microbiology, Mersin University, Faculty of Medicine, Mersin, Türkiye
[3]Department of Anesthesiology and Reanimation, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Ankara, Türkiye

ABSTRACT

Objective: This study aimed to compare the diagnostic performance and cost-effectiveness of classical culture, chromogenic agar, and the Smart-Cycle I-CORE real-time polymerase chain reaction (PCR) method for detecting *vancomycin-resistant enterococci (VRE)* in intensive care unit (ICU) patient and environmental samples.

Methods: In a prospective surveillance design conducted in adult medical, surgical, and general ICUs, perianal/rectal swab samples from patients hospitalized ≥48 hours and high-touch environmental surface samples were obtained. Each specimen was tested in parallel using classical culture, chromogenic agar, and real-time PCR targeting vanA/vanB. Sensitivity, specificity, positive and negative predictive values (PPV/NPV), turnaround time, and per-test costs were calculated.

Results: In patient samples, PCR, chromogenic agar, and culture achieved sensitivity/specificity of 100%/100%, 100%/80%, and 95.2%/86.4%, respectively. In environmental samples, PCR showed 100%/100%, chromogenic agar 100%/92%, and culture 94%/100%, respectively. PCR provided a markedly shorter time-to-result (~4 h), compared with chromogenic agar (~24 h) and classical culture (~48 h). vanA was the predominant genotype (≈82%), followed by vanB (≈18%). Although PCR was the most costly method, its rapid turnaround time contributed to earlier isolation decisions and to a reduction in environmental positivity rates.

Conclusion: Smart-Cycle I-CORE PCR offers the highest diagnostic accuracy and the fastest reporting among currently available surveillance methods, while chromogenic agar represents a reliable and cost-effective option. A two-step strategy—chromogenic agar for

ÖZ

Amaç: Bu araştırmada, yoğun bakım ünitesi (YBÜ) hasta ve çevresel örneklerinde *vankomisin dirençli enterokok (VRE)* saptanmasında klasik kültür, kromojenik agar ve Smart-Cycle I-CORE gerçek zamanlı polimeraz zincir reaksiyonu (PZR) yöntemlerinin tanısal performans ve maliyet etkinliklerinin karşılaştırılması amaçlanmıştır.

Yöntemler: Prospektif tasarımla dahili, cerrahi ve genel YBÜ'lerde ≥48 saat yatan yetişkin hastalardan perianal/rektal sürüntü örnekleri ile sık temaslı yüzeylerden çevresel örnekler alınmıştır. Tüm örnekler klasik kültür, kromojenik agar ve vanA/vanB genlerini hedefleyen gerçek zamanlı PZR ile eşzamanlı olarak çalışılmıştır. Duyarlılık, özgüllük, pozitif ve negatif öngörü değerleri ile sonuç alma süresi ve test başı maliyet hesaplanmıştır.

Bulgular: Hasta örneklerinde PZR'nin duyarlılığı ve özgüllüğü %100/%100, kromojenik agarın %100/%80, klasik kültürün %95,2/%86,4 olarak bulunmuştur. Çevresel örneklerde ise PZR %100/%100, kromojenik agar %100/%92, klasik kültür %94/%100 doğruluk göstermiştir. Sonuç alma süresi PZR'de ~4 saat, kromojenik agarda ~24 saat ve klasik kültürde ~48 saat olarak saptanmıştır. En sık saptanan genotip vanA (%82) olup bunu vanB (%18) izlemiştir. PZR, maliyet açısından en yüksek olmakla birlikte hızlı sonuç vermesi nedeniyle erken izolasyon kararlarına katkı sağlamıştır.

Sonuç: Smart-Cycle I-CORE PZR yöntemi, *VRE* sürveyansında en yüksek tanısal doğruluk ve en kısa sonuç süresini sağlamaktadır. Kromojenik agar ise tarama için güvenilir ve maliyet açısından uygun bir alternatiftir. "Kromojenik agar ile tarama, PZR ile doğrulama"

Cite this article as: Uğuz M, Ersöz G, Çirkin Doruk B, Mendil Erdoğan N. Diagnostic comparison of PCR, chromogenic agar and culture for *vancomycin-resistant enterococci* in intensive care units. Gazi Med J. 2026;37(2):203-208

Address for Correspondence/Yazışma Adresi: Mustafa Uğuz, Department of Infectious Diseases and Clinical Microbiology, University of Health Sciences Türkiye, Mersin City Hospital, Mersin, Türkiye
E-mail / E-posta: drmustafauguz@gmail.com
ORCID ID: orcid.org/0000-0002-3428-6137

Received/Geliş Tarihi: 12.12.2025
Accepted/Kabul Tarihi: 06.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

routine screening with PCR confirmation—balances accuracy, speed, and cost in ICU *VRE* surveillance.

Keywords: *Vancomycin-resistant enterococci*, real-time PCR, chromogenic agar, surveillance, intensive care units

ÖZ

yaklaşımı, yoğun bakım koşullarında hız, doğruluk ve maliyet dengesi açısından önerilebilir.

Anahtar Sözcükler: *Vankomisin dirençli enterokok*, gerçek zamanlı PZR, kromojenik agar, sürveyans, yoğun bakım üniteleri

INTRODUCTION

Vancomycin-resistant enterococci (VRE) have emerged as important healthcare-associated pathogens since their first description in the late 1980s. Their ability to acquire and transfer resistance determinants, survive for prolonged periods on inanimate surfaces, and rapidly disseminate under the selective pressure of broad-spectrum antibiotic use has led to endemic circulation particularly in intensive care units (ICUs) (1-3).

Enterococci are Gram-positive cocci commonly found in the gastrointestinal microbiota. *Enterococcus faecalis (E. faecalis)* and *Enterococcus faecium (E. faecium)* represent the most clinically relevant species and may persist on medical equipment and environmental surfaces, thereby contributing to nosocomial transmission. Of particular concern is the increasing vancomycin resistance observed predominantly among *E. faecium* isolates in recent years, posing a significant public health challenge worldwide (4,5).

Rapid and accurate detection of *VRE* colonization is essential for timely implementation of infection control measures, including patient isolation and environmental decontamination, to prevent hospital outbreaks. Conventional culture-based methods remain widely available and relatively inexpensive; however, they require prolonged incubation periods that may delay infection control interventions. Chromogenic agar has been introduced as a more rapid screening method allowing presumptive identification of *VRE* colonies based on color differentiation, although its diagnostic performance may vary depending on laboratory conditions. In contrast, molecular assays such as real-time polymerase chain reaction (PCR) targeting resistance genes including vanA and vanB provide high sensitivity and specificity and can substantially reduce diagnostic turnaround time (6,7).

Although the molecular epidemiology of *VRE* has evolved over the past decade, the fundamental diagnostic strategies used for *VRE* detection—culture-based screening, chromogenic agar, and PCR—remain central components of infection control programs in many healthcare settings. The data used in the present study were obtained from a prospective surveillance program conducted in the ICUs of a tertiary-care university hospital. Therefore, the primary aim of this study was to compare the diagnostic performance, turnaround time, and relative cost of classical culture, chromogenic agar, and Smart-Cycle I-CORE real-time PCR for detecting *VRE* in patient and environmental samples, to provide evidence supporting the optimization of surveillance strategies in critical care environments.

MATERIALS AND METHODS

Study Design and Ethical Approval

This prospective diagnostic validation study was conducted in the ICUs of a tertiary-care university hospital to compare classical culture, chromogenic agar, and Smart-Cycle I-CORE real-time PCR for *VRE* surveillance. The study protocol was approved by the Mersin University Clinical Research Ethics Committee (approval number: 2008/91, date: 17/10/2008) and was carried out in accordance with the Declaration of Helsinki. Written informed consent was waived because the study did not include identifiable patient data.

Setting and Sample Collection

The study was performed in the medical, surgical, and general ICUs of Mersin University Medical Faculty Hospital. Adult patients hospitalized for ≥48 hours were eligible. During the study period, 210 patient samples and 185 environmental samples were collected. Environmental swabs were obtained from high-touch surfaces in rooms of *VRE*-positive or suspected patients (e.g., bed rails, monitor keypads, nightstand surfaces).

Specimen Sampling and Transport

Patient samples consisted of perianal or rectal specimens collected with sterile flocked swabs and transported in appropriate media according to manufacturer's recommendations. Environmental sampling was performed with sterile swabs applied to frequently touched surfaces under similar transport conditions. Initial screening was conducted 48 hours after ICU admission. Weekly follow-up samples were obtained from *VRE*-positive patients. All specimens were processed within two hours of receipt.

Diagnostic Procedures

Classical culture, chromogenic agar culture, and real-time PCR were performed simultaneously on each specimen.

Classical Culture: Enterococcus-selective media (e.g., Enterococcosel/BEA) and 6.5% NaCl media were inoculated and incubated at 37 °C for 24–48 h. Identification was performed using Gram staining, catalase testing, and PYR testing. Vancomycin susceptibility was assessed phenotypically according to CLSI criteria.

ChromID *VRE* chromogenic agar (bioMérieux, France) was inoculated and incubated at 37 °C for 24 h. Color differentiation provided presumptive identification (pink–red colonies suggestive of *E. faecium*; blue–green colonies suggestive of *E. faecalis*). Presumptive *VRE* isolates were confirmed using molecular testing.

Real-time PCR was performed using the Smart-Cycle I-CORE Real-Time PCR System (Cepheid, USA) to detect the *vanA* and *vanB* genes. DNA extraction and amplification followed the manufacturer's instructions. Cycle threshold (Ct) ≤35 was interpreted as positive. Positive and negative controls were included in each run.

Quality Control and Internal Validation

Quality control included *E. faecium* ATCC 51559 (vanA positive) and *E. faecalis* ATCC 29212 (vancomycin susceptible). At least one positive and one negative control were included per assay run. Runs

with invalid controls were repeated. Daily calibration and weekly maintenance of instruments were documented.

Cost Analysis

Direct diagnostic costs were calculated per test based on institutional cost components, including reagents, consumables, personnel time (processing, analysis, reporting), and instrument utilization (energy and service allocation). Costs were expressed in Turkish Lira (”). All diagnostic methods were compared using identical unit prices.

Statistical Analysis

Statistical analysis was performed using SPSS version 25 (IBM Corp., USA). χ^2 or Fisher’s exact tests were applied to compare categorical variables. Agreement between methods was assessed using Pearson correlation analysis. Diagnostic performance metrics (sensitivity, specificity, PPV, NPV) were calculated using standard definitions. A p-value <0.05 was considered statistically significant.

RESULTS

Between May and October 2010, a total of 420 ICU patients were screened for *VRE* colonization using perianal swab specimens. Of these, 268 (63.8%) were male and 152 (36.2%) were female. *VRE* colonization was detected in 21 patients (5.0%), including 11 males (52.4%) and 10 females (47.6%). The median age of colonized patients was 56.1 ± 5.07 years (range, 30–80). Mean age was 61.8 ± 12.3 years in males and 49.0 ± 30.8 years in females.

Intensive Care Unit Distribution and Clinical Characteristics

Among all screened patients, 146 (34.7%) were admitted to medical ICUs, 137 (32.6%) to surgical ICUs, and 137 (32.6%) to general ICUs. Of the 21 *VRE*-positive cases, 3 (14.3%) were identified in medical ICUs, 9 (42.9%) in surgical ICUs, and 9 (42.9%) in general ICUs.

Chronic comorbidities were present in 18 patients (85.7%); 3 patients (14.3%) had no comorbidities. Of these, at least two comorbidities were identified in 7 patients (33.3%). Diabetes mellitus was the most frequent underlying disease (33.3%), followed by solid organ malignancies (23.8%) and cardiovascular diseases (19.0%).

Total parenteral nutrition was administered to 18 patients (85.7%), and enteral nutrition was used concomitantly in 14 patients (66.7%). Previous surgery was recorded in 16 patients (76.2%), most commonly gastrointestinal procedures (42.8%), followed by genitourinary procedures (19.0%), cranial procedures (9.5%), and thoracic procedures (4.0%).

Central venous catheterization was documented in 20 patients (95.2%) at the time of *VRE* detection. Diarrhea was present in 6 patients (28.6%). Immunosuppression was observed in 9 patients (42.9%).

Within the preceding three months, 10 patients (47.6%) had been hospitalized and 4 (19.0%) had stayed in an ICU. The mean ICU length of stay among *VRE*-positive patients was 17.4 ± 11.0 days (range, 6–50) (Table 1).

Table 1. Demographic and clinical characteristics of *VRE*-positive patients (n = 21).

Parameter	n (%)	Details / mean ± SD
Age (years)	–	56.1 ± 5.07 (range, 30–80)
Sex (male/female)	11 (52.4) / 10 (47.6)	–
ICU type	–	Medical: 3 (14.3%), surgical: 9 (42.9%), general: 9 (42.9%)
Total patients screened	420	–
VRE positivity	21 (5.0)	–
Comorbidities	18 (85.7)	DM: 7 (33.3%), malignancy: 5 (23.8%), cardiovascular: 4 (19.0)
Total parenteral nutrition	18 (85.7)	–
Enteral feeding	14 (66.7)	–
Previous surgery	16 (76.2)	GI: 9 (42.8%), GU: 4 (19%), cranial: 2 (9.5%), Thoracic: 1 (4%)
Central venous catheter	20 (95.2)	–
Diarrhea	6 (28.6)	–
Immunosuppression	9 (42.9)	–
Previous hospitalization (last 3 months)	10 (47.6)	Prior ICU stay: 4 (19.0)
ICU length of stay (days)	–	17.4 ± 11.0 (range, 6–50)
Common infection types*	–	Pneumonia (n = 11), surgical site (n = 9), bacteremia (n = 8), UTI (n = 3)
VRE species	–	*E. faecium* 81%, *E. faecalis* 19%
Environmental positivity	21/113 (18.3)	Bed: 9, nightstand: 5, monitor/pump: 5, Cart: 2

*Infection types listed in this table (e.g., pneumonia, surgical site infection, bacteremia, and urinary tract infection) represent the primary clinical conditions leading to ICU admission and are not necessarily infections caused by *VRE*.
GI: Gastrointestinal surgery, GU: Genitourinary surgery, ICU: Intensive care unit, DM: Diabetes mellitus, *VRE: Vancomycin-resistant enterococci, SD: Standard deviation.*

Antibiotic Use

All *VRE*-positive patients were receiving parenteral antimicrobial therapy at the time of detection. The most frequently administered agents were carbapenems (19%), carbapenem plus glycopeptide combinations (23.8%), ampicillin–sulbactam (9.5%), carbapenem plus aminoglycoside combinations (9.5%),and ampicillin–sulbactam plus metronidazole (9.5%), whereas glycopeptide monotherapy was used in 9.5% of cases. Among patients receiving glycopeptides, six (85.7%) were treated with teicoplanin and one (14.3%) was treated with vancomycin. Antimicrobial treatment indications included bacteremia (n = 8), urinary tract infection, and other suspected or confirmed infections.

E-test results showed that all isolates exhibited high-level vancomycin resistance (MIC >256 µg/mL), while teicoplanin resistance was observed in 75% of isolates (MIC >16 µg/mL).

Environmental Samples

Among 113 environmental samples obtained from rooms of *VRE*-positive patients, 21 (18.3%) yielded *VRE*. Environmental colonization persisted during follow-up in 14 patients and lasted for a mean of 7 ± 5.4 days (range, 7–21). No environmental contamination was detected at the end of the study. The most frequently contaminated surfaces were bed rails (n = 9), nightstands (n = 5), and monitor/pump surfaces (n = 5), followed by clinical carts (n = 2).

Only one patient with perianal *VRE* colonization developed clinical infection; therefore, separate risk factor analyses for colonization versus infection could not be performed. Environmental positivity was not significantly associated with the duration of colonization (p = 0.6). The infection types listed in Table 1 (e.g., pneumonia, surgical site infection, bacteremia, and urinary tract infection) represent the primary clinical diagnoses leading to ICU admission, rather than infections caused by *VRE*.

Table 2. Distribution of *VRE* species.

Species	n	%
E. faecalis	4	19.0
E. faecium	17	81.0
Total	21	100.0

E. faecium: Enterococcus faecium, *E. faecalis: Enterococcus faecalis, VRE: Vancomycin-resistant enterococci.*

Species Distribution and Antimicrobial Susceptibility

Of the 21 *VRE* isolates, 81% were *E. faecium* and 19% were *E. faecalis* (Table 2). All isolates demonstrated high-level gentamicin resistance.

Diagnostic Performance of the Methods

In environmental samples, the turnaround time of PCR was 1 hour, compared with 38.8 ± 6.6 hours for Enterococcosel agar and 24.5 ± 5.9 hours for chromogenic agar. No significant difference was observed between Enterococcosel agar and chromogenic agar for environmental specimens (p > 0.05).

In patient samples, the turnaround time of Enterococcosel agar was 60 ± 4.0 hours, whereas chromogenic agar required a significantly shorter duration of 26.8 ± 3.2 hours (p < 0.038).

When PCR was considered the reference standard, patient samples yielded sensitivities, specificities, PPVs, and NPVs of 100% for PCR; 95.2%, 84.6%, 92%, and 96%, respectively, for Enterococcosel agar; and 100%, 80%, 94%, and 100%, respectively, for chromogenic agar (Table 3).

In environmental samples, with PCR as the reference standard, Enterococcosel agar showed a sensitivity of 94%, specificity of 100%, PPV of 100%, and NPV of 80%, whereas chromogenic agar showed a sensitivity of 100%, specificity of 92%, PPV of 87%, and NPV of 100% (Table 4).

Cost Analysis

The average direct diagnostic cost per test was calculated to be 181 ” for PCR, 113.15 ” for chromogenic agar, and 182.4 ” for classical culture. Although PCR had a higher direct cost per test than chromogenic agar, it provided substantially faster results. This shorter turnaround time may facilitate earlier infection control interventions such as patient isolation and environmental decontamination. From a practical perspective, chromogenic agar may serve as an effective screening method due to its lower cost, whereas PCR may be used as a confirmatory test when rapid and highly accurate detection is required.

Table 3. Diagnostic performance and cost per patient sample.

Method	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Cost (”/sample)
PCR	100	100	100	100	181
Enterococcosel agar	95.2	86.4	92	96	182.4
Chromogenic agar	100	80	94	100	113.15

PCR: Polymerase chain reaction, PPV: Positive predictive value, NPV: Negative predictive value.

Table 4. Diagnostic performance and cost per environmental sample.

Method	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Cost (”/sample)
PCR	100	100	100	100	181
Enterococcosel agar	94	100	100	80	182.4
Chromogenic agar	100	92	87	100	113.15

PCR: Polymerase chain reaction, PPV: Positive predictive value, NPV: Negative predictive value.

DISCUSSION

This study compared the diagnostic performance, turnaround time, and cost of three commonly used modalities—classical culture, chromogenic agar, and Smart-Cycle I-CORE real-time PCR—for detecting *VRE* colonization in ICUs. Our findings demonstrated that PCR achieved the highest diagnostic accuracy, whereas chromogenic agar emerged as a practical and cost-effective alternative for routine surveillance due to its ease of interpretation and lower implementation costs.

An important consideration when interpreting the findings of this study is the time period during which the data were collected. Although the surveillance was conducted in 2010, the diagnostic approaches evaluated in this study—classical culture, chromogenic agar, and PCR targeting the *vanA/vanB* genes—remain fundamental tools in current clinical microbiology laboratories. Therefore, the results primarily reflect differences in diagnostic performance and turnaround time rather than temporal changes in *VRE* epidemiology.

E. faecium and *E. faecalis* are well-recognized causes of healthcare-associated infections, particularly among critically ill patients with risk factors such as broad-spectrum antibiotic exposure, invasive procedures, and immunosuppression (5-8). The predominance of *VRE* in the medical and surgical ICUs of our cohort is consistent with existing epidemiological trends, and the higher freq*uency* of *E. faecium* isolates is in line with increasingly reported resistance profiles in the literature.

Classical culture remains widely accessible and inexpensive; however, its prolonged incubation period, often exceeding 48 hours, may delay the implementation of infection control interventions. Chromogenic agar offers more rapid presumptive identification based on color differentiation and can be used as an efficient rule-out tool given its high negative predictive value (9,10). In the present study, chromogenic agar performed adequately as a screening method, representing a feasible option for large-scale surveillance.

Real-time PCR provides the major advantage of directly detecting of *vanA* and *vanB* genes, enabling rapid isolation measures. Previous studies have shown that molecular-based algorithms reduce diagnostic turnaround time from 24–48 hours to 3–5 hours compared with conventional culture, contributing to earlier interruption of the transmission chain (11-13). In the current study, PCR consistently returned results within 4 hours, facilitating patient isolation within approximately 6 hours, thereby supporting timely infection prevention efforts.

Environmental persistence of *VRE* on dry surfaces for prolonged durations highlights the importance of environmental decontamination in prevention strategies (14-16). In our investigation, the environmental positivity rate decreased from 6.7% to 1.2% after the implementation of environmental and isolation precautions, indicating the effectiveness of targeted cleaning interventions.

Although PCR was more expensive than chromogenic agar, it provided substantially faster results (17-19). Consequently, a two-step diagnostic approach—initial screening by chromogenic agar followed by PCR confirmation—may offer an optimal balance between cost and diagnostic reliability, depending on laboratory capacity and workload.

The present findings align with the 2023 guideline by the Turkish Ministry of Health on the prevention of *VRE*, which emphasizes early diagnosis, effective isolation, and meticulous environmental decontamination as key components in controlling endemic *VRE* transmission.

Study Limitations

This study has several limitations, including its single-center design and a relatively small number of isolates. Nevertheless, our results provide important insights into the applicability and cost-effectiveness of molecular diagnostic methods in ICU settings in Türkiye. Larger multicenter studies are warranted to further validate these findings and guide national infection control strategies.

This study has several limitations. First, it was conducted in a single tertiary-care center with a relatively small number of isolates, which may limit the generalizability of the findings. Second, the differentiation between colonization and infection could not be fully evaluated because of the small number of cases of infection. Third, molecular characterization was limited to detection of vanA and vanB, without further genotypic analysis. Therefore, multicenter prospective studies are required to validate our results and better define the epidemiological characteristics of *VRE* circulation in intensive care settings. Another limitation of this study is that the dataset was collected in 2010. Although the fundamental diagnostic methods remain unchanged, the molecular epidemiology of *VRE* may have evolved over time.

CONCLUSION

This study presents a comparative analysis of three diagnostic methods used for the detection of *VRE* in ICUs. Our findings indicate that the Smart-Cycle I-CORE PCR system provides the highest diagnostic accuracy and the shortest turnaround time, whereas chromogenic agar represents an appropriate and practical option for active surveillance due to its lower cost and ease of use.

Early detection, which enabling which enables the rapid implementation of isolation measures, offers a major advantage in preventing nosocomial transmission. Given the high mortality risk predominantly associated with *E. faecium* strains, rapid diagnosis and appropriate isolation remain essential components of infection control strategies.

Although PCR appears more expensive in the short term, it has the potential to reduce additional costs associated with delayed isolation in the long run. Therefore, a two-step diagnostic strategy—screening by chromogenic agar followed by PCR confirmation—may provide an optimal balance among speed, accuracy, and cost, depending on institutional resources.

The selection of diagnostic strategies that ensure an appropriate balance between diagnostic accuracy, rapidity, and cost-effectiveness is crucial for infection prevention, patient safety, and effective resource management in intensive care settings.

Ethics

Ethics Committee Approval: The study protocol was approved by the Mersin University Clinical Research Ethics Committee (approval number: 2008/91, date: 17/10/2008) and was carried out in accordance with the principles of the Declaration of Helsinki.

Informed Consent: Written informed consent was waived because the study was observational and did not involve any identifiable personal data.

Footnotes

Authorship Contributions

Concept: M.U., G.E., Design: M.U., G.E., B.Ç.D., N.M.E., Data Collection or Processing: M.U., G.E., Analysis or Interpretation: M.U., G.E., B.Ç.D., N.M.E., Literature Search: M.U., G.E., B.Ç.D., N.M.E., Writing: M.U., G.E., B.Ç.D., N.M.E.

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

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

REFERENCES

1. Pezzani MD, Arieti F, Rajendran NB, Barana B, Cappelli E, De Rui ME, et al. Frequency of bloodstream infections caused by six key antibiotic-resistant pathogens for prioritization of research and discovery of new therapies in Europe: a systematic review. Clin Microbiol Infect. 2024; 30 Suppl 1: S4-13.
2. Fox E, Ha D, Bounthavong M, Meng L, Mui E, Holubar M, et al. Risk factors and outcomes associated with persistent vancomycin resistant Enterococcal Bacteremia. BMC Infect Dis. 2022; 22: 855.
3. Eichel VM, Last K, Brühwasser C, von Baum H, Dettenkofer M, Götting T, et al. Epidemiology and outcomes of vancomycin-resistant enterococcus infections: a systematic review and meta-analysis. J Hosp Infect. 2023; 141: 119-28.
4. Hornuss D, Göpel S, Walker SV, Tobys D, Häcker G, Seifert H, et al. Epidemiological trends and susceptibility patterns of bloodstream infections caused by Enterococcus spp. in six German university hospitals: a prospectively evaluated multicentre cohort study from 2016 to 2020 of the R-Net study group. Infection. 2024; 52: 1995-2004.
5. Pan SC, Wang JT, Chen YC, Chang YY, Chen ML, Chang SC. Incidence of and risk factors for infection or colonization of vancomycin-resistant enterococci in patients in the intensive care unit. PLoS One. 2012; 7: e47297.
6. Seo JY, Kim PW, Lee JH, Song JH, Peck KR, Chung DR, et al. Evaluation of PCR-based screening for vancomycin-resistant enterococci compared with a chromogenic agar-based culture method. J Med Microbiol. 2011; 60: 945-9.
7. He YH, Ruan GJ, Hao H, Xue F, Ma YK, Zhu SN, et al. Real-time PCR for the rapid detection of vanA, vanB and vanM genes. J Microbiol Immunol Infect. 2020; 53: 746-50.
8. Hristova PM, Marinova-Bulgaranova TV, Strateva TV, et al. Risk factors for gut colonization with vancomycin-resistant enterococci among Bulgarian critically ill patients. Gut Pathog. 2023; 15: 37.
9. Boschert AL, Arndt F, Hamprecht A, Wolke M, Walker SV. Comparison of Five Different Selective Agar for the Detection of Vancomycin-Resistant Enterococcus faecium. Antibiotics (Basel). 2023; 12: 666.
10. Tan C, Linkenheld-Struk A, Williams V, Kozak R, Dhabaan G, Maze Dit Mieusement L, et al. The role of the environment in transmission of vancomycin-resistant Enterococcus: a proof-of-concept study. Antimicrob Steward Healthc Epidemiol. 2022; 2: e178.
11. Marner ES, Wolk DM, Carr J, Hewitt C, Dominguez LL, Kovacs T, et al. Diagnostic accuracy of the Cepheid GeneXpert vanA/vanB assay ver. 1.0 to detect the vanA and vanB vancomycin resistance genes in Enterococcus from perianal specimens. Diagn Microbiol Infect Dis. 2011; 69: 382-9.
12. Ayaydın Z, Azarkan AB, Bilik ÖA, Tekin AC. Evaluation of screening results for vancomycin-resistant enterococci: three-year surveillance. Iran J Public Health. 2024; 53: 1746-53.
13. Salvador BC, Lucchetta RC, Sarti FM, Ferreira FF, Tuesta EF, Riveros BS, et al. Cost-effectiveness of molecular method diagnostic for rapid detection of antibiotic-resistant bacteria. Value Health Reg Issues. 2022; 27: 12-20.
14. Mareković I, Markanović M, Lešin J, Ćorić M. Vancomycin-resistant enterococci: current understandings of resistance in relation to transmission and preventive strategies. Pathogens. 2024; 13: 966.
15. Kramer A, Schwebke I, Kampf G. How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. BMC Infect Dis. 2006; 6: 130.
16. Knobling B, Franke G, Belmar Campos C, Büttner H, Christner M, Klupp EM, et al. Tolerance of clinical vancomycin-resistant Enterococcus faecium isolates against UV-C light from a mobile source. Antimicrob Resist Infect Control. 2023; 12: 63.
17. Lee TA, Hacek DM, Stroupe KT, Collins SM, Peterson LR. Three surveillance strategies for vancomycin-resistant enterococci in hospitalized patients: detection of colonization efficiency and a cost-effectiveness model. Infect Control Hosp Epidemiol. 2005; 26: 39-46.
18. Akinbobola A. Assessment of surface contamination by vancomycin resistant Enterococcus species in selected hospitals in the Akoko region of Ondo State, Nigeria. Int J Pathogen Res. 2025; 14: 107-12.
19. Zhou X, Willems RJL, Friedrich AW, Rossen JWA, Bathoorn E. Enterococcus faecium: from microbiological insights to practical recommendations for infection control and diagnostics. Antimicrob Resist Infect Control. 2020; 9: 130.

Gazi Med J 2026;37(2):209-214

DOI: http://dx.doi.org/10.12996/gmj.2026.4624

Subclinical Right Ventricular and Atrial Dysfunction in Heart Failure with Preserved Ejection Fraction: A Speckle Tracking Comparison

Korunmuş Ejeksiyon Fraksiyonlu Kalp Yetersizliğinde Subklinik Sağ Ventrikül ve Sağ Atriyum Disfonksiyonu: Benek Takibi Yöntemi ile Karşılaştırmalı Bir Değerlendirme

Yakup Yalçın[1], Özden Seçkin[2], Serkan Ünlü[2], Betül Ayça Yamak[3], Mustafa Cemri[2], Mehmet Rıdvan Yalçın[2]

[1]Department of Cardiology, Hakkâri State Hospital, Hakkâri, Türkiye
[2]Department of Cardiology, Gazi University Hospital, Ankara, Türkiye
[3]Department of Cardiology, Hopa State Hospital, Artvin, Türkiye

ABSTRACT

Objective: Functional involvement of the right heart in heart failure (HF) with preserved ejection fraction (HFpEF) has gained increasing attention; however, early functional changes may not be adequately detected using standard echocardiographic techniques. The purpose of this work was to assess ventricular and atrial performance in individuals with HFpEF by combining conventional echocardiographic evaluation with deformation-based strain analysis.

Methods: In this cross-sectional study, patients with HFpEF hospitalized for acute HF between January 2020 and January 2022 were evaluated. All patients had a left ventricular (LV) EF of ≥50% and underwent comprehensive transthoracic echocardiographic assessment. The structure and function of the right ventricular (RV) and right atrium were analysed using conventional echocardiographic measurements, Doppler and tissue Doppler indices, and speckle-tracking–derived strain parameters. Findings were compared with those of an age- and sex-matched control group.

Results: LV dimensions and systolic function were similar between the HFpEF and control groups. In contrast, patients with HFpEF exhibited significantly larger RV diastolic and right atrial end-systolic areas. Conventional RV functional parameters demonstrated limited sensitivity for detecting early dysfunction. Two-dimensional (2D) speckle-tracking echocardiography revealed significantly reduced RV global longitudinal strain and free-wall strain in HFpEF patients. Additionally, right atrial reservoir and contraction strain values were significantly lower, whereas conduit strain values did not differ between groups.

ÖZ

Amaç: Korunmuş ejeksiyon fraksiyonlu kalp yetersizliğinde (HFpEF) sağ kalp tutulumu giderek daha fazla önem kazanmaktadır; ancak erken dönemde ortaya çıkan fonksiyonel bozukluklar, konvansiyonel ekokardiyografik parametrelerle saptanamayabilir. Bu çalışmada, HFpEF tanılı hastalarda sağ ventrikül ve sağ atriyum fonksiyonlarının, konvansiyonel ekokardiyografi ve iki boyutlu benek takibi yöntemiyle elde edilen miyokardiyal deformasyon parametreleri kullanılarak değerlendirilmesi amaçlanmıştır.

Yöntemler: Bu kesitsel çalışmada, Ocak 2020–Ocak 2022 tarihleri arasında dekompanse kalp yetersizliği nedeniyle hastaneye yatırılan HFpEF tanılı hastalar değerlendirildi. Tüm hastaların sol ventrikül ejeksiyon fraksiyonu ≥%50 idi ve kapsamlı transtorasik ekokardiyografik inceleme yapıldı. Sağ ventrikül ve sağ atriyumun yapısal ve fonksiyonel özellikleri; konvansiyonel ekokardiyografik ölçümler, Doppler ve doku Doppler parametreleri ile iki boyutlu benek takibi yöntemiyle elde edilen strain değerleri kullanılarak analiz edildi. Bulgular, yaş ve cinsiyet açısından benzer özelliklere sahip kontrol grubu ile karşılaştırıldı.

Bulgular: Sol ventrikül boyutları ve sistolik fonksiyon parametreleri HFpEF ve kontrol grupları arasında benzerdi. Buna karşılık, HFpEF grubunda sağ ventrikül diyastolik alanı ve sağ atriyum sistol sonu alanı anlamlı olarak daha büyüktü. Konvansiyonel sağ ventrikül fonksiyon parametrelerinin erken fonksiyonel bozukluğu saptamadaki duyarlılığı sınırlıydı. İki boyutlu (2B) benek takibi ekokardiyografisi ile HFpEF hastalarında sağ ventrikül global longitudinal strain ve serbest duvar strain değerlerinin anlamlı olarak azaldığı saptandı. Ayrıca sağ atriyum

Cite this article as: Yalçın Y, Seçkin Ö, Ünlü S, Yamak BA, Cemri M, Yalçın MR. Subclinical right ventricular and atrial dysfunction in heart failure with preserved ejection fraction: a speckle tracking comparison. Gazi Med J. 2026;37(2):209-214

Address for Correspondence/Yazışma Adresi: Yakup Yalçın, Department of Cardiology, Hakkâri State Hospital, Hakkâri, Türkiye
E-mail / E-posta: yakupyalcn@gmail.com
ORCID ID: orcid.org/0000-0002-8034-5847

Received/Geliş Tarihi: 12.01.2026
Accepted/Kabul Tarihi: 09.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

Conclusion: 2D speckle tracking echocardiography enables detection of subclinical RV and right atrial dysfunction in patients with HFpEF, which may not be apparent on conventional echocardiographic assessment. Incorporating myocardial deformation analysis into routine evaluation may provide incremental insight into right heart involvement in this patient population.

Keywords: Heart failure, right ventricular, right atrium, preserved ejection fraction

ÖZ

rezervuar ve kontraksiyon fazı strain değerleri düşük bulunurken, kondüit fazı strain değerlerinde gruplar arasında anlamlı fark izlenmedi.

Sonuç: 2B benek takibi ekokardiyografisi, HFpEF hastalarında konvansiyonel ekokardiyografik değerlendirme ile saptanamayabilecek subklinik sağ ventrikül ve sağ atriyum disfonksiyonunu ortaya koyabilmektedir. Miyokardiyal deformasyon analizinin rutin değerlendirmeye eklenmesi, bu hasta grubunda sağ kalp tutulumunun daha kapsamlı biçimde değerlendirilmesine katkı sağlayabilir.

Anahtar Sözcükler: Kalp yetersizliği, sağ ventrikül, sağ atriyum, korunmuş ejeksiyon fraksiyonu

INTRODUCTION

Heart failure (HF) is a condition in which characteristic clinical manifestations arise as a consequence of underlying disturbances in cardiac structure and/or function. Epidemiological evidence shows that the frequency of HF increases substantially with advancing age, affecting more than 10% of individuals over 70 years. In developed countries, this age-related increase continues to make a major contribution to the overall burden of cardiovascular disease (1,2).

While the prognostic and clinical relevance of right ventricular (RV) function is well recognised in patients with HF with reduced ejection fraction (HFrEF), interest in the role of the RV has more recently expanded to include HF with preserved EF (HFpEF) (3).

Structural and functional changes affecting the left ventricular (LV), whether they are a cause or a consequence of HF, can gradually alter RV geometry and function. With disease progression, these interactions may drive RV remodeling, ultimately resulting in deterioration of RV structure and performance and, in advanced stages, the development of RV failure (4).

Evaluation of RV function using conventional echocardiographic approaches remains challenging, primarily because of the ventricle's intricate structural characteristics and non-uniform geometry. Although cardiac magnetic resonance imaging (cMRI) is considered a reference technique for detailed evaluation of RV structure and function, its routine use in daily clinical practice remains constrained by limited availability and practical cconsiderations (5).

Although LV diastolic dysfunction forms the cornerstone of HFpEF diagnosis, increasing evidence suggests that RV and right atrial dysfunction may influence symptom severity and haemodynamic burden, and may potentially affect clinical outcomes. However, subclinical right-sided involvement may remain undetected using conventional echocardiographic parameters. Therefore, advanced myocardial deformation imaging may provide additional mechanistic insight beyond diagnostic confirmation.

MATERIALS AND METHODS

Study Design and Population

A cross-sectional cohort design was used to investigate adults with HFpEF who were hospitalized for acute HF. Consecutive patients aged 18 years or older admitted to the Cardiology Department of Gazi University between January 2020 and January 2022 diagnosed with acute HF were screened for eligibility. Patients who underwent standard TTE examination and demonstrated an LV EF of at least 50% were included. At the time of hospital admission, all enrolled participants were in New York Heart Association class III or IV.

Patients with prosthetic heart valves or insufficient echocardiographic image quality were not included in the analysis. Before enrolment, all participants were informed about the study procedures and provided written consent to take part. Clinical evaluations and therapeutic interventions were performed in accordance with current scientific evidence and established clinical practice guidelines.

HFpEF was identified based on established diagnostic criteria. In this framework, HFpEF was defined by the presence of typical HF symptoms and signs, LV EF of at least 50%, and objective findings of structural and/or functional abnormalities consistent with LV diastolic dysfunction (6).

This study was approved by the Ethics Committee of Gazi University (approval number: 216, date: 27.12.2021).

Echocardiographic Imaging Protocol and Speckle Tracking Analysis

All TTE studies were performed by an experienced cardiologist with a Vivid E95 ultrasound system. All echocardiographic examinations were performed during hospitalization for acute HF, after initial clinical and haemodynamic stabilization. Image acquisition was performed to include a minimum of three consecutive cardiac cycles, and all recordings were subsequently analysed offline using EchoPAC software (GE Healthcare, Chicago, IL, USA).

Measurements of LV end-systolic and end-diastolic dimensions, chamber volumes, and LV EF, as well as left atrial diameters and area, were obtained in accordance with current TTE guidelines.

RV size was characterized by linear measurements and by quantification of systolic and diastolic areas obtained from an RV-focused apical four-chamber (A4C) view. The RV fractional area change (RVFAC) was derived by calculating the difference between systolic and diastolic areas, normalizing the resulting measurement to the diastolic area, and reporting the results as percentages. Tricuspid annular plane systolic excursion (TAPSE) was determined in the A4C view using M-mode imaging, with the cursor positioned at the lateral tricuspid annulus.

The myocardial performance index (MPI) was derived as the ratio of total isovolumetric time to ejection time using both pulsed-wave Doppler and tissue Doppler techniques. Doppler TTE evaluation included assessment of the RV outflow tract velocity and of tricuspid inflow velocities obtained from the A4C view. Tissue Doppler measurements were performed at the lateral tricuspid annulus in

the A4C view. With the sample volume positioned between the tricuspid valve leaflets, E- and A-wave velocities, the E/A ratio, and deceleration time were recorded. Systolic annular velocity (S') and diastolic annular velocities were also measured at the lateral tricuspid annulus using tissue Doppler imaging. In addition, peak systolic velocity during the isovolumetric contraction phase was assessed at the same site.

The longitudinal and transverse dimensions of the right atrium (RA) were assessed from the A4C view. Longitudinal measurements were recorded along the axis extending from the centre of the tricuspid valve to the midpoint of the superior RA wall, whereas transverse measurements were performed at the mid-atrial level along a line drawn from the lateral RA wall toward the interventricular septum. The RA area was measured in the A4C view during ventricular systole, corresponding to the phase of maximal atrial size, with careful exclusion of the systemic venous inflow sites and atrial appendage. The Inferior vena cava diameter and its respiratory variability were evaluated using the subcostal approach. Measurements of the RV outflow tract were obtained at end-diastole from the parasternal short-axis view, and the RV septal wall thickness was measured during diastole from the parasternal long-axis view.

For speckle-tracking analysis, dedicated apical views optimized for visualization of the RV and RA were acquired. To assess RV myocardial deformation, at least 15 reference points were manually positioned along the endocardial border, extending from the lateral tricuspid annulus to the septal tricuspid annulus, thereby defining the endocardial contour. Following identification of the myocardial region of interest, strain analysis was carried out with dedicated speckle tracking software. RV global longitudinal strain (RVGLS) and RV free wall longitudinal strain (RVFWS) were subsequently calculated. For RA strain analysis, tracking points were aligned with the R waves on the electrocardiogram to ensure a consistent temporal reference.

Statistical Analysis

Statistical analyses were performed using IBM SPSS software version 25.0 (IBM Corp., Armonk, NY, USA). Continuous variables were reported as mean ± standard deviation or as median with minimum and maximum values, as appropriate. Categorical data were presented as counts and percentages. Between-group comparisons of continuous variables were performed using the Student's t-test, whereas categorical variables were compared using the chi-square test. Statistical significance was defined as a p-value ≤0.05.

RESULTS

A total of 94 patients with HFpEF were initially assessed for eligibility. Nineteen individuals were excluded because atrial fibrillation was present, and an additional 15 were excluded due to inadequate echocardiographic image quality for speckle-tracking analysis. Consequently, the final study cohort comprised 60 patients with HFpEF, with a mean age of 62.2 ± 14.3 years, 60% of whom were female.

Baseline demographic variables, comorbidities, and conventional echocardiographic parameters were comparable between the HFpEF and control groups (Table 1). Hypertension was the most prevalent comorbidity in the HFpEF cohort (96.7%), followed by hyperlipidaemia (61.7%) and diabetes mellitus (53.3%). LV dimensions, volumes, and indices of systolic performance showed no meaningful differences between the two groups (Table 2). Measures of left atrial size and volume were also similar. By contrast, the evaluation of the right-sided cardiac chambers demonstrated clear between-group differences.

Patients in the HFpEF group had significantly larger RV diastolic and RA end-systolic areas versus controls with p-values ≤ 0.002 for both comparisons (Table 3). Doppler-derived parameters demonstrated higher RV E and A velocities, an increased E/A ratio, and a higher MPI

Table 1. Patient demographics and baseline echocardiographic measurements.

Variables	HFpEF (n = 60)	Control (n = 60)	p-value
Age (year)	62.2 ± 14.3	61.6 ± 13.7	0.81
Sex (F/M)	36 (60%)/24(40%)	34 (57%)/26 (43.3%)	0.71
BMI (kg/m^2)	29.9 ± 4.9	28.8 ± 5.2	0.24
CAD (n, %)	9 (15)	10 (16.7)	0.83
DM (n, %)	32 (53.3)	29 (48.3)	0.58
HT (n, %)	58 (96.7)	55 (91.7)	0.24
HPL (n, %)	37 (61.7)	33 (55)	0.46
CKD (n, %)	11 (18.3)	7 (8.3)	0.31
Septum thickness (cm)	1.3 ± 0.4	1.2 ± 0.4	0.17
Posterior wall thickness (cm)	1.20 ± 0.4	1.1 ± 0.3	0.12
Aorta-sinotubular junction (cm)	3.4 ± 0.7	3.3 ± 0.8	0.47
Ascending aorta (cm)	3.9 ± 0.8	3.8 ± 0.7	0.46
LV mass index (g/m^2)	115 ± 31.6	109.9 ± 27.1	0.34

BMI: Body mass index, CAD: Coronary artery disease, CKD: Chronic kidney disease, DM: Diabetes mellitus, HFpEF: Heart failure with preserved ejection fraction, HPL: Hyperlipidaemia, HT: Hypertension, LV: Left ventricular, F: Female, M: Male.

in patients with HFpEF. Tissue Doppler analysis showed significantly lower E', A', and S' velocities and a markedly elevated E/E' ratio in the HFpEF group (all $p < 0.05$). TAPSE excursion was significantly reduced in HFpEF patients, whereas RVFAC did not differ significantly between the groups.

Assessment with speckle-tracking echocardiography demonstrated impaired RV deformation in patients with HFpEF. Both RVGLS and RVFWS were significantly lower in the HFpEF group compared with controls ($p = 0.015$ and $p = 0.003$, respectively; Table 4). Furthermore, RA reservoir and contraction-phase strain values were significantly reduced in HFpEF patients, whereas conduit-phase strain did not differ significantly between the two groups.

Table 2. Echocardiographic measurements of the left ventricular and left atrium.

Variables	HFpEF (n=60)	Control (n=60)	p-value
Volume and size measurements			
LV end diastolic diameter (cm)	4.5 ± 0.7	4.4 ± 0.7	0.43
LV end-diastolic volume index (mL/m^2)	50.7 ± 14.1	49.6 ± 15.0	0.68
LV ejection fraction (%)	65.5 ± 5.6	64.9 ± 6.4	0.59
LA diameter (cm)	4.7 ±0.9	4.5 ± 0.8	0.20
LA area (cm^2)	19.9 ± 4.4	18.6 ± 4.7	0.12
LA volume index (mL/m^2)	33.9 ± 9.5	31.9 ±10.3	0.27

HFpEF: Heart failure with preserved ejection fraction, LA: Left atrium, LV: Left ventricular.

Table 3. Echocardiographic measurements of the right ventricular and right atrium.

Variables	HFpEF (n = 60)	Control (n = 60)	p-value
Volume and size measurements			
RV basal diameter (cm)	4.3 ± 0.9	4.1± 0.8	0.20
RV mid-cavity diameter (cm)	3.2 ± 0.6	3.1 ± 0.5	0.32
RV longitudinal diameter (cm)	6.9 ± 1.0	6.7 ± 0.9	0.25
RA longitudinal axis (cm)	4.9 ± 0.8	4.6 ± 0.9	0.056
RA short axis (cm)	4.0 ± 0.7	3.8 ± 0.7	0.12
RA end-systolic area (cm^2)	15.8 ± 4.0	13.6 ± 3.8	0.002
RV diastolic area (cm^2)	16.7 ± 4.0	14.1 ± 3.1	0.001
RV systolic area (cm^2)	10.2 ± 3.2	9.2 ± 3.1	0.08
RVFAC (%)	38.3 ± 9.2	41.2 ± 9.8	0.09
TAPSE (cm)	15.8 ± 3.8	18.1 ± 4.2	0.002
Doppler measurements			
E (m/s)	1.1 ± 0.3	0.8 ± 0.2	0.001
A (m/s)	0.6 ± 0.2	0.5 ± 0.2	0.007
E/A ratio	1.8 ± 0.5	1.6 ± 0.4	0.02
MPI	0.41 ± 0.13	0.36 ± 0.14	0.04
Deceleration time (ms)	229.5 ± 76.8	251 ± 91.4	0.16
Tissue Doppler measurements			
E' (m/s)	10 ± 4	12 ± 4.2	0.007
A' (m/s)	13 ± 4.9	15 ± 5.0	0.031
E'/A'	0.8 ± 0.3	0.7 ± 0.4	0.12
S' (m/s)	9 ± 3	11 ± 4	0.002
E/E'	11.1 ± 2.9	8.5 ± 2.4	<0.001
MPI	0.43 ± 0.11	0.36 ± 0.2	0.019
ICA (m/s^2)	2.9 ± 0.9	3.4 ± 1.4	0.022

ICA: Isovolumetric contraction acceleration, HFpEF: Heart failure with preserved ejection fraction, MPI: Myocardial performance index, RA: Right atrium, RV: Right ventricular, RVFAC: Right ventricular fractional area change, TAPSE: Tricuspid annular plane systolic excursion

Table 4. Right ventricular and right atrial strain measurements.

Variables	HFpEF (n = 60)	Control (n = 60)	p-value
RVGLS (%)	-17.9 ± 3.9	-19.7 ± 4.1	0.015
RVFWS (%)	-19.2 ± 4.5	-21.8 ± 5.4	0.003
RA reservoir phase strain (%)	34.6 ± 9.8	39.5 ± 10.9	0.01
RA contraction phase strain (%)	-13.6 ± 5.8	-16.5 ± 6.2	0.009
RA conduit phase strain (%)	20.2 ± 8.1	22.9 ± 8.5	0.07

HFpEF: Heart failure with preserved ejection fraction, RA: Right atrium, RVGLS: Right ventricular global longitudinal strain, RV: Right ventricular, RVFWS: Right ventricular free wall strain.

DISCUSSION

RV and right atrial functions were evaluated in participants with HFpEF hospitalized for acute HF. Echocardiographic assessment at admission enabled the evaluation of ventricular and atrial performance during systolic and diastolic phases, with specific emphasis on advanced myocardial deformation imaging.

HFpEF has a rising prevalence and currently constitutes a substantial proportion of the total HF population. Although LV EF is preserved, HFpEF is associated with a significant clinical burden and adverse outcomes, highlighting the importance of identifying subclinical myocardial dysfunction at an early stage (7). TTE is the most widely used approach for the assessment of right heart structure and function, largely because of its broad accessibility, noninvasive nature, and suitability for serial evaluations (8). Nevertheless, accurate assessment of RV function remains challenging, mainly because of the ventricle's intricate geometry, intrathoracic position, and close anatomical interaction with the LV.

In daily clinical practice, a range of echocardiographic parameters – such as TAPSE, MPI, Doppler-based inflow velocities, and area-derived measurements – is routinely utilised. However, these measurements may fail to adequately capture subtle myocardial abnormalities, particularly at earlier stages of the disease process (9). In our study, LV dimensions and indices of systolic performance did not show meaningful differences between the groups, which is consistent with the defining characteristics of HFpEF. Similarly, conventional RV parameters reflecting contractile and filling function demonstrated limited ability to detect early right heart dysfunction, in line with previous reports (10,11).

In contrast, myocardial deformation analysis based on speckle tracking revealed impaired ventricular mechanics in patients with HFpEF. Both RVGLS and RVFWS were significantly lower in the HFpEF group compared with controls. These results suggest that deformation-based assessment can identify early alterations in systolic mechanics that may not be captured by conventional echocardiographic measures. Consistent with previous studies, strain parameters appear to provide a more sensitive reflection of intrinsic myocardial performance and may therefore represent early indicators of RV involvement in HFpEF (12-14).

Right atrial function was also assessed using speckle-tracking-based deformation analysis. In the HFpEF cohort, the reservoir and contraction strain components of the RA were significantly reduced, indicating early involvement of atrial myocardial mechanics. By contrast, strain during the conduit phase showed no meaningful difference between the study groups, a finding that may reflect its predominantly passive role in ventricular filling. Taken together, these observations highlight the complementary role of RA strain assessment in evaluating overall right-sided cardiac performance and suggest that atrial deformation indices may provide additional insight into the pathophysiological mechanisms underlying HFpEF.

Right atrial strain reflects the triphasic function of the atrium, including reservoir, conduit, and contraction phases. Reductions in reservoir and contraction strain observed in our study may indicate impaired atrial compliance and reduced booster pump function in the setting of elevated filling pressures in HFpEF. In contrast, the relatively preserved conduit strain may be explained by its more passive dependence on ventricular relaxation and loading conditions.

Although our study was not designed to assess long-term outcomes, previous investigations have demonstrated that RV dysfunction in HFpEF is associated with increased morbidity and mortality. The presence of subclinical impairment detected by strain imaging may therefore represent an earlier stage in the spectrum of right heart involvement. Whether these deformation abnormalities translate into adverse clinical outcomes remains to be determined. Prospective studies incorporating mortality and rehospitalization endpoints are warranted.

Study Limitations

This study has several limitations. First, its cross-sectional design did not allow for assessment of clinical outcomes such as mortality or rehospitalization; therefore, the prognostic significance of the observed strain abnormalities remains uncertain. Second, the relatively modest sample size may limit generalizability. In addition, cardiac evaluation was based solely on transthoracic echocardiography, and invasive haemodynamic measurements were not performed. Prospective studies incorporating longitudinal outcome data are needed to clarify the clinical implications of right-heart deformation abnormalities in HFpEF.

CONCLUSION

Deformation-based echocardiographic analysis identified significant changes in RVGLS, RVFWS, and atrial strain parameters in individuals with HFpEF that were not apparent on conventional echocardiographic measures. These observations indicate that myocardial deformation imaging offers incremental value beyond standard echocardiographic indices when evaluating right-sided cardiac involvement. Integration of ventricular and atrial strain analysis into routine echocardiographic practice may

improve diagnostic precision and enable a more comprehensive characterization of right heart involvement in this population.

Ethics

Ethics Committee Approval: This study was approved by the Ethics Committee of Gazi University (approval number: 216, date: 27.12.2021).

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

Footnotes

Authorship Contributions

Surgical and Medical Practices: Y.Y., Ö.S., S.Ü., B.A.Y., M.C., M.R.Y., Concept: Y.Y., Ö.S., S.Ü., B.A.Y., M.C., M.R.Y., Design: Y.Y., Ö.S., S.Ü., B.A.Y., M.C., M.R.Y., Data Collection or Processing: Y.Y., Ö.S., S.Ü., B.A.Y., M.C., M.R.Y., Analysis or Interpretation: Y.Y., Ö.S., S.Ü., B.A.Y., M.C., M.R.Y., Literature Search: Y.Y., Ö.S., S.Ü., B.A.Y., M.C., M.R.Y., Writing: Y.Y., Ö.S., S.Ü., B.A.Y., M.C., M.R.Y.

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

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

REFERENCES

1. Brouwers FP, de Boer RA, van der Harst P, Voors AA, Gansevoort RT, Bakker SJ, et al. Incidence and epidemiology of new onset heart failure with preserved vs. reduced ejection fraction in a community-based cohort: 11-year follow-up of PREVEND. Eur Heart J. 2013; 34: 1424-31.
2. van Riet EE, Hoes AW, Wagenaar KP, Limburg A, Landman MA, Rutten FH. Epidemiology of heart failure: the prevalence of heart failure and ventricular dysfunction in older adults over time. A systematic review. Eur J Heart Fail. 2016; 18: 242-52.
3. Ghio S, Temporelli PL, Klersy C, Simioniuc A, Girardi B, Scelsi L, et al. Prognostic relevance of a non-invasive evaluation of right ventricular function and pulmonary artery pressure in patients with chronic heart failure. Eur J Heart Fail. 2013; 15: 408-14.
4. Mohammed SF, Hussain I, AbouEzzeddine OF, Takahama H, Kwon SH, Forfia P, et al. Right ventricular function in heart failure with preserved ejection fraction: a community-based study. Circulation. 2014; 130: 2310-20.
5. Tadic M. Multimodality evaluation of the right ventricle: an updated review. Clin Cardiol. 2015; 38: 770-6.
6. Pieske B, Tschöpe C, de Boer RA, Fraser AG, Anker SD, Donal E, et al. How to diagnose heart failure with preserved ejection fraction: the HFA-PEFF diagnostic algorithm: a consensus recommendation from the Heart Failure Association (HFA) of the European Society of Cardiology (ESC). Eur Heart J. 2019; 40: 3297-317.
7. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC), with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J. 2021; 42: 3599-726.
8. Schneider M, Binder T. Echocardiographic evaluation of the right heart. Wien Klin Wochenschr. 2018; 130: 413-20.
9. Galderisi M, Cosyns B, Edvardsen T, Cardim N, Delgado V, Di Salvo G, et al. Standardization of adult transthoracic echocardiography reporting in agreement with recent chamber quantification, diastolic function, and heart valve disease recommendations: an expert consensus document of the European Association of Cardiovascular Imaging. Eur Heart J Cardiovasc Imaging. 2017; 18: 1301-10.
10. Kjaergaard J, Snyder EM, Hassager C, Oh JK, Johnson BD. Impact of preload and afterload on global and regional right ventricular function and pressure: a quantitative echocardiography study. J Am Soc Echocardiogr. 2006; 19: 515-21.
11. Berglund F, Piña P, Herrera CJ. Right ventricle in heart failure with preserved ejection fraction. Heart. 2020; 106: 1798-804.
12. Rothschild E, Baruch G, Kaplan A, Laufer-Perl M, Beer G, Kapusta L, et al. The prognostic value of right ventricular strain and mechanical dispersion on mortality in patients with normal left ventricle function. Int J Cardiol. 2023; 372: 130-7.
13. Melenovsky V, Hwang SJ, Lin G, Redfield MM, Borlaug BA. Right heart dysfunction in heart failure with preserved ejection fraction. Eur Heart J. 2014; 35: 3452-62.
14. Obokata M, Reddy YNV, Borlaug BA. The role of echocardiography in heart failure with preserved ejection fraction: What do we want from imaging? Heart Fail Clin. 2019; 15: 241-56.

Original Investigation | Özgün Araştırma

Gazi Med J 2026;37(2):215-229

DOI: http://dx.doi.org/10.12996/gmj.2026.4594

Combining Clinical Variables with 18F-FDG-PET/CT Metrics Enhances Overall Survival Prediction in Gastric Cancer

Klinik Değişkenlerle 18F-FDG PET/CT Metriklerinin Birleştirilmesi, Mide Kanserinde Genel Sağkalım Tahminini Güçlendiriyor

Mehmet Tarık Tatoğlu[1], Aydın Acarbay[2], Ebru İbişoğlu[1], Ayşe Nur Toksöz Yıldırım[3,4], Ferda Yerdelen Tatoğlu[5], Hatice Uslu[1,4], Filiz Özülker[6,7]

[1]Department of Nuclear Medicine, Göztepe Prof. Dr. Süleyman Yalçın City Hospital, İstanbul, Türkiye
[2]Department of Medical Oncology, Göztepe Prof. Dr. Süleyman Yalçın City Hospital, İstanbul, Türkiye
[3]Department of Pathology, Göztepe Prof. Dr. Süleyman Yalçın City Hospital, İstanbul, Türkiye
[4]Department of Nuclear Medicine, İstanbul Medeniyet University, Faculty of Medicine, İstanbul, Türkiye
[5]Department of Econometrics, İstanbul University, Faculty of Economics, İstanbul, Türkiye
[6]Department of Nuclear Medicine, University of Health Sciences Türkiye, Kartal Dr. Lütfi Kırdar City Hospital, İstanbul, Türkiye
[7]Department of Nuclear Medicine, University of Health Sciences, Faculty of Medicine İstanbul, Türkiye

ABSTRACT

Objective: This study aimed to evaluate whether adding a broad set of pre-treatment [18F] fluorodeoxyglucose positron emission tomography/computed tomography (18F-FDG PET/CT) metabolic and volumetric parameters to routine clinical variables improves the prediction of overall survival (OS) in patients with gastric cancer (GC). Secondary objectives were to assess the prognostic value of blood- and spleen-normalized metabolic indices and to explore associations between PET metrics and HER2 status.

Methods: In this retrospective cohort, pre-treatment 18F-FDG PET/CT data were analyzed to extract standardized uptake value (SUV)- and volume-based PET metrics, BLR_mean, and SLR_mean. Clinical variables, pathological features, treatment details, HER2 status, and survival outcomes were obtained from institutional records. OS was calculated based on the date of initial management. Prognostic performance was evaluated using Cox models, calibration metrics, time-dependent area under the curve (AUC), and decision curve analysis (DCA). Nested models (clinical-only vs. clinical+PET) were compared to determine the incremental value.

Results: SUV- and volume-based PET metrics showed variable but directionally consistent associations with OS. Metabolic tumor volume (MTV_40) and total lesion glycolysis (TLG_40) demonstrated trends

ÖZ

Amaç: Bu çalışmanın amacı, geniş bir yelpazede elde edilen tedavi öncesi [18F] florodeoksiglukoz pozitron emisyon tomografi/bilgisayarlı tomografi (18F-FDG PET/CT) metabolik ve volumetrik parametrelerinin rutin klinik değişkenlere eklenmesinin mide kanserinde (GC) genel sağkalım (OS) öngörüsünü iyileştirip iyileştirmediğini değerlendirmektir. İkincil amaçlar; kan ve dalak normalizasyonuna dayalı metabolik indekslerin prognostik değerini incelemek ve PET metrikleri ile HER2 durumu arasındaki ilişkileri araştırmaktır.

Yöntemler: Bu retrospektif kohortta, tedavi öncesi 18F-FDG PET/CT verilerinden standart tutulum değeri (SUV)-temelli ve hacimsel PET metrikleri ile BLR_mean ve SLR_mean hesaplandı. Klinik değişkenler, patolojik özellikler, tedavi ayrıntıları, HER2 durumu ve sağkalım sonuçları kurum kayıtlarından elde edildi. OS, başlangıç tedavisinin tarihine göre hesaplandı. Prognostik performans Cox modelleri, kalibrasyon ölçütleri, zamana bağlı eğri altı alan (AUC) ve karar eğrisi analizi (DCA) kullanılarak değerlendirildi. Artımsal değeri belirlemek için klinik model ile klinik+PET içeren iç içe modeller karşılaştırıldı.

Bulgular: SUV- ve hacim temelli PET metrikleri, genel sağkalım ile değişken ancak yön olarak tutarlı ilişkiler gösterdi. MTV_40 ve TLG_40 daha kötü sonuçlara yönelik eğilim sergiledi; ancak bu etkiler çok değişkenli analizlerde tutarlı şekilde istatistiksel anlamlılığa ulaşmadı.

Cite this article as: Tatoğlu MT, Acarbay A, İbişoğlu E, Toksöz Yıldırım AN, Yerdelen Tatoğlu F, Uslu H, et al. Combining clinical variables with 18F-FDG-PET/CT metrics enhances overall survival prediction in gastric cancer. Gazi Med J. 2026;37(2):215-229

Address for Correspondence/Yazışma Adresi: Mehmet Tarık Tatoğlu, Department of Nuclear Medicine, Göztepe Prof. Dr. Süleyman Yalçın City Hospital, İstanbul, Türkiye
E-mail / E-posta: m.tatoglu@saglik.gov.tr
ORCID ID: orcid.org/0000-0002-1680-4973

Received/Geliş Tarihi: 26.11.2025
Accepted/Kabul Tarihi: 23.02.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

toward worse outcomes, although these effects did not consistently reach statistical significance in multivariable analyses. Blood- and spleen-normalized parameters (BLR_mean and SLR_mean) showed stronger effects in the PET-only model but became attenuated after adjustment for clinical covariates. Incorporating PET parameters into the clinical model modestly improved discrimination and yielded acceptable calibration. HER2-positive tumors exhibited higher metabolic activity; however, no significant interaction was observed between HER2 status and the prognostic effect of PET metrics. Across clinically relevant decision thresholds (10–40%), the combined Clinical+PET model achieved higher net benefit than the clinical model alone. The combined model demonstrated a higher net benefit at 12 and 24 months.

Conclusion: Pretreatment 18F-FDG PET/CT appears to provide additional prognostic information beyond routine clinical variables in GC. The inclusion of SUV-based, volumetric, and normalized metabolic parameters modestly improves risk stratification and is associated with favorable decision-analytic performance. These findings support integrating quantitative PET metrics into prognostic evaluation frameworks for patients undergoing management for GC.

Keywords: Gastric cancer, positron-emission tomography, fluorodeoxyglucose F18, prognosis, survival analysis

ÖZ

Kan ve dalak normalizasyonlu parametreler (BLR_mean ve SLR_mean) PET-only modelde daha güçlü etkiler gösterse de klinik değişkenlere göre ayarlama sonrası bu etkiler zayıfladı. PET parametrelerinin klinik modele eklenmesi ayırıcılığı mütevazı düzeyde artırdı ve kabul edilebilir düzeyde kalibrasyon sağladı. HER2 pozitif tümörler daha yüksek metabolik aktivite gösterdi; ancak HER2 durumunun PET metriklerinin prognostik etkisini anlamlı düzeyde değiştirdiğine dair etkileşim bulunmadı. Klinik olarak anlamlı karar eşiklerinde (10–40%), klinik+PET modeli klinik modele kıyasla daha yüksek net fayda sağladı.

Sonuç: Özetle, tedavi öncesi 18F-FDG PET/CT'nin mide kanserinde rutin klinik değişkenlere ek olarak ilave prognostik bilgi sağlayabileceği görülmektedir. SUV-temelli, hacimsel ve normalizasyonlu metabolik parametrelerin kullanımı risk sınıflandırmasını mütevazı düzeyde iyileştirmekte ve karar analitiği açısından olumlu bir performans sergilemektedir. Bu bulgular, mide kanseri yönetiminde prognostik değerlendirme çerçevelerine nicel PET metriklerinin dahil edilmesinin göz önünde bulundurulabileceğini desteklemektedir.

Anahtar Sözcükler: Mide kanseri, pozitron emisyon tomografi bilgisayarlı tomografi, fluorodeoksiglukoz F18, prognoz, sağkalım analizi

INTRODUCTION

GC is still one of the leading causes of cancer-related death worldwide, despite improvements in diagnosis and treatment (1). The 2019 WHO classification emphasizes the biological diversity of gastric adenocarcinoma, distinguishing intestinal, diffuse, mixed, and other subtypes with distinct patterns of behavior and prognosis (2). This variability diminishes the predictive utility of standard clinicopathological variables and underscores the necessity for quantitative markers that more precisely represent tumor biology. [18F] fluorodeoxyglucose positron emission tomography/computed tomography (18F-FDG PET/CT) is used in selected patients to identify distant metastases and refine treatment decisions. Several studies have shown that staging 18F-FDG PET/CT can reveal otherwise unsuspected metastatic disease and modify management in patients considered for radical therapy (3,4). These findings suggest that PET-derived metabolic information may also offer predictions that go beyond conventional anatomical staging. 18F-FDG PET/CT quantifies tumor glucose uptake through metabolic indices such as standardized uptake value maximum (SUVmax) and SUVmean. A meta-analysis demonstrated that higher pre-treatment SUVmax is associated with worse OS, although study results vary with tumor subtype and methodology (5). Cohort studies likewise report that SUV-based metrics correlate with more aggressive disease and poorer outcomes (6). Volumetric PET parameters, including MTV and TLG, may better represent total metabolic burden than SUV alone. In resectable and locally advanced disease, higher MTV or TLG consistently predicted shorter OS or recurrence-free survival and in several reports outperformed SUVmax as prognostic indicators (7-9). Several reports indicate that MTV and TLG retain prognostic value even within biologically defined subgroups, for example, in c-MET–positive GCs, suggesting that these volumetric indices add information beyond conventional staging (10). In parallel, there is increasing interest in intratumoral metabolic heterogeneity. Initial data suggest that non-uniform FDG uptake within the primary lesion, summarized by various heterogeneity indices, can predict survival independently of MTV and TLG (11). Taken together, these observations raise the possibility that combining volumetric parameters with heterogeneity measures may better reflect underlying tumor biology. HER2 status is also an important modifier of GC behavior. Although available series are relatively small, several have reported higher FDG uptake and distinct metabolic patterns in HER2-positive tumors, with signals of an association with clinical outcome (12,13). However, evidence remains limited and heterogeneous, and is rarely integrated into multivariable survival models. Overall, while metabolic and volumetric PET parameters show prognostic value in GC, it is not yet clear how much they improve established clinical risk models or whether integrating multiple PET-derived indices enhances prediction of OS in real-world settings (5-11). Few studies have examined calibration or clinical utility using decision-analytic methods.

In this study, we evaluated the incremental prognostic value of a broad panel of 18F-FDG PET/CT parameters—including SUVmax, SUVmean, SUV corrected for lean body mass peak (SULpeak), MTV_40, and TLG_40 as well as normalized measures such as BLR_mean and SLR_mean—when added to clinical prognostic models. We also assessed their relationship with HER2 expression and analyzed discrimination, calibration, and decision-analytic performance.

MATERIALS AND METHODS

Patient Selection

This retrospective observational study included consecutive patients with biopsy-confirmed primary gastric adenocarcinoma who underwent baseline 18F-FDG PET/CT for initial staging at Göztepe Prof. Dr. Süleyman Yalçın City Hospital between April 2016 and March 2025. Clinical, imaging, and follow-up data were

obtained from institutional electronic records. No formal sample size calculation was performed; all eligible patients within the study period were included.

Eligibility criteria included age ≥18 years, histologically confirmed primary gastric adenocarcinoma, and availability of a pretreatment 18F-FDG PET/CT acquired before the initiation of oncologic therapy. Patients with missing imaging data, unavailable pathology reports, or insufficient follow-up information were excluded. In total, 92 patients were initially screened. Three patients were excluded: one with missing pre-treatment PET/CT, one with non-adenocarcinoma histology, and one with incomplete baseline clinical data. The final study cohort consisted of 89 patients.

The study was approved by the Noninterventional Clinical Research Ethics Committee of Göztepe Prof. Dr. Süleyman Yalçın City Hospital (approval number: 2025/0053; date: 31.07.2025) and was conducted in accordance with the Declaration of Helsinki. Owing to the retrospective design of the study, informed consent was waived.

PET/CT Imaging Protocol

All PET/CT examinations were performed on the same integrated device using a standard institutional protocol. Patients fasted for at least 6 hours, and a weight-based dose of 18F-FDG (3.7 MBq/kg) was administered intravenously. Image acquisition started 60–70 minutes after injection.

Low-dose CT (140 kV, 40–60 mAs, 5-mm slices) was used for attenuation correction and anatomical localization. PET data were acquired from the skull base to mid-thigh and reconstructed in axial, coronal, and sagittal planes using consistent parameters throughout the study period.

Quantitative PET Metrics

Tumor segmentation was performed on the primary gastric lesion using semiautomated three-dimensional volumes of interest, and all VOIs were jointly reviewed and finalized by two nuclear medicine physicians with more than 12 years of PET/CT experience. To reduce potential bias, all quantitative PET metrics were extracted using a standardized protocol. The retrospective, single-center nature of the study, heterogeneity in treatment pathways, and missing HER2 data represent the main potential sources of bias. We sought to limit selection bias by including all eligible consecutive patients and to minimize measurement bias by using consistent PET/CT acquisition and reconstruction protocols throughout the study period. Survival bias was reduced by defining overall survival (OS) as beginning at the start of cancer management.

The following quantitative parameters were recorded:

SUVmax, SUVmean, SUVpeak, SULmax_James, SULpeak_James, SULmean_James, metabolic tumor volume at 40% of SUVmax (MTV_40) and total lesion glycolysis (TLG_40), BLR_mean = tumor SUVmean / bone marrow SUVmean, and SLR_mean = tumor SUVmean / spleen SUVmean. Tumor size on CT was measured as the largest axial diameter.

Lean body mass was estimated using the James equation.

For men: LBM = 1.10 × weight (kg) − 128 × [weight (kg)/height (m)]^2.

For women: LBM = 1.07 × weight (kg)− 148 × [weight (kg)/height (m)]^2.

SUL values were calculated by normalizing SUV to estimated lean body mass.

For bone marrow assessment, three-dimensional volumes of interest were placed within the T11–L4 vertebral bodies on PET/CT. A 75% SUVmax isocontour was used to delineate intramedullary uptake while excluding cortical bone and focal benign or metastatic lesions. After verification in the axial, coronal, and sagittal planes, the SUVmean values from the selected vertebrae were averaged and recorded as the bone marrow SUVmean.

Survival Endpoints

The primary endpoint was OS. OS was defined as the time (in months) from diagnosis or treatment initiation (OS_start in the dataset) to death from any cause or last known follow-up.

There was no loss to follow-up; censoring occurred only due to administrative end-of-study cutoffs.

Statistical Analysis

1. Descriptive Statistics

Continuous variables were described using the mean and standard deviation (SD), the median and interquartile range (IQR), and the observed minimum–maximum values. Categorical variables were summarized as counts and percentages.

2. Univariate Cox Regression

As an initial screening step, each PET/CT-derived quantitative metric (SUV/SUL measures, MTV_40, TLG_40, BLR_mean, SLR_mean, and tumor size) was entered separately into a Cox proportional hazards model. For these analyses, we estimated hazard ratios (HRs) with 95% confidence intervals (CIs) and reported the corresponding p-values.

3. Multivariable Cox Modeling

The main analysis consisted of three predefined Cox models:

a) Clinical Model

Included covariates were age, sex, and tumor differentiation (coded as 1 = well, 2 = moderate, 3 = poor/low). Lymph node metastasis was recorded as present or absent. Distant metastasis was defined as peritoneal carcinomatosis or the presence of at least one metastatic lesion in the bone, liver, lung, adrenal glands, or brain.

b) PET-Only Model

The included PET variables were SULmax_James, SULpeak_James, SULmean_James, MTV_40, TLG_40, BLR_mean, SLR_mean, and tumor size.

c) Clinical+PET Combined Model

The model was constructed by adding all PET variables to the clinical model. For each model, HR, 95% CI, p-values, and the concordance index (c-index) were obtained.

Potential confounding was addressed by prespecified multivariable Cox models including all clinically relevant covariates. Missing HER2 values (n = 19) were handled by listwise deletion for analyses involving HER2; no imputation was performed.

Model Handling and Software

HER2 status was missing for 19 patients; all other clinical and PET/CT variables had complete data.

There was no loss to follow-up; censoring occurred only due to administrative end-of-study cut-offs.

Differentiation was encoded ordinally, and sex was binarized (M = 0; W = 1).

All analyses were performed in Python 3.10 using the lifelines package (v0.30) and the CoxPHFitter function.

Survival Modelling and Model Performance

Cox proportional hazards models were fitted for three predefined predictor sets: (i) clinical variables (age, gender, tumor differentiation, lymph node metastasis, distant metastasis); (ii) PET-only variables (SULmax, SULpeak, SULmean, MTV_40, TLG_40, BLR_mean, SLR_mean, tumor size); and (iii) a combined Clinical+PET set including all of these covariates. Individual risk scores were derived from the model's linear predictor. For each model, patients were dichotomized into low- and high-risk groups using the cohort median, and OS was compared with Kaplan–Meier curves and two-sided log-rank tests (Figure 1a–c). Model discrimination was summarized using the c-index.

Model Calibration at Multiple Time Horizons

For model calibration at 12, 36, and 60 months, predicted survival probabilities from the Clinical+PET combined Cox model were obtained by evaluating the baseline survival function at each time point and applying the individual linear predictors. Predicted probabilities were grouped into 10 deciles, and the corresponding observed survival within each decile was estimated using Kaplan–Meier curves at the same horizon. Calibration was visualized using scatter plots of observed versus predicted survival, overlaid with a LOESS-smoothed curve and the ideal 45° reference line (Figure 2a–c).

Decision-Curve Analysis

To evaluate the potential clinical utility of the prognostic models, we performed decision curve analysis (DCA) at 12 and 24 months. Model-based risks of death by time t were obtained from the Cox models, and patients censored before t were excluded from that horizon-specific analysis. Net benefit was calculated for threshold probabilities between 10% and 40% using the standard formula:

$$NB = [(TP / N) - (FP / N) \times (pt / (1 - pt))]$$

Decision curves for the Clinical, PET-only, and Clinical+PET models were compared with the "treat all" and "treat none" reference strategies.

Presentation of multivariable effects

For each of the three multivariable Cox models (Clinical, PET-only, and Clinical+PET), HRs with 95% CIs were estimated and graphically summarized using forest plots.

Bootstrap-corrected calibration at 60 months

To further assess and adjust for potential optimism in model calibration, we performed a bootstrap-based out-of-bag (OOB) correction at the 60-month horizon for the Clinical+PET model. In each of 200 bootstrap replicates, a Cox model was refitted on the bootstrap sample, and 60-month survival probabilities were predicted only for OOB patients. For each patient, OOB predictions were then averaged across replicates to yield a bootstrap-corrected survival estimate. Apparent and OOB-corrected predictions were each grouped into 10 deciles, and observed survival within each decile was estimated from Kaplan–Meier curves. Calibration error was quantified using the mean absolute difference between predicted and observed probabilities (E_avg) and its 90th percentile (E90).

Distribution of predicted risk

For descriptive purposes, 60-month death risks derived from the Clinical+PET Cox model [1−S(60|X)] were plotted as a histogram to illustrate the distribution of patient-level predictions. The cohort median predicted risk was displayed as a reference line.

Risk Distribution Analysis

Risk distribution plots were generated to illustrate the spread of predicted probabilities of death among patients. For each time horizon (12, 24, 36, and 60 months), predicted survival probabilities

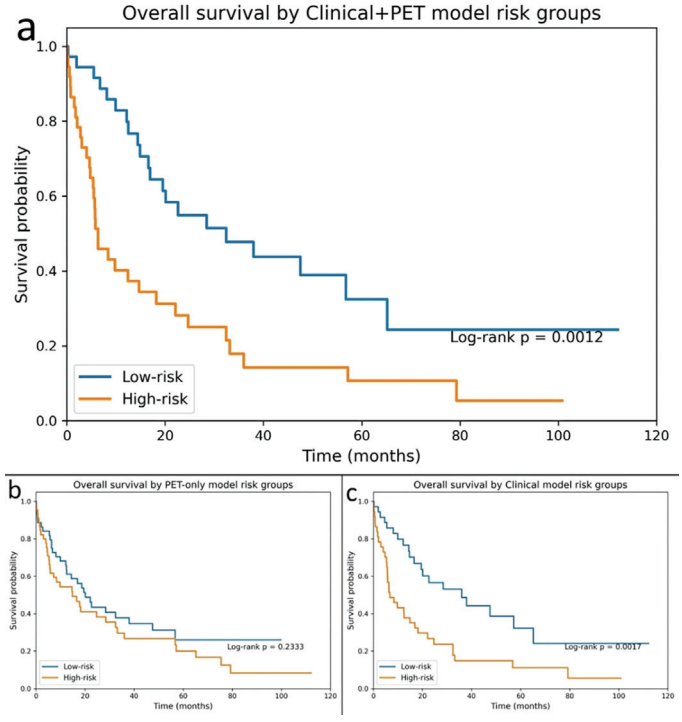

Figure 1. Overall survival stratified by three Cox-based risk models.

(a) Clinical+PET combined model: patients were divided into low- and high-risk groups based on the median Cox linear predictor derived from clinical variables (age, sex, histologic differentiation, lymph node metastasis, and distant metastasis) together with PET-derived quantitative metrics (SULmax, SULpeak, SULmean, MTV_40, TLG_40, BLR_mean, SLR_mean, and tumor size).

(b) PET-only model: risk groups defined according to the median PET-based Cox linear predictor.

(c) Clinical model: stratification based on the Cox linear predictor derived from clinical factors alone.

PET: Positron emission tomography.

were computed using the combined Clinical+PET Cox model by evaluating the baseline survival function at the specified time points. Individual predicted risks were defined as 1– S(t|X*), sort*ed in ascending order, and displayed as waterfall plots, with the cohort median risk indicated by a horizontal reference line. For tertile-based Kaplan–Meier analyses, patients were stratified according to the distribution of the Clinical+PET model–estimated 60-month mortality risk using tertile cutoffs of approximately 0.70 and 0.91.

Performance Metrics

Time-dependent discrimination was assessed at 12, 24, 36, and 60 months using an inverse-probability-of-censoring-weighted estimator applied to each model's linear predictor. Restricted mean survival time (RMST) was calculated for up to 60 months in low- and high-risk groups defined by the median risk score of the Clinical+PET model. For visual stratification of survival, the predicted 60-month risk was computed from the baseline survival function and used to generate tertile-based Kaplan–Meier curves.

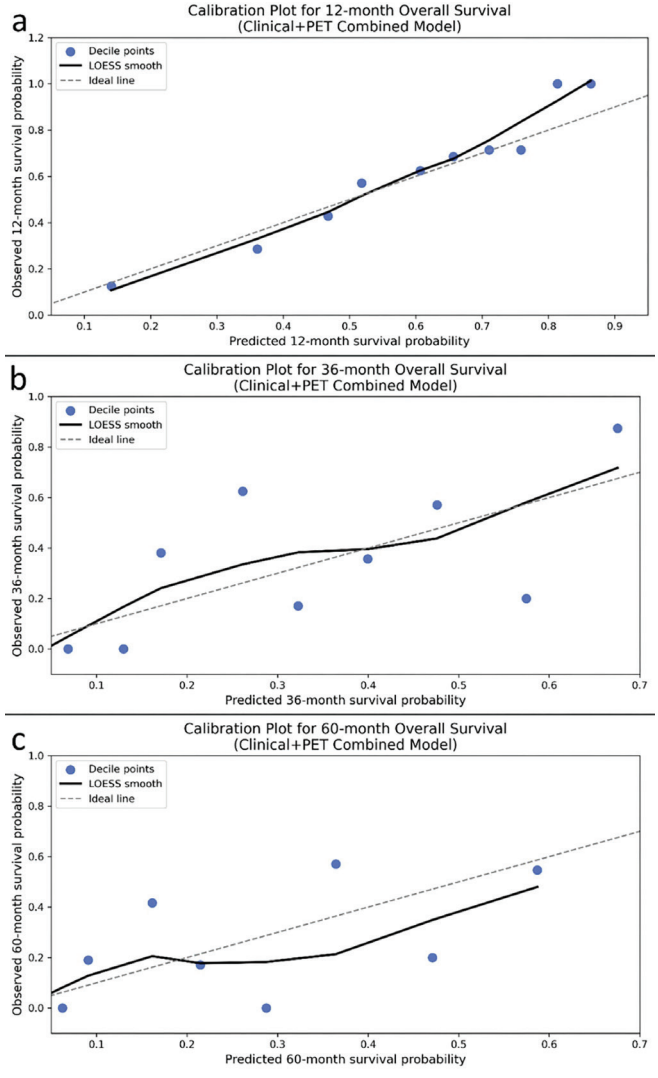

Figure 2. Calibration plots for predicting 12-, 36-, and 60-month overall survival using the Clinical+PET combined model.

PET: Positron emission tomography.

Follow-Up Time Estimation (Reverse Kaplan–Meier)

Follow-up duration was estimated using the reverse Kaplan–Meier method, in which censoring and death indicators are exchanged to obtain the distribution of potential follow-up time. The median follow-up and its 95% CI were extracted from the survival function of the reversed-event process.

Exploratory ROC Analysis for HER2 Prediction

To assess the discriminatory ability of PET-derived quantitative metrics for HER2 status, we included patients with complete HER2 data. Continuous PET parameters (SUVmax, SUVpeak, SUVmean, MTV_40, TLG_40, SUL-derived metrics, BLR_mean, SLR_mean, and tumor size) were evaluated individually using Receiver Operating Characteristic (ROC) analysis. Area under the curve (AUC) values and 95% CIs were estimated using bootstrap resampling (n = 1000 iterations). Given the limited number of HER2-positive cases (n = 12), this analysis was considered exploratory.

Study Design and Reporting

This observational study was conducted in accordance with the STROBE reporting principles. The design, data collection procedures, variables, and analytical methods were structured to meet these guidelines, and the completed STROBE checklist is included among the submission files.

RESULTS

Patient Characteristics

The study included 89 patients. The mean age was 64.2 ± 12.6 years, and 75.3% of the cohort were male. Lymph node metastasis was present in 67.4%, distant metastasis in 30.3%, and peritoneal carcinomatosis in 12.4% of patients. Detailed distributions of continuous and categorical variables are provided in Table 1. During follow-up, 64 deaths occurred, corresponding to an event rate of 71.9%.

Size refers to the longest diameter of the primary tumor, as measured on CT. SUV-based metrics were extracted from baseline 18F-FDG PET/CT scans acquired before treatment. MTV and TLG were calculated using a 40% SUVmax threshold. Continuous variables are reported as mean ± SD, median with IQR, and minimum–maximum values.

Univariate Analyses

In univariable Cox regression analyses, SULmean_James (HR 1.075, p = 0.048) was significantly associated with worse OS, while SULpeak_James (p = 0.053) and TLG_40 (p = 0.057) showed borderline significance. Among the background-normalized ratios, SLR_mean suggested the most pronounced increase in risk (HR 2.97; 95% CI, 0.67–13.12), whereas BLR_mean indicated a potential protective effect (HR 0.64); however, both estimates were statistically non-significant due to wide CIs (p = 0.150 and p = 0.304, respectively). Conventional SUV parameters and the remaining volumetric metric (MTV_40) demonstrated weaker and generally non-significant trends (Table 2).

Univariable Cox proportional hazards models were used to assess the relationships between quantitative baseline PET/CT parameters and OS. HRs are presented with 95% CIs and Wald p-values.

Table 1. Baseline clinical, metabolic, and clinicopathologic characteristics of the cohort.

A. Continuous variables

Variable	n	Mean	SD	Median	IQR	Min	Max
Age (years)	89	64.225	12.587	65.000	18.000	37.000	88.000
BMI (kg/m²)	89	25.965	4.456	26.235	5.345	14.901	38.200
Weight (kg)	89	72.348	13.531	72.000	18.000	40.000	104.000
Height (cm)	89	166.899	8.666	167.000	10.000	143.000	185.000
Tumor size (mm)	89	57.663	27.026	53.000	30.000	16.000	153.000
SUVmax	89	9.518	6.424	7.660	5.390	2.700	35.640
SUVpeak	89	7.759	5.545	5.980	4.820	2.140	30.970
SUVmean	89	5.406	3.790	4.240	3.450	1.490	21.490
MTV_40	89	45.673	40.074	34.690	38.970	2.940	177.980
TLG_40	89	306.412	505.852	143.140	201.270	7.370	2736.410
SULmean (James)	89	4.105	2.980	3.100	2.790	1.140	15.260
SULpeak (James)	89	5.830	4.325	4.580	4.060	1.630	22.300
SULmax (James)	89	7.146	4.991	5.730	4.920	2.070	25.050
BLR_mean	89	1.240	0.333	1.212	0.394	0.000	1.930
SLR_mean	89	0.909	0.187	0.909	0.175	0.187	1.478

B. Categorical variables

Variable	Category	N	Percent
Gender	Man	67	75.28
Gender	Woman	22	24.72
Lesion site	Antrum	31	34.83
Lesion site	Corpus	16	17.98
Lesion site	Cardia and corpus	15	16.85
Lesion site	Corpus and antrum	11	12.36
Lesion site	Cardia	9	10.11
Lesion site	Cardia–fundus–corpus	3	3.37
Lesion site	Cardia, corpus, and antrum	1	1.12
Lesion site	Cardia and fundus	1	1.12
Lesion site	Corpus and fundus	1	1.12
Lesion site	Cardia, fundus, corpus, antrum	1	1.12
Peritonitis carcinomatosa	0	78	87.64
Peritonitis carcinomatosa	1	11	12.36
Lymph node metastasis	1	60	67.42
Lymph node metastasis	0	29	32.58
Bone metastasis	0	81	91.01
Bone metastasis	1	8	8.99
Liver metastasis	0	76	85.39
Liver metastasis	1	13	14.61
Lung metastasis	0	87	97.75
Lung metastasis	1	2	2.25
Adrenal metastasis	0	88	98.88
Adrenal metastasis	1	1	1.12
Pathological subtype	Tubular	49	55.1
Pathological subtype	Poorly cohesive	27	30.3

Table 1. Continued.

B. Categorical variables							
Variable	**n**	**Mean**	**SD**	**Median**	**IQR**	**Min**	**Max**
Pathological subtype	Mucinous	6	6.70				
Pathological subtype	Signet ring cell	4	4.50				
Pathological subtype	Medullary	2	2.20				
Pathological subtype	Hepatoid	1	1.10				
Differentiation	1 (well)	4	4.50				
Differentiation	2 (moderate)	51	57.3				
Differentiation	3 (poor)	34	38.2				
HER2	0	58	65.20				
HER2	1	12	13.50				
HER2	Missing	19	21.30				

Min: Minimum, Max: Maximum, SD: Standard deviation, IQR: Interquartile range, BMI: Body mass index.

Each parameter was analyzed in a separate Cox model; n_used denotes the number of patients with available data for that variable (total analyzed cohort, n = 89). MTV_40 and TLG_40 were calculated using a 40% SUVmax threshold.

Multivariable Analyses

1. Clinical Model

In the clinical model, two variables were independently associated with OS:

Age (HR = 1.04; 95% CI, 1.01–1.06; p = 0.006)

Presence of distant metastasis (HR = 2.16; 95% CI, 1.14–4.09; p = 0.018)

The model's discriminative performance was c-index = 0.669.

2. PET-Only Model

In the *PET*-only model, two background-normalized ratios stood out as independent predictors. Higher BLR_mean was associated with a lower risk of mortality (HR 0.26; p = 0.010), whereas SLR_mean showed the opposite pattern, with a markedly increased risk (HR 6.82; p = 0.032). By contrast, conventional SUV/SUL measures and the volumetric parameters (MTV_40 and TLG_40) did not reach independent significance.

3. Clinical + PET Combined Model

Incorporating PET-derived variables into the clinical model modestly improved prognostic performance. Within the combined model, MTV_40 (HR ≈ 1.02; p = 0.068) and BLR_mean (HR ≈ 0.32; p = 0.079) showed borderline associations with OS. In contrast, in the PET-only model, BLR_mean (p = 0.010) and SLR_mean (p = 0.032) were statistically significant predictors. The Clinical+PET combined model achieved *th*e highest discrimination among the three predefined models (c-index = 0.728). Full results are presented in Table 3.

Multivariable Cox models were fitted using three predefined predictor sets: a clinical model (age, gender, tumour differentiation, lymph-node metastasis, any distant metastasis), a PET-only model (SULmax, SULpeak, SULmean, MTV_40, TLG_40, BLR_mean, SLR_mean, tumour size), and a combined Clinical + PET model including all variables from both sets. For each covariate, HRs with 95% CIs and Wald p-values are reported.

Table 2. Univariable associations between baseline PET/CT metrics and overall survival.

Variable	**HR**	**CI_lower**	**CI_upper**	**p-value**	**n_ used**
SULmean_James	1.075	1.001	1.155	0.048	89
SULpeak_James	1.048	0.999	1.100	0.053	89
TLG_40	1.000	1.000	1.001	0.057	89
SULmax_James	1.042	0.998	1.087	0.059	89
Size	1.009	1.000	1.018	0.063	89
SUVmean	1.052	0.995	1.112	0.074	89
MTV_40	1.006	0.999	1.012	0.075	89
SUVpeak	1.034	0.997	1.074	0.076	89
SUVmax	1.029	0.996	1.064	0.084	89
SLR_mean	2.973	0.674	13.117	0.150	89
BLR_mean	0.641	0.274	1.497	0.304	89

CI: Confidence interval, TLG: Total lesion glycolysis.

Overall Interpretation

Among the three prespecified models, the combined Clinical+PET model provided the highest prognostic performance, whereas the PET-only and clinical models demonstrated more limited but complementary discriminatory performance. Kaplan–Meier curves for these risk models are presented in Figure 1a–c.

Survival curves were compared using the log-rank test. Time shown in months.

Calibration of the Clinical+PET model showed good overall agreement between predicted and observed survival across all three time horizons (12, 36, and 60 months).

At 12 months, predicted probabilities aligned closely with observed outcomes, with minimal systematic deviation (Figure 2a).

At 36 months, the model maintained broadly acceptable calibration, although some variability was noted in mid-range risk strata (Figure 2b).

At 60 months, calibration remained acceptable, but the model tended to modestly overestimate survival in higher predicted-risk groups, consistent with a mild upward drift of the LOESS curve relative to the ideal 45° line (Figure 2c).

Each point represents one decile of predicted survival probabilities, averaged across patients within that decile. Observed survival was estimated from Kaplan–Meier curves at the corresponding horizons (12, 36, or 60 months). The dashed diagonal line denotes perfect calibration, while the solid line shows a LOESS-smoothed estimate of the observed–predicted relationship. The model showed close agreement between predicted and observed outcomes at 12 months, moderate variability at 36 months, and mild overestimation of survival probabilities in higher predicted-probability ranges at 60 months.

DCA indicated that the Clinical+PET model performed best at 12 and 24 months, offering greater net benefit than either the clinical-only or the PET-only model within the accepted 0.10–0.40 clinical threshold range.

The combined model demonstrated a superior net-benefit profile at 12 months, particularly for threshold probabilities between approximately 0.15 and 0.35 (Figure 3a).

At 24 months, all three models performed similarly, but the combined model again yielded a slightly higher net benefit across most thresholds (Figure 3b).

In multivariable Cox regression analysis (Figure 4), age and distant metastasis remained independent clinical predictors of OS. Among PET-derived parameters, higher BLR_mean was associated with improved survival, whereas higher SLR_mean indicated increased mortality risk. In the combined Clinical+PET model, clinical factors retained their prognostic significance, whereas PET metrics did not substantially alter effect directions but improved overall model discrimination (c-index ≈0.73).

At 60 months, apparent calibration of the Clinical+PET model showed a mean absolute calibration error (E_avg) of 0.129 and a 90th-percentile error (E90) of 0.279. Bootstrap OOB–corrected estimates were very similar (E_avg = 0.136, E90 = 0.294), indicating

Table 3. Multivariable Cox proportional hazards models for overall survival.

Covariate	Model	HR	HR_lower_95	HR_upper_95	p	coef	se (coef)
Age	Clinical	1.035	1.010	1.061	0.006	0.034	0.013
Gender_bin	Clinical	0.911	0.460	1.804	0.789	-0.093	0.349
Differentiation	Clinical	1.336	0.792	2.254	0.278	0.290	0.267
LN_met	Clinical	1.429	0.738	2.764	0.289	0.357	0.337
Any_distant_met	Clinical	2.162	1.143	4.089	0.018	0.771	0.325
SULmax_James	PET	0.741	0.327	1.678	0.472	-0.300	0.417
SULpeak_James	PET	1.161	0.559	2.411	0.689	0.149	0.373
SULmean_James	PET	1.431	0.620	3.303	0.402	0.358	0.427
TLG_40	PET	0.999	0.998	1.001	0.460	-0.001	0.001
MTV_40	PET	1.007	0.993	1.021	0.314	0.007	0.007
BLR_mean	PET	0.264	0.095	0.731	0.010	-1.332	0.520
SLR_mean	PET	6.823	1.179	39.493	0.032	1.920	0.896
Size	PET	1.007	0.990	1.024	0.447	0.007	0.009
Age	Clinical+PET	1.033	1.005	1.063	0.022	0.033	0.014
Gender_bin	Clinical+PET	1.103	0.530	2.296	0.793	0.098	0.374
Differentiation	Clinical+PET	1.250	0.699	2.235	0.452	0.223	0.297
LN_met	Clinical+PET	1.044	0.427	2.555	0.925	0.043	0.457
Any_distant_met	Clinical+PET	2.085	1.005	4.324	0.048	0.735	0.372
SULmax_James	Clinical+PET	0.742	0.265	2.080	0.571	-0.298	0.526
SULpeak_James	Clinical+PET	1.262	0.517	3.079	0.609	0.233	0.455
SULmean_James	Clinical+PET	1.313	0.520	3.321	0.564	0.273	0.473
TLG_40	Clinical+PET	0.999	0.997	1.001	0.319	-0.001	0.001
MTV_40	Clinical+PET	1.019	0.999	1.040	0.068	0.019	0.010
BLR_mean	Clinical+PET	0.318	0.089	1.140	0.079	-1.144	0.651
SLR_mean	Clinical+PET	3.987	0.523	30.389	0.182	1.383	1.036
Size	Clinical+PET	0.994	0.972	1.015	0.559	-0.006	0.011

Differentiation was coded as an ordinal variable (1 = well-differentiated, 2 = moderately differentiated, 3 = poorly differentiated).
BLR_mean = tumor-to-liver- bone marrow SUV_{mean} ratio; SLR_mean = tumor-to-spleen SUVmean ratio.
PET/CT: Positron emission tomography/computed tomography.

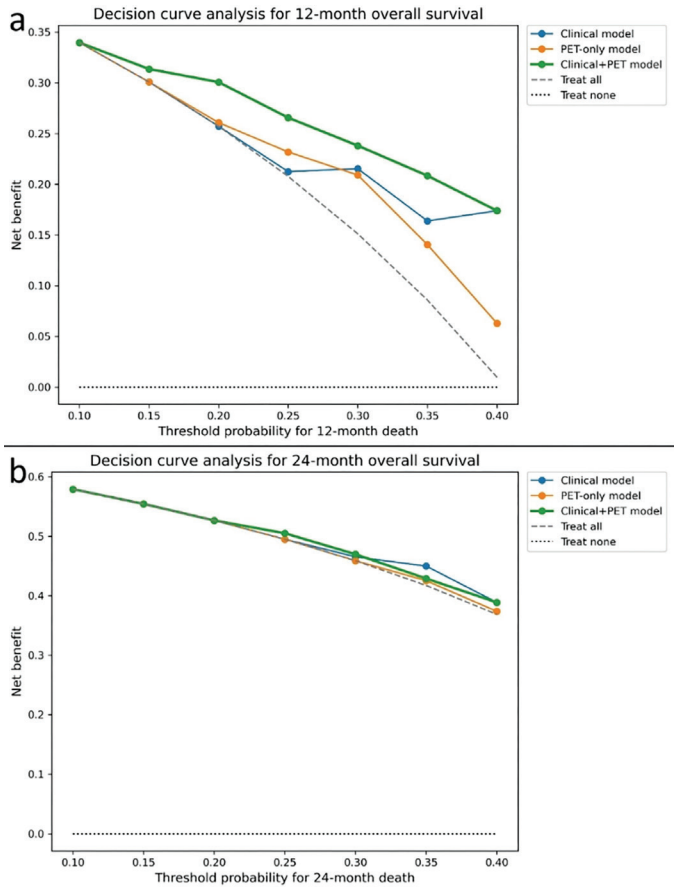

Figure 3. Decision curve analysis for 12- and 24-month overall survival.

(a) Decision curve analysis at 12 months, comparing three prespecified Cox models: the clinical model (age, gender, differentiation, lymph node metastasis, distant metastasis).

The PET-only model included the predefined PET-derived variables, whereas the combined Clinical+PET model included both clinical and PET-derived variables.

The net benefit is plotted across threshold probabilities from 0.10 to 0.40, with reference lines for the 'treat-all' and 'treat-none' strategies.

The Clinical+PET model provides a consistently higher net benefit across most thresholds.

(b) Decision curve analysis at 24 months, showing a similar pattern, with the Clinical+PET model maintaining a small but consistent net-benefit advantage over the other two models.

PET: Positron emission tomography, MTV: Metabolic tumor volume, Total lesion glycolysis, SLR_mean: tumor-to-spleen SUVmean ratio.

only modest optimism and broadly preserved calibration after correction (Figures 5a–b).

The distribution of predicted 60-month mortality risk from the Clinical+PET model was right-skewed, indicating that most patients clustered within the lower predicted risk range. Predicted risks had a median of 0.81 (range, 0.33–1.00). This pattern aligns with the overall high-risk profile of the cohort (Figure 6).

Each bar represents the number of patients within a given probability interval. The dashed vertical line indicates the cohort median predicted risk (0.81). The distribution displays a right-shifted pattern, consistent with the high-risk profile of the cohort.

The risk plots showed that predicted mortality varied widely across patients. At 12, 24, 36, and 60 months, the predicted risk did not follow a symmetric distribution; instead, most patients had moderate risk values, with only a small fraction in the high-risk tail. Median predicted risk increased as the horizon extended, in line with the accumulation of events over time. The skew was most marked at 60 months, reflecting the greater long-term mortality burden in this cohort (Figure 7).

In the time-dependent AUC analysis, the Clinical+PET model maintained AUC values between 0.85 and 0.89 at all evaluated time points, and these values were higher than those of the Clinical-only and PET-only models. Taken together, a multidimensional comparison of model performance across discrimination, calibration, and clinical utility metrics is presented in Figure 8. In the RMST analysis, low- and high-risk groups differed clearly, with mean restricted survival times of 33.9 months and 17.7 months, respectively. Kaplan–Meier curves based on tertiles of the predicted 60-month risk showed a stepwise separation, consistent with a strong prognostic signal of the model.

Median follow-up, calculated using the reverse Kaplan–Meier method, was 60.0 months (95% CI, 42.1–79.7 months) (Figure 9).

The reverse Kaplan–Meier method was used to estimate the distribution of potential follow-up time, treating deaths as censored observations and censoring events as failures. The vertical dashed line denotes the median follow-up of 60.0 months (95% CI, 42.1–79.7 months). Shaded areas represent the 95% confidence bands.

HER2 Status Prediction (Exploratory Analysis)

A subset of 70 patients with available HER2 data (HER2–: n=58; HER2+: n=12) was used to explore whether PET-derived metabolic and volumetric parameters could discriminate HER2 status. Overall discrimination was limited, with AUC values ranging from 0.536 to 0.609 across SUV/SUL- and volume-based metrics. The highest (yet still modest) performance was observed for SUVmean (AUC 0.609; 95% CI 0.406–0.796) and SULmean_James (AUC 0.606; 95% CI 0.404–0.787). BLR_mean and SLR_mean showed no discriminatory ability (AUC < 0.36). These results should be interpreted as exploratory due to the small number of HER2-positive cases (Table 4).

We conducted receiver operating characteristic (ROC) analysis for HER2 positivity based on individual PET-derived metrics (n=70; HER2–=58, HER2+=12). Values represent AUCs with 95% CIs. Given the small number of HER2-positive patients, results should be interpreted cautiously.

DISCUSSION

In this study, we evaluated whether a broad panel of 18F-FDG PET/CT metabolic and volumetric parameters improves the prediction of OS beyond established clinical variables in patients with GC. The need for more refined prognostication is well recognized. Gastric cancer (GC) remains a substantial global health concern (1), and the biological heterogeneity described in the 2019 WHO classification continues to hinder consistent risk stratification (2). These differences in tumor behavior encourage the search for quantitative

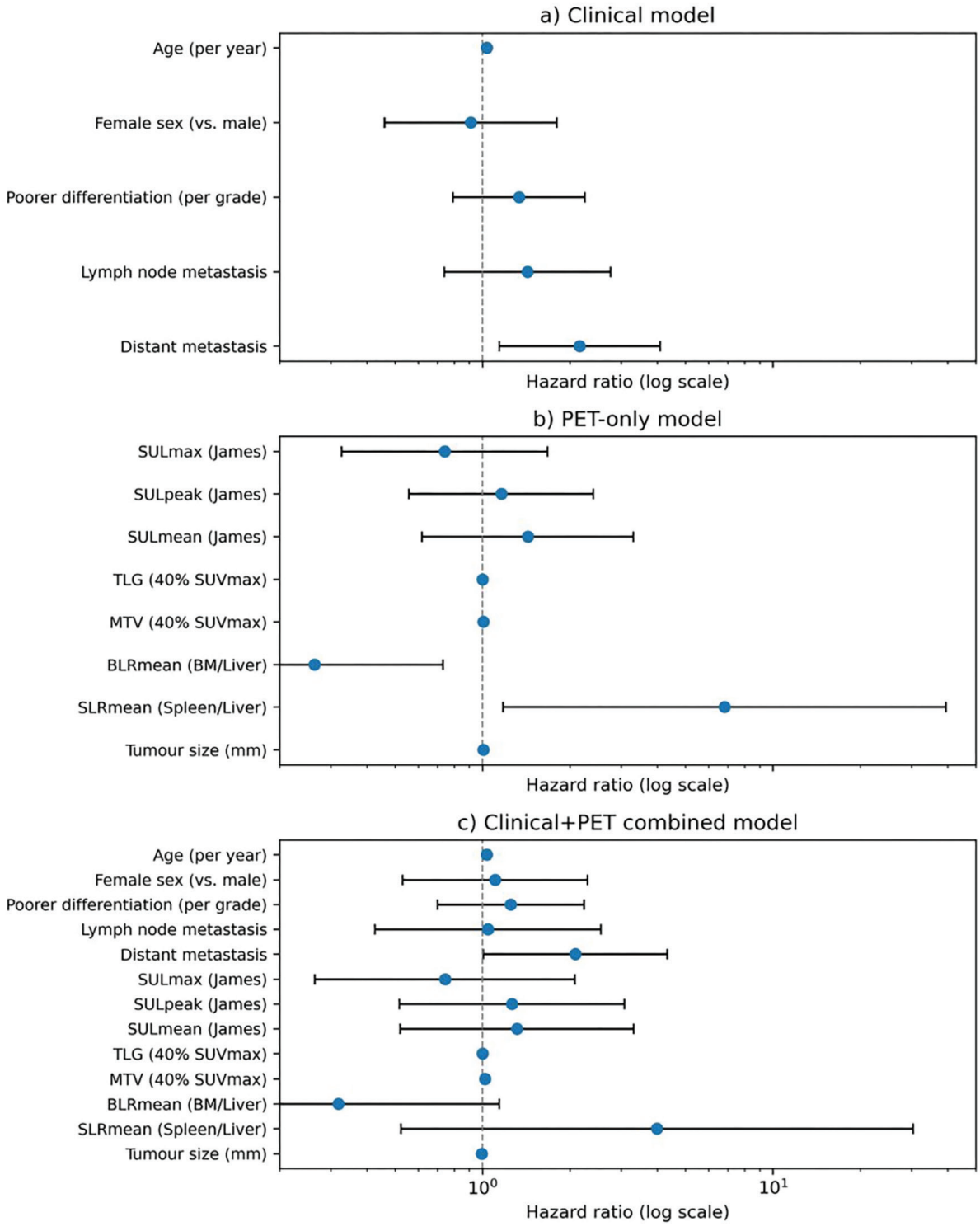

Figure 4. Multivariable Cox regression models for overall survival.

The forest plots show HRs and 95% CIs for (a) the clinical model, (b) the PET-only model, and (c) the combined Clinical+PET model. HRs are displayed on a logarithmic scale. Blue circles denote point estimates, and horizontal lines represent 95% CIs. The dashed vertical line indicates HR = 1. Clinical variables included age, sex, differentiation grade, lymph-node metastasis, and distant metastasis; PET variables included SULmax, SULpeak, SULmean, MTV_40, TLG_40, BLR_mean, SLR_mean, and tumor size.

PET: Positron emission tomography, MTV: Metabolic tumor volume, TLG: Total lesion glycolysis, SLR_mean: tumor-to-spleen SUVmean ratio, BLR_mean: tumor-to-bone marrow SUVmean ratio.

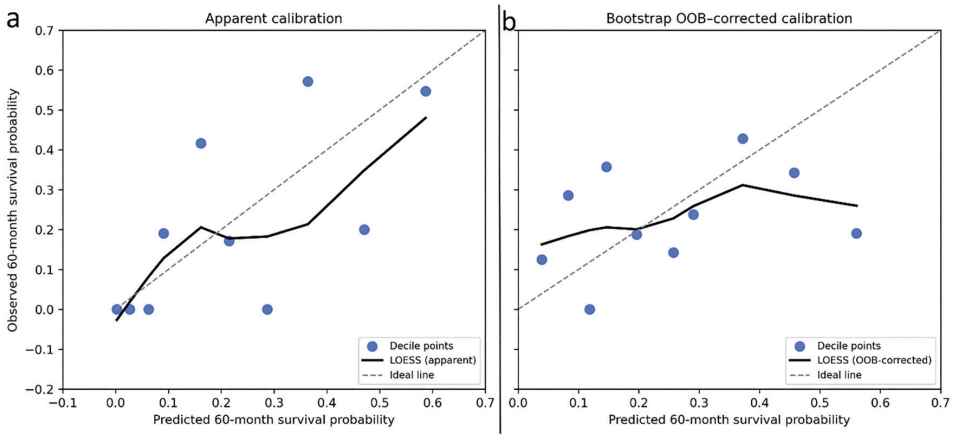

Figure 5. Apparent and bootstrap-corrected calibration of the Clinical+PET model at 60 months.

(a) The Clinical+PET Cox model was calibrated to predict the OS at 60 months. Predicted survival probabilities were grouped into 10 deciles, and observed survival within each decile was estimated from Kaplan–Meier curves at 60 months. Scatter points represent decile-specific predicted versus observed survival; the dashed diagonal line indicates perfect calibration, and the solid line shows a LOESS-smoothed calibration curve.

(b) Bootstrap out-of-bag (OOB)–corrected calibration at 60 months based on 200 bootstrap resamples. For each patient, OOB predictions were averaged across replicates to obtain a corrected survival estimate, which was then summarized in deciles and plotted analogously. Mean absolute calibration error (E_avg) was 0.129 for the apparent curve and 0.136 after OOB correction, with corresponding E90 values of 0.279 and 0.294.

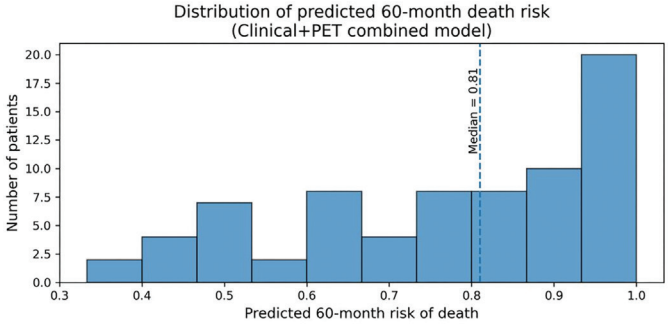

Figure 6. Distribution of predicted 60-month risk of death for the Clinical+PET combined Cox model.

PET: Positron emission tomography.

biomarkers that capture tumor aggressiveness and complement routine clinicopathological features (14).

The potential value of 18F-FDG PET/CT in this setting is supported by evidence showing that PET can reveal occult metastases and influence management in a substantial subset of patients considered for radical treatment (3,4,15). In addition to staging, metabolic data provide prognostic information that may indicate the underlying biology of the tumor. A number of prior studies have found that SUV-based metrics, especially SUVmax, correlate with worse clinical outcomes (5,6). Because these values depend on technical factors such as uptake timing and image reconstruction (7-9,16-18), they offer only a partial view of tumor metabolism and may vary across institutions.

Volumetric parameters such as MTV and TLG, which reflect the overall metabolic burden, have shown more consistent prognostic value (19). Studies in stage III and locally advanced GC frequently report that higher MTV or TLG is associated with inferior survival or an increased likelihood of early recurrence (7-10). Our findings

Table 4. Performance of PET/CT quantitative parameters for predicting HER2 status (exploratory analysis).

Variable	AUC	CI_lower	CI_upper	n_used
SUVmax	0.596	0.391	0.792	70
SUVpeak	0.593	0.385	0.799	70
SUVmean	0.609	0.406	0.796	70
MTV_40	0.536	0.320	0.733	70
TLG_40	0.579	0.346	0.793	70
SULmax_James	0.601	0.400	0.785	70
SULpeak_James	0.600	0.399	0.790	70
SULmean_James	0.606	0.404	0.787	70
BLR_mean	0.356	0.151	0.570	70
SLR_mean	0.345	0.167	0.539	70
Size	0.560	0.351	0.769	70

CI: Confidence interval, AUC: Area under the curve.

are consistent with this literature. MTV_40 and TLG_40 showed trends toward contributing to the model, although these effects did not consistently reach statistical significance. Their inclusion was associated with a modest improvement in discrimination, while calibration remained acceptable. These results are consistent with earlier observations that volumetric tumor burden remains prognostically relevant even in biologically distinct subsets, such as c-MET–positive tumors (10). The link between metabolic burden and c-MET activation supports a biological rationale for the observed associations.

Heterogeneity in FDG uptake has also emerged as a feature associated with aggressive tumor biology. Early work demonstrated that intratumor metabolic heterogeneity predicts survival independent of conventional metabolic or volumetric indices (11). Additional prognostic evidence has also been reported from interim and

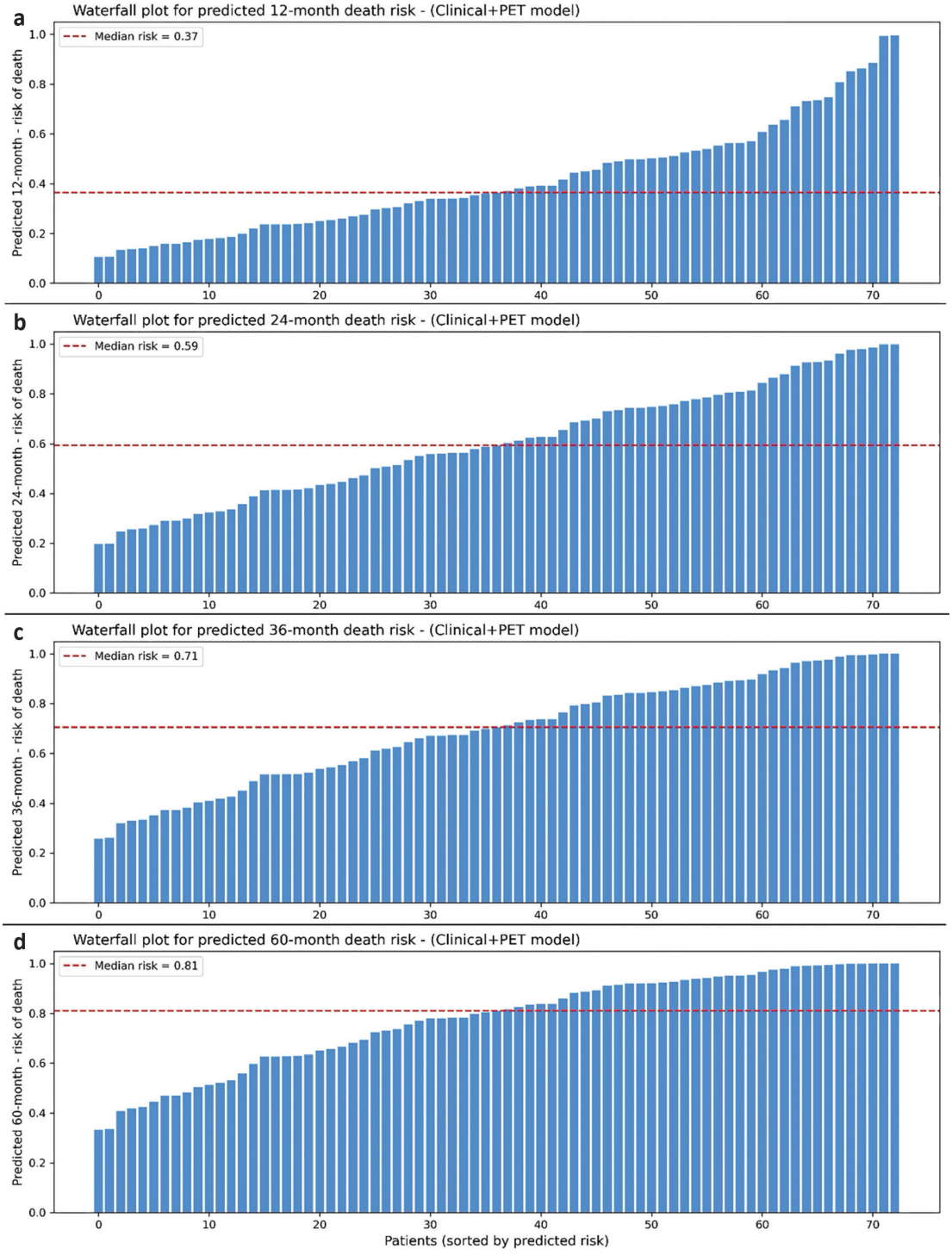

Figure 7. Waterfall plots of predicted death risk at 12, 24, 36, and 60 months derived from the Clinical+PET combined Cox model. Each panel shows the distribution of individual patient risk estimates at a specific time horizon (a: 12 months, b: 24 months, c: 36 months, d: 60 months). Predicted risks of death were obtained from the Clinical+PET Cox model by evaluating the baseline survival function at each time point. Patients are sorted in ascending order of predicted risk. Dashed red lines indicate the median predicted risk for each horizon.

PET: Positron emission tomography.

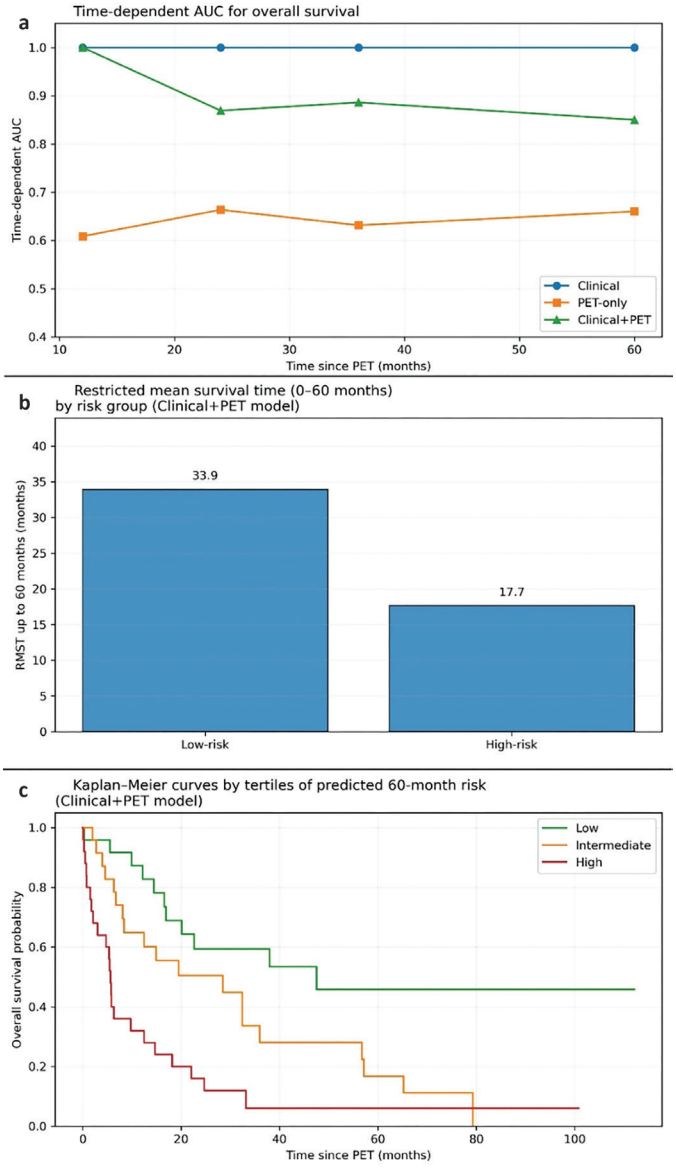

Figure 8. Multidimensional performance evaluation of the prognostic models.

(a) Time-dependent AUC values at 12, 24, 36, and 60 months for the Clinical, PET-only, and Clinical+PET Cox models, computed using an IPCW-based estimator. The Clinical+PET model demonstrated the highest and most stable discrimination across all horizons.

(b) Restricted mean survival time (RMST) up to 60 months comparing low- and high-risk groups derived from the Clinical+PET model's linear predictor. Low-risk patients had substantially longer event-free survival than high-risk patients.

(c) Kaplan–Meier survival curves stratified by tertiles of predicted 60-month mortality risk (Clinical+PET model). Clear monotonic separation was observed across the low-, intermediate-, and high-risk strata, consistent with strong overall prognostic discrimination.

PET: Positron emission tomography.

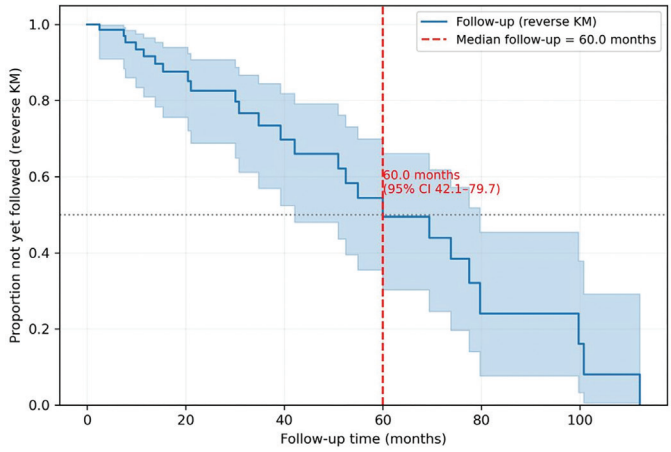

Figure 9. Reverse Kaplan–Meier curve for follow-up time.

restaging PET studies in metastatic or recurrent gastric cancer (20,21). More recent analyses confirmed that heterogeneity-based metrics provide important prognostic information, likely reflecting variations in cell density, hypoxia, or necrosis (22). In our cohort, although voxel-level heterogeneity metrics were not computed, BLR_mean emerged as a statistically significant PET-derived predictor and may act as a practical surrogate for metabolic variability, albeit within the limitations of our sample size.

Histological diversity remains an important determinant of FDG avidity. Diffuse-type and mucinous GCs typically show lower FDG uptake than intestinal-type tumors, which can complicate prognostic interpretation (23). Such histological differences need to be considered when PET-derived biomarkers are incorporated into survival models. In our cohort, PET-based parameters showed directionally consistent associations with outcome after adjustment for clinicopathological variables, although these associations were attenuated. This suggests that metabolic tumor burden may capture risk across histological subtypes.

The link between FDG uptake and HER2 status has also attracted attention. Several reports have described higher SUV values in HER2-positive GCs (12,13) or have proposed distinct metabolic patterns according to HER2 expression (24-27), although these studies are heterogeneous and generally involve modest sample sizes. In the present analysis, PET-derived metabolic parameters showed prognostic trends that were not fully explained by HER2 status. However, the number of HER2-positive cases was limited, and these findings should be interpreted with caution.

More recently, developments in image analysis have expanded the range of PET-based prognostic tools. Radiomics approaches, in particular, have been used to explore associations with survival, lymphovascular invasion, and treatment response (28-33). Machine learning approaches, including deep learning models, have demonstrated encouraging performance in estimating survival or immunotherapy response (31). Although these methods were outside the scope of our study, our findings may serve as a foundation for future multimodal prognostic models that incorporate handcrafted and derived PET features.

PET has also been used extensively to assess response during neoadjuvant therapy. Several reports have shown that early metabolic changes after the first cycles of chemotherapy are associated with histopathological tumor regression and longer-term outcomes (34-37), although some subgroups appear to benefit less from this approach (35). Our study centered on pretreatment scans rather than treatment response, but early metabolic change remains a promising area for future research, particularly when evaluated alongside volumetric PET measures.

Radiomics-based and advanced quantitative analyses have also been used to predict metastatic risk. Baseline PET radiomics features were associated with subsequent distant metastases in gastro-esophageal cancers in a recent prospective investigation (38). Our results suggest that PET-derived signatures may help to identify patients at higher risk who could be candidates for more intensive systemic therapy or closer surveillance.

Taken together with previous work, this supports a complementary, rather than standalone, role for PET-based metrics. In our dataset, adding PET parameters to clinical factors improved model discrimination and increased net benefit across relevant decision thresholds while maintaining acceptable calibration. These observations argue for incorporating PET information into prognostic models, especially when treatment decisions depend on estimated risk.

Study Limitations

It was a retrospective, single-center study, which may limit the generalizability of the findings. Diffuse-type tumors with intrinsically low FDG uptake may have attenuated the apparent prognostic value of metabolic parameters. We did not investigate texture-based radiomics features, which could provide complementary information but would require robust standardization before clinical use. Treatment heterogeneity, particularly with respect to HER2-targeted regimens, also restricted the depth of subgroup analyses, and HER2 status was unavailable in a subset of patients. However, all PET/CT studies were acquired and processed using uniform protocols and a standardized segmentation workflow, which likely reduced measurement variability and supported the internal consistency of the results.

CONCLUSION

Our results indicate that 18F-FDG PET/CT–derived metabolic and volumetric parameters offer complementary prognostic information beyond routine clinical factors in GC. The consistent associations observed for MTV_40, TLG_40, SULpeak, and BLR_mean support their use as practical imaging biomarkers that can strengthen survival models. Integrating these PET-derived measures with clinical variables may improve risk stratification and support more informed therapeutic decision-making in selected patients.

Ethics

Ethics Committee Approval: This study was approved by the Noninterventional Clinical Research Ethics Committee of Göztepe Prof. Dr. Süleyman Yalçın City Hospital (approval number: 2025/0053; date: 31.07.2025) and was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice and Good Laboratory Practice standards.

Informed Consent: Waived due to the retrospective design of the study.

Footnotes

Authorship Contributions

Concept: M.T.T., A.A., E.İ., A.N.T.Y., Design: M.T.T., H.U., Data Collection or Processing: M.T.T., A.A., Analysis or Interpretation: M.T.T., E.İ., F.Y.T., Literature Search: M.T.T., F.Ö., Writing: M.T.T.

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

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

REFERENCES

1. GBD 2019 Cancer Risk Factors Collaborators. The global burden of cancer attributable to risk factors, 2010-19: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2022; 400: 563-91.
2. Nagtegaal ID, Odze RD, Klimstra D, Paradis V, Rugge M, Schirmacher P, et al. The 2019 WHO classification of tumours of the digestive system. Histopathology. 2020; 76: 182-8.
3. Bosch KD, Chicklore S, Cook GJ, Davies AR, Kelly M, Gossage JA, et al. Staging FDG PET-CT changes management in patients with gastric adenocarcinoma who are eligible for radical treatment. Eur J Nucl Med Mol Imaging. 2020; 47: 759-67.
4. Findlay JM, Antonowicz S, Segaran A, El Kafsi J, Zhang A, Bradley KM, et al. Routinely staging gastric cancer with 18F-FDG PET-CT detects additional metastases and predicts early recurrence and death after surgery. Eur Radiol. 2019; 29: 2490-8.
5. Wu Z, Zhao J, Gao P, Song Y, Sun J, Chen X, et al. Prognostic value of pretreatment standardized uptake value of F-18-fluorodeoxyglucose PET in patients with gastric cancer: a meta-analysis. BMC Cancer. 2017; 17: 275.
6. Grabinska K, Pelak M, Wydmanski J, Tukiendorf A, d'Amico A. Prognostic value and clinical correlations of 18-fluorodeoxyglucose metabolism quantifiers in gastric cancer. World J Gastroenterol. 2015; 21: 5901-9.
7. Na SJ, o JH, Park JM, Lee HH, Lee SH, Song KY, et al. Prognostic value of metabolic parameters on preoperative 18F-fluorodeoxyglucose positron emission tomography/computed tomography in patients with stage III gastric cancer. Oncotarget. 2016; 7: 63968-80.
8. Song J, Li Z, Chen P, Yu J, Wang F, Yang Z, et al. A 18FDG PET/CT-based volume parameter is a predictor of overall survival in patients with local advanced gastric cancer. Chin J Cancer Res. 2019; 31: 632-40.
9. Sun G, Cheng C, Li X, Wang T, Yang J, Li D. Metabolic tumor burden on postsurgical PET/CT predicts survival of patients with gastric cancer. Cancer Imaging. 2019; 19: 18.
10. Liu G, Hu Y, Cheng X, Wang Y, Gu Y, Liu T, et al. Volumetric parameters on 18F-FDG PET/CT predict the survival of patients with gastric cancer associated with their expression status of c-MET. BMC Cancer. 2019; 19: 790.
11. Liu G, Yin H, Cheng X, Wang Y, Hu Y, Liu T, et al. Intra-tumor metabolic heterogeneity of gastric cancer on 18F-FDG PETCT indicates patient survival outcomes. Clin Exp Med. 2021; 21: 129-38.
12. Kim JS, Young Park S. (18)F-FDG PET/CT of advanced gastric carcinoma and association of HER2 expression with standardized uptake value. Asia Ocean J Nucl Med Biol. 2014; 2: 12-8.
13. Celli R, Colunga M, Patel N, Djekidel M, Jain D. Metabolic Signature on 18F-FDG PET/CT, HER2 Status, and Survival in Gastric Adenocarcinomas. J Nucl Med Technol. 2016; 44: 234-8.
14. Abbas M, Habib M, Naveed M, Karthik K, Dhama K, Shi M, et al. The relevance of gastric cancer biomarkers in prognosis and pre- and post- chemotherapy in clinical practice. Biomed Pharmacother. 2017; 95: 1082-90.

15. Shimada H, Okazumi S, Koyama M, Murakami K. Japanese Gastric Cancer Association Task Force for Research Promotion: clinical utility of ^{18}F-fluoro-2-deoxyglucose positron emission tomography in gastric cancer. A systematic review of the literature. Gastric Cancer. 2011; 14: 13-21.
16. Vriens D, Visser EP, de Geus-Oei LF, Oyen WJ. Methodological considerations in quantification of oncological FDG PET studies. Eur J Nucl Med Mol Imaging. 2010; 37: 1408-25.
17. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. J Nucl Med. 2009; 50 Suppl 1: 122S-50S.
18. de Langen AJ, Vincent A, Velasquez LM, van Tinteren H, Boellaard R, Shankar LK, et al. Repeatability of 18F-FDG uptake measurements in tumors: a metaanalysis. J Nucl Med. 2012; 53: 701-8.
19. Kwon HR, Pahk K, Park S, Kwon HW, Kim S. Prognostic value of metabolic information in advanced gastric cancer using preoperative 18F-FDG PET/CT. Nucl Med Mol Imaging. 2019; 53: 386-95.
20. Lee SY, Seo HJ, Kim S, Eo JS, Oh SC. Prognostic significance of interim 18 F-fluorodeoxyglucose positron emission tomography-computed tomography volumetric parameters in metastatic or recurrent gastric cancer. Asia Pac J Clin Oncol. 2018; 14: e302-9.
21. Kim SH, Song BI, Kim HW, Won KS, Son YG, Ryu SW. Prognostic value of restaging F-18 fluorodeoxyglucose positron emission tomography/computed tomography to predict 3-year post-recurrence survival in patients with recurrent gastric cancer after curative resection. Korean J Radiol. 2020; 21: 829-37.
22. Wang J, Yu X, Shi A, Xie L, Huang L, Su Y, et al. Predictive value of 18F-FDG PET/CT multi-metabolic parameters and tumor metabolic heterogeneity in the prognosis of gastric cancer. J Cancer Res Clin Oncol. 2023; 149: 14535-47.
23. Chon HJ, Kim C, Cho A, et al. The clinical implications of FDG-PET/CT differ according to histology in advanced gastric cancer. Gastric Cancer. 2019;22:113-22.
24. Park JS, Lee N, Beom SH, Kim HS, Lee CK, Rha SY, et al. The prognostic value of volume-based parameters using 18F-FDG PET/CT in gastric cancer according to HER2 status. Gastric Cancer. 2018; 21: 213-24.
25. Wang JL, Shi AQ, Fan CC, Wang YZ, Guo GR, Liu JY. [Correlations of 18F-FDG PET/CT metabolic parameters and metabolic heterogeneity with human epidermal growth factor receptor 2 expression in patients with gastric cancer]. Zhongguo Yi Xue Ke Xue Yuan Xue Bao. Acta Academiae Medicinae Sinicae. 2022 Aug; 44: 628-35.
26. Acar Tayyar MN, Tamam MÖ, Babacan GB, Şahin MC, Özçevik H, Şengiz Erhan S, et al. The relationship between HER2 status acquired from pathological data and metabolic parameters from pre-treatment [18F]FDG PET/CT in gastric adenocarcinomas. Rev Esp Med Nucl Imagen Mol (Engl Ed). 2025; 44: 500080.
27. Liu Q, Li J, Xin B, Sun Y, Wang X, Song S. Preoperative 18F-FDG PET/CT radiomics analysis for predicting HER2 expression and prognosis in gastric cancer. Quant Imaging Med Surg. 2023; 13: 1537-49.
28. Jiang Y, Yuan Q, Lv W, Xi S, Huang W, Sun Z, et al. Radiomic signature of 18F fluorodeoxyglucose PET/CT for prediction of gastric cancer survival and chemotherapeutic benefits. Theranostics. 2018; 8: 5915-28.
29. Amrane K, Thuillier P, Bourhis D, Le Meur C, Quere C, Leclere JC, et al. Prognostic value of pre-therapeutic FDG-PET radiomic analysis in gastro-esophageal junction cancer. Sci Rep. 2023; 13: 5789.
30. Adili D, Mohetaer A, Zhang W. Diagnostic accuracy of radiomics-based machine learning for neoadjuvant chemotherapy response and survival prediction in gastric cancer patients: a systematic review and meta-analysis. Eur J Radiol. 2024; 173: 111249.
31. Jiang Y, Zhang Z, Wang W, Huang W, Chen C, Xi S, et al. Biology-guided deep learning predicts prognosis and cancer immunotherapy response. Nat Commun. 2023; 14: 5135.
32. Yang L, Chu W, Li M, Xu P, Wang M, Peng M, et al. Radiomics in gastric cancer: First clinical investigation to predict lymph vascular invasion and survival outcome using 18F-FDG PET/CT Images. Front Oncol. 2022; 12: 836098.
33. Zhi H, Xiang Y, Chen C, Zhang W, Lin J, Gao Z, et al. Development and validation of a machine learning-based 18F-fluorodeoxyglucose PET/CT radiomics signature for predicting gastric cancer survival. Cancer Imaging. 2024; 24: 99.
34. Güç ZG, Turgut B, Avci A, Cengiz F, Eren Kalender M, Alacacioğlu A. Predicting pathological response and overall survival in locally advanced gastric cancer patients undergoing neoadjuvant chemotherapy: the role of PET/computed tomography. Nucl Med Commun. 2022; 43: 560-7.
35. Morgagni P, Bencivenga M, Colciago E, Tringali D, Giacopuzzi S, Framarini M, et al. Limited usefulness of 18F-FDG PET/CT in predicting tumor regression after preoperative chemotherapy for noncardia gastric cancer: The Italian Research Group for Gastric Cancer (GIRCG) Experience. Clin Nucl Med. 2020; 45: 177-81.
36. Schneider PM, Eshmuminov D, Rordorf T, Vetter D, Veit-Haibach P, Weber A, et al. 18FDG-PET-CT identifies histopathological non-responders after neoadjuvant chemotherapy in locally advanced gastric and cardia cancer: cohort study. BMC Cancer. 2018; 18: 548.
37. Park S, Ha S, Kwon HW, Kim WH, Kim TY, Oh DY, et al. Prospective Evaluation of Changes in Tumor Size and Tumor Metabolism in Patients with Advanced Gastric Cancer Undergoing Chemotherapy: Association and Clinical Implication. J Nucl Med. 2017; 58: 899-904.
38. Hinzpeter R, Mirshahvalad SA, Kulanthaivelu R, Kohan A, Ortega C, Metser U, et al. Gastro-esophageal cancer: can radiomic parameters from baseline 18F-FDG-PET/CT predict the development of distant metastatic disease? Diagnostics (Basel). 2024; 14: 1205.

Gazi Med J 2026;37(2):230-237

DOI: http://dx.doi.org/10.12996/gmj.2026.4626

Addressing Patient Questions on Spinal Muscular Atrophy: Performance of Large Language Models in a Genetic Context

Hastaların Spinal Musküler Atrofi Hakkındaki Soruları: Genetik Açıdan Büyük Dil Modellerinin Performansı

Dilsu Dicle Erkan[1], Mehmet Alikaşifoğlu[2]

[1]Clinic of Medical Genetics, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Ankara, Türkiye
[2]Department of Medical Genetics, Hacettepe University, Faculty of Medicine, Ankara, Türkiye

ABSTRACT

Objective: Large language models (LLMs) are increasingly used by the public to obtain medical and genetic information. Given the genetic complexity and public health relevance of spinal muscular atrophy (SMA), this study aimed to evaluate the quality, readability, and actionability of LLM-generated responses to SMA-related frequently asked questions (FAQs).

Methods: Fifteen SMA-related FAQs were identified in Turkish using Google's "People Also Ask" feature and categorized into disease definition, genetic screening, and genetic diagnosis and treatment. Each question was submitted to the free versions of ChatGPT, Gemini, and DeepSeek. Responses were evaluated using the modified DISCERN instrument and a 5-point Likert scale for information quality; the Flesch–Kincaid reading ease and grade level for readability; and the Patient Education Materials Assessment Tool (PEMAT) for understandability and actionability.

Results: Median DISCERN scores were 3.00 across all LLMs, indicating moderate information quality, and there was no significant difference among models ($p = 0.069$). Readability differed significantly, with ChatGPT producing responses at a lower Flesch–Kincaid grade level than that of Gemini and DeepSeek ($p = 0.001$). PEMAT understandability and actionability scores varied by question category, with significant differences observed for questions on disease definition and genetic screening ($p < 0.05$).

Conclusion: LLMs generate SMA-related information with moderate quality and understandability; however, variability in readability, actionability, and topic-specific performance limits their suitability

ÖZ

Amaç: Büyük dil modelleri (LLM'ler), tıbbi ve genetik bilgiye erişim amacıyla toplum tarafından giderek daha fazla kullanılmaktadır. Spinal musküler atrofi (SMA) hastalığının anlaması zorlayıcı genetik terimleri ve halk sağlığı açısından önemi göz önüne alındığında, bu çalışmada LLM'ler tarafından üretilen SMA ile ilişkili sıkça sorulan sorulara (SSS) verilen yanıtların kalite, okunabilirlik ve uygulanabilirlik açısından değerlendirilmesi amaçlanmıştır.

Yöntemler: Google'ın "People Also Ask" özelliği kullanılarak 15 adet SMA ile ilişkili Türkçe SSS belirlenmiş ve bu sorular hastalığın tanımı, genetik tarama testleri ve genetik tanı ile tedavi olmak üzere üç başlık altında sınıflandırılmıştır. Her bir soru, ChatGPT, Gemini ve DeepSeek'in ücretsiz sürümleri kullanılarak sırayla sorulmuştur. Yanıtlar; bilgi kalitesi için modifiye DISCERN aracı ve 5 puanlı Likert ölçeği, okunabilirlik için Flesch–Kincaid Okunabilirlik Düzeyi ve Sınıf Seviyesi, anlaşılabilirlik ve uygulanabilirlik için ise Hasta Eğitim Materyalleri Değerlendirme Aracı (PEMAT) kullanılarak değerlendirilmiştir.

Bulgular: Tüm LLM'lerde ortanca DISCERN puanı 3,00 olarak saptanmış ve bu değer orta düzeyde bilgi kalitesine işaret etmiş olup modeller arasında anlamlı fark izlenmemiştir ($p = 0,069$). Okunabilirlik açısından anlamlı farklılıklar saptanmış ve ChatGPT'nin Gemini ve DeepSeek'e kıyasla daha düşük Flesch–Kincaid Sınıf Seviyesi ile yanıtlar ürettiği görülmüştür ($p = 0,001$). PEMAT anlaşılabilirlik ve uygulanabilirlik puanları soru kategorilerine göre farklılık göstermiş; özellikle hastalık tanımı ve genetik tarama sorularında anlamlı farklar saptanmıştır ($p < 0,05$).

Cite this article as: Erkan DD, Alikaşifoğlu M. Addressing patient questions on spinal muscular atrophy: performance of large language Models in a genetic context. Gazi Med J. 2026;37(2):230-237

Address for Correspondence/Yazışma Adresi: Dilsu Dicle Erkan, Clinic of Medical Genetics, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Ankara, Türkiye
E-mail / E-posta: dilsudicle@gmail.com
ORCID ID: orcid.org/0000-0002-2400-8344

Received/Geliş Tarihi: 15.01.2026
Accepted/Kabul Tarihi: 01.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

for standalone use in genetic counseling. While these tools may serve as supplementary educational resources, they should not replace clinician-led genetic counseling, particularly in contexts requiring individualized risk assessment and decision-making.

Keywords: Spinal muscular atrophy, genetic counseling, artificial intelligence, health information quality, large language models, patient education

ÖZ

Sonuç: LLM'ler SMA ile ilişkili bilgileri orta düzeyde kalite ve anlaşılabilirlik ile sunabilmektedir; ancak okunabilirlik, uygulanabilirlik ve konuya özgü performanstaki değişkenlik, bu yanıtların genetik danışmada tek başına kullanımını sınırlamaktadır. LLM'ler tarafından üretilen içerikler tamamlayıcı bir eğitim kaynağı olarak değerlendirilebilir; ancak bireyselleştirilmiş risk değerlendirmesi ve karar verme süreçleri gerektiren durumlarda klinisyen tarafından yürütülen genetik danışmanın yerini almamalıdır.

Anahtar Sözcükler: Spinal musküler atrofi, genetik danışma, yapay zeka, sağlık bilgisi kalitesi, büyük dil modelleri, hasta eğitimi

INTRODUCTION

Spinal muscular atrophy (SMA) is a severe autosomal recessive neuromuscular disorder characterized by progressive degeneration of anterior horn motor neurons (1). SMA can lead to symmetrical muscle weakness, respiratory insufficiency, and, in severe forms, early mortality. The disease is caused by biallelic pathogenic variants in the survival motor neuron-1 (*SMN1*) gene, resulting in a deficiency of the SMN protein. SMA represents one of the most common inherited causes of infant mortality, with an estimated global incidence of approximately 1 in 10,000 live births and a carrier frequency of about 1 in 40–60 individuals (2,3). Clinical severity spans a broad spectrum and is traditionally classified into phenotypic subtypes based on age of onset and achievement of motor milestones, ranging from severe infantile-onset disease to milder later-onset forms.

From a public health and genetic medicine perspective, SMA has a unique position due to the availability of effective disease-modifying therapies and benefits of early diagnosis (4). These developments have increased the importance of timely genetic counseling, carrier screening, and newborn screening programs (5). In Türkiye, SMA has become a major focus of both clinical practice and public awareness. Although the increased focus on SMA in Türkiye is partly related to the high prevalence of consanguineous marriages, population-based data indicate that the elevated SMA carrier frequency cannot be explained by consanguinity alone. Genetic counseling for SMA is further complicated by diagnostic challenges such as "2 + 0" silent carriers, who may be misclassified by standard copy-number testing, creating differences in genetic counseling before or after diagnostic tests rather than screening programs (5,6). Misunderstanding or oversimplification of these concepts may result in false reassurance, unnecessary anxiety, or suboptimal reproductive decision-making. These findings show the need for clear and accurate information when addressing genetically complex conditions, such as SMA.

In parallel with these, the way patients and families seek health information has undergone a substantial transformation (7). Large language models (LLMs) are now widely used by the public to obtain medical information, including explanations of genetic diseases, inheritance patterns, screening tests, and treatment options (7-9). Given the emotional impact of SMA, which often involves reproductive planning, newborn diagnosis, and therapeutic decisions, patients and families increasingly turn to LLMs for rapid, accessible explanations and guidance. In this context, LLM-generated information may function not only as general education but also as a supplement to medical and genetic counseling.

The expanding reliance on LLMs for health information raises critical concerns regarding the quality, accuracy, readability, and actionability of the information provided. Inaccurate or misleading responses about SMA may have significant consequences, particularly in settings where users seek guidance on carrier status, prenatal or preimplantation testing, newborn screening results, or emerging therapies. Despite the growing use of LLMs in other medical specialities, data evaluating the quality of LLM-generated responses specifically for medical genetics remain limited (10-13).

Therefore, assessing the information quality of LLMs in the context of SMA is of particular importance. This study aims to evaluate and compare the quality, readability, and usability of responses generated by the most commonly used LLMs in response to frequently asked questions (FAQs) about SMA.

MATERIALS AND METHODS

Ethical approval was not required for this study because it involved no human participants, patient data, or identifiable personal information. In this study we analysed publicly available online content and responses generated by artificial intelligence (AI) systems.

Question Identification and Categorization

FAQs related to SMA were identified using a structured Google search strategy. Searches were performed using the Google Chrome browser (version 143.0.7499.170) on November 16, 2025, in incognito mode to minimize personalization bias. Initially, common search terms were explored using Google searches, including "spinal muscular atrophy," "SMA," and their Turkish-language equivalents. Among these terms, "SMA" was the most frequently searched keyword and was therefore selected as the primary search term (Figure 1). To capture patient- and public-oriented information needs, the search term "SMA" was entered into the Google search tool (www.google.com.tr), and the questions listed under the "People Also Ask" section were reviewed to identify Turkish-language questions. The "People Also Ask" feature reflects queries that are most commonly searched by users, and has been widely used to represent real-world public interest and information-seeking behavior. A total of 15 non-repetitive FAQs were selected from this section for analysis (Table 1). These questions were chosen to reflect common online information needs of individuals seeking medical and genetic information on SMA.

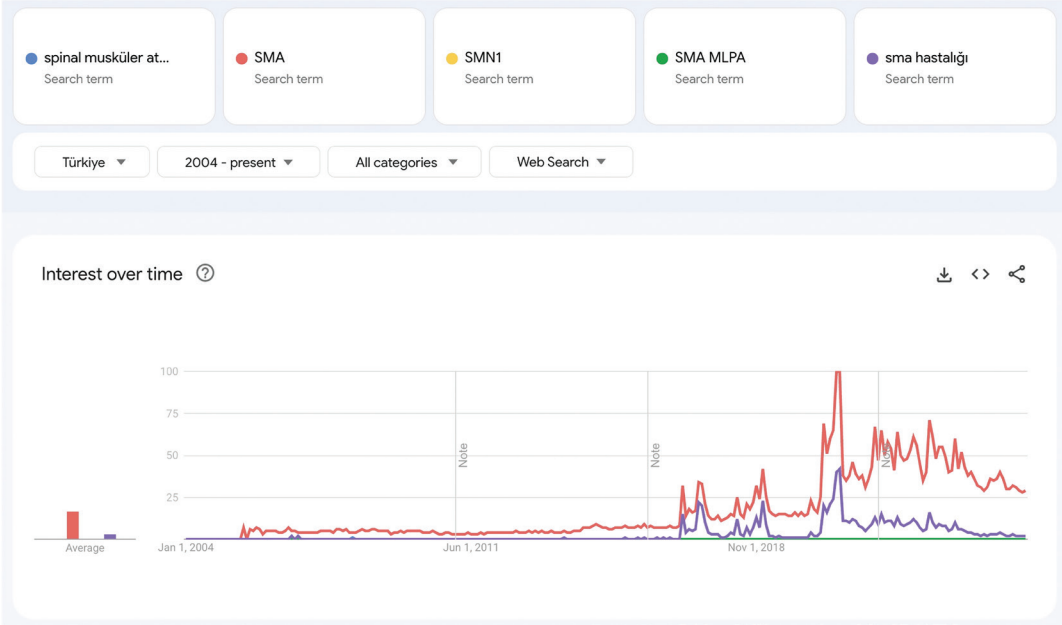

Figure 1. Google Trends analysis showing relative search interest over time in Türkiye for spinal muscular atrophy–related terms, including "spinal musküler atrofi," "SMA," "*SMN1*," "SMA MLPA," and "SMA hastalığı," from 2004 to the present. Search interest is normalized on a scale from 0 to 100, with higher values indicating greater relative search volume. Among the evaluated terms, "SMA" demonstrated the highest and most sustained search interest, supporting its selection as the primary search term for identifying SMA-related frequently asked questions.

MLPA: Multiplex ligation-dependent probe amplification, SMA: Spinal muscular atrophy, *SMN1*: Survival motor neuron-1.

Table 1. Patient-oriented frequently asked questions on spinal muscular atrophy categorized by topic translated to English.

Groups	Questions
Definition of SMA	What does SMA disease mean?
	From whom is the *SMA* gene inherited?
	Is SMA caused by consanguineous marriage?
	Why are there so many patients with SMA in Türkiye?
	Whose children can be affected by SMA?
Genetic screening tests for SMA	How is SMA carrier screening performed before marriage?
	What happens if the SMA test performed by a family physician is positive?
	Why is the SMA test performed only on the male partner?
	What can be done to prevent a baby from having SMA?
	Is SMA tested in the heel-prick blood sample taken from newborns?
Genetic diagnostic tests and gene therapy for SMA	What happens if the SMA test is positive before marriage?
	What happens if the mother or father is an SMA carrier?
	Can SMA be detected during pregnancy?
	Can patients with SMA recover?
	Is *SMA* gene therapy a definitive cure?

SMA: Spinal muscular atrophy.

The selected questions were then categorized into three predefined thematic groups, each comprising five questions: the definition of SMA; genetic screening tests for SMA; and genetic diagnosis and treatment of SMA.

Large Language Models and Generation of Responses

Each of the 15 questions was independently posed to the free versions of three widely accessible LLMs: ChatGPT, Gemini, and DeepSeek. These tools were selected based on reported usage prevalence in Türkiye for public information seeking, while other

platforms (e.g., Copilot, Meta AI, Apple Intelligence) are also available (14). All queries were entered using identical wording for each model. No follow-up prompts, clarifications, or iterative refinements were applied. The initial responses generated by each LLM were recorded verbatim and used for analysis. All responses were collected in Turkish, and the webpage was refreshed before each question. Additionally, questions were collected within the same time period to minimize potential variability related to model updates.

Evaluation Tools and Outcome Measures

Each response generated by the LLMs was independently evaluated by a single board-certified clinical geneticist using multiple established instruments assessing information quality, readability, and usability. Information quality was assessed using the modified DISCERN instrument, which evaluates the reliability and quality of health information, with higher scores indicating better overall quality, and a 5-point Likert scale assessing accuracy and completeness, where higher scores reflect greater concordance with expert consensus (15,16). Readability was analyzed using the Flesch–Kincaid Reading Ease score, in which higher scores indicate easier-to-read text, and the Flesch–Kincaid Grade Level, which estimates the U.S. school grade required to comprehend the material, with lower values indicating greater readability (17). Usability and clarity were further assessed using the Patient Education Materials Assessment Tool (PEMAT), which evaluates whether health information is understandable and actionable for patients, with higher percentages representing better performance in both understandability and actionability domains (18).

For example, the first FAQ, "What does SMA mean?" ("SMA hastalığı ne demek?") was independently queried across all evaluated LLMs. The response generated by ChatGPT correctly defined SMA as a genetic neuromuscular disorder characterized by progressive muscle weakness and atrophy and identified involvement of motor neurons in the spinal cord, autosomal recessive inheritance related to the *SMN1* gene, and common clinical features such as limb weakness, delayed motor milestones, and respiratory involvement (Supplementary Table 1). The response also provided a brief overview of SMA subtypes and current treatment approaches. This response received a DISCERN score of 3, reflecting moderate information quality, as it presented generally accurate and balanced information but did not address uncertainty, alternative sources, or limitations in detail. The Likert score was 4, indicating good accuracy and completeness of a general, patient-oriented explanation. Readability analysis yielded a Flesch–Kincaid Reading Ease score of 46.6 and a Flesch–Kincaid Grade Level of 8.7, indicating that the content requires at least middle-school level reading proficiency. The PEMAT understandability score was 84.61%, indicating that the information was generally clear and comprehensible to patients. However, the PEMAT actionability score was 0% because the response provided descriptive information and did not include guidance on actionable next steps. This evaluation approach was applied to all responses generated by each LLM across the remaining questions.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics version 23 (IBM Corp., Armonk, NY, USA). Normality of data distribution was assessed using the Shapiro-Wilk test. For comparisons of normally distributed scores across three or more question categories, one-way analysis of variance (ANOVA) was applied. For non-normally distributed scores across three or more question categories, the Kruskal-Wallis test was used, with post-hoc multiple comparisons conducted using Dunn's test. Comparisons of normally distributed scores across different AI models were performed using repeated-measures ANOVA, with Bonferroni correction applied for multiple comparisons. For non-normally distributed scores across LLMs, the Friedman test was used, followed by Dunn's test for post-hoc pairwise comparisons. Results are presented as mean ± standard deviation for normally distributed data and as median (minimum-maximum) for non-normally distributed data. A two-sided p-value of < 0.05 was considered statistically significant.

AI-based tools were used to assist with language editing and manuscript preparation; however, all content was reviewed, verified, and finalized by the authors. Ethical approval was not required for this study, as it involved no human participants, patient data, or identifiable personal information.

RESULTS

Overall Performance of Large Language Model Across All Spinal Muscular Atrophy Faqs

Across all 15 SMA-related FAQs, information quality was similar among LLMs. Median DISCERN scores were 3.00 (2.00–3.00) for ChatGPT, 3.00 (2.00–4.00) for Gemini, and 3.00 (0.00–3.00) for DeepSeek, with no significant difference ($p = 0.069$). These scores were clustered around 3.00, suggesting that the information quality of LLM-generated SMA FAQs was moderate. Median Likert-scale scores differed across models ($p = 0.015$), with ChatGPT and Gemini scoring 4.00 (2.00–5.00), and DeepSeek scoring 3.00 (1.00–4.00); post-hoc analyses revealed no significant pairwise differences (Table 2).

Median PEMAT understandability scores were 84.61 (83.33–92.85) for ChatGPT, 91.66 (81.25–92.85) for Gemini, and 84.61 (73.33–92.85) for DeepSeek ($p = 0.786$). Median PEMAT actionability scores were identical across models at 60.00 ($p = 0.105$). Median Flesch–Kincaid Reading Ease scores were 64.30 (22.80–81.90), 59.40 (39.90–69.50), and 58.00 (46.60–75.20) for ChatGPT, Gemini, and DeepSeek, respectively ($p = 0.344$).

Mean Flesch–Kincaid Grade Level scores differed significantly among models ($p = 0.001$), with lower values for ChatGPT (6.64 ± 2.17) compared with Gemini (8.37 ± 1.56) and DeepSeek (8.12 ± 1.19); no difference was observed between Gemini and DeepSeek (Figure 2e).

Language Model Across Performance Across Spinal Muscular Atrophy Question Categories

Within individual question categories, no significant differences in DISCERN or Likert scores were observed among LLMs. However, within-model analyses demonstrated significant category-based differences for DeepSeek in both DISCERN ($p = 0.034$) and Likert scores ($p = 0.032$), whereas ChatGPT and Gemini showed no such differences (Table 3).

PEMAT understandability scores differed across models for SMA disease-definition questions ($p = 0.015$), with medians of 92.30

(84.61–92.30) for ChatGPT, 92.30 (84.61–92.85) for Gemini, and 76.92 (73.33–83.33) for DeepSeek. A significant overall difference among the models was observed for genetic screening test questions (p = 0.034), whereas no difference was found for genetic diagnosis and treatment questions (p = 0.128).

For SMA disease definition questions, mean Flesch–Kincaid Reading Ease scores differed among models (p = 0.020), with values of 70.24 ± 14.03 for ChatGPT, 57.26 ± 10.90 for Gemini, and 58.12 ± 5.91 for DeepSeek. Mean Flesch–Kincaid Grade Level scores between disease-definition questions (p = 0.002) and genetic screening test questions (p = 0.007), with lower grade levels observed for ChatGPT in both categories. No significant readability differences were observed for genetic diagnosis and treatment questions.

PEMAT actionability scores did not differ among models within any category. However, significant within-model differences across categories were observed for all three AI tools (ChatGPT p = 0.013; Gemini p = 0.004; DeepSeek p = 0.023), and the lowest actionability scores were consistently found for SMA disease-definition questions.

DISCUSSION

This study evaluated the performance of LLMs in generating responses to SMA–related FAQs from a clinical genetics perspective. Across all evaluated metrics, overall information quality and understandability did not differ significantly among ChatGPT, Gemini, and DeepSeek. In contrast, readability differed across models, with ChatGPT producing responses at a lower reading grade level. Actionability varied by FAQ

Table 2. Comparison of responses generated by artificial intelligence tools according to DISCERN, Likert, PEMAT, and Flesch-Kincaid readability metrics.

	ChatGPT	Gemini	DeepSeek	Test statistic	p-value
DISCERN score	3.00 (2.00 – 3.00)	3.00 (2.00 – 4.00)	3.00 (0.00 – 3.00)	5.353	0.069[x]
Likert scale	4.00 (2.00 – 5.00)	4.00 (2.00 – 5.00)	3.00 (1.00 – 4.00)	8.359	**0.015**[x]
PEMAT understandability score	84.61 (83.33 – 92.85)	91.66 (81.25 – 92.85)	84.61 (73.33 – 92.85)	0.481	0.786[x]
Flesch–Kincaid reading ease score	64.30 (22.80 – 81.90)	59.40 (39.90 – 69.50)	58.00 (46.60 – 75.20)	2.133	0.344[x]
Flesch–Kincaid grade level	6.64 ± 2.17[a]	8.37 ± 1.56[b]	8.12 ± 1.19[b]	8.328	**0.001**[y]
PEMAT actionability	60.00 (0.00 – 100.00)	60.00 (0,00 – 80.00)	60.00 (0.00 – 83.33)	4.514	0.105[x]

AI: Artificial intelligence.
[x]Friedman test, [y]Repeated-measures analysis of variance.
Data are presented as median (minimum-maximum) or mean ± standard deviation.
[a-b]AI tools sharing the same superscript letter do not differ significantly.

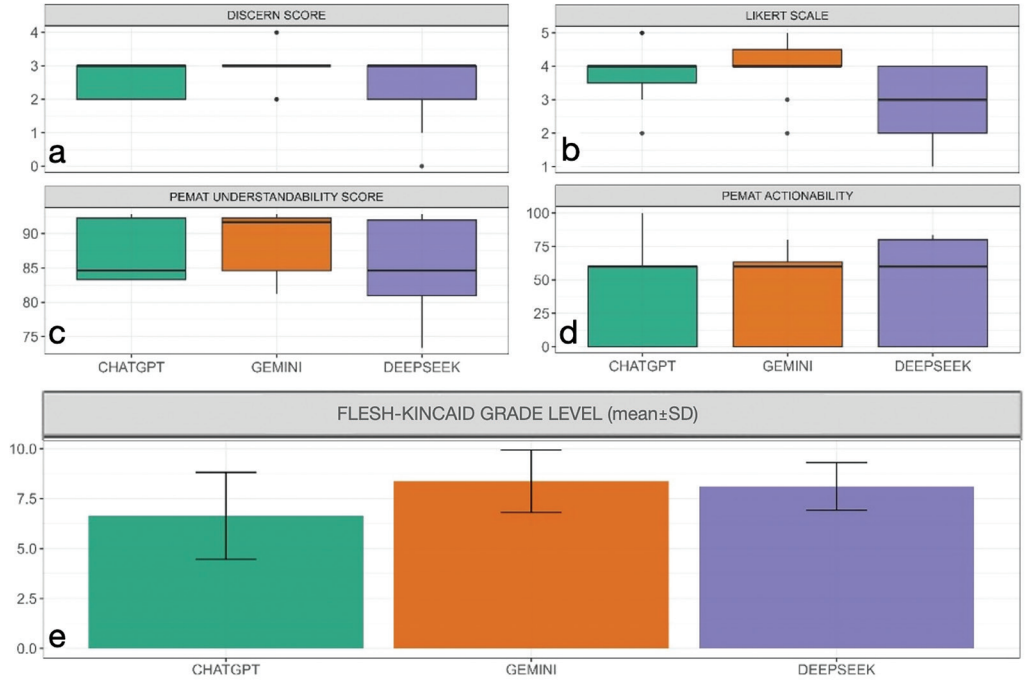

Figure 2. Comparison of large language model performance across multiple evaluation metrics for spinal muscular atrophy related frequently asked questions. (a) Distribution of DISCERN scores, b) Distribution of Likert scale scores, c) PEMAT understandability scores, d) PEMAT actionability scores, e) Mean Flesch–Kincaid Grade Level scores (± standard deviation). Box plots display medians, interquartile ranges, and ranges, while bars represent mean values with standard deviation where indicated. Comparisons are shown for ChatGPT, Gemini, and DeepSeek.

category, with higher scores observed for screening, diagnosis, and treatment-related questions than for disease definition questions. These findings indicate that LLM-generated responses to SMA FAQs vary according to both the model used and the type of question asked.

The absence of significant differences in information quality scores across LLMs reflects a broader limitation of current generative models in addressing complex genetic conditions such as SMA. Because genetic counseling involves important concepts including carrier status, inheritance risk, and genotype–phenotype variability, similar information quality scores do not necessarily translate into equivalent clinical usefulness (19,20). Although LLMs may support preliminary information seeking, variability in readability, actionability, and topic-specific performance shows the need for clinician led contextualization (21-23).

Readability is a determinant of patient understanding in SMA, where counseling is repeated across the lifespan and often occurs under emotional stress and limited health literacy. Differences in

Table 3. Comparison of responses generated by artificial intelligence tools according to DISCERN, Likert, PEMAT, and Flesch-Kincaid readability metrics within and across question groups.

Score	AI tool	Definition of SMA	Genetic screening tests for SMA	Genetic diagnostic tests and gene therapy for SMA	Test statistic	p-value
DISCERN score	ChatGPT	3.00 (2.00–3.00)	3.00 (2.00–3.00)	2.00 (2.00–3.00)	1.556	0.459[t]
	Gemini	3.00 (2.00–3.00)	3.00 (3.00–3.00)	3.00 (2.00–4.00)	1.400	0.497[t]
	DeepSeek	3.00 (3.00–3.00)[a]	2.00 (0.00–3.00)[b]	3.00 (2.00–3.00)[ab]	6.746	**0.034**[t]
	Test statistic	1.000	4.800	4.308		
	p-value[x]	0.607	0.091	0.116		
Likert scale	ChatGPT	4.00 (3.00–5.00)	4.00 (3.00–4.00)	4.00 (2.00–5.00)	1.173	0.556[t]
	Gemini	4.00 (3.00–5.00)	4.00 (4.00–5.00)	4.00 (2.00–5.00)	0.410	0.815[t]
	DeepSeek	4.00 (4.00–4.00)	3.00 (1.00–4.00)	2.00 (2.00–4.00)	6.863	**0.032**[t]
	Test statistic	0.545	5.375	3.500		
	p-value[x]	0.761	0.068	0.174		
PEMAT understandability score	ChatGPT	92.30 (84.61–92.30)[AB]	83.33 (83.33–84.61)	92.30 (83.33–92.85)	7.198	**0.027**[t]
	Gemini	92.30 (84.61–92.85)[A]	91.66 (84.61–92.85)	83.33 (81.25–92.30)	3.490	0.175[t]
	DeepSeek	76.92 (73.33–83.33)[aB]	85.71 (83.33–92.30)[ab]	91.66 (84.61–92.85)[b]	9.403	**0.009**[t]
	Test statistic	8.444	6.778	4.111[a]		
	p-value[x]	**0.015**	**0.034**	0.128		
Flesch-Kincaid reading ease score	ChatGPT	70.24 ± 14.03[A]	64.38 ± 0.84	46.20 ± 18.11	2.629	0.160[z]
	Gemini	57.26 ± 10.90[B]	60.88 ± 4.07	58.82 ± 9.86	0.213	0.811[z]
	DeepSeek	58.12 ± 5.91[AB]	60.38 ± 9.06	56.10 ± 7.04	0.413	0.671[z]
	Test statistic	6.695	0.791	4.357		
	p-value[x]	**0.020**	0.486	0.104		
Flesch-Kincaid grade level	ChatGPT	5.42 ± 2.01[A]	6.04 ± 0.33[A]	8.46 ± 2.44	3.837	0.051[z]
	Gemini	9.16 ± 1.79[B]	7.78 ± 0.80[B]	8.16 ± 1.85	1.048	0.381[z]
	DeepSeek	8.44 ± 0.86[AB]	7.56 ± 0.97[AB]	8.36 ± 1.64	0.815	0.466[z]
	Test statistic	15.259	9.629	0.124		
	p-value[x]	**0.002**	**0.007**	0.885		
PEMAT actionability score	ChatGPT	0.00 (0.00–0.00)[a]	60.00 (60.00–60.00)[b]	60.00 (0.00–100.00)[b]	8.714	**0.013**[t]
	Gemini	0.00 (0.00–0.00)[a]	66.66 (60.00 –80.00)[b]	60.00 (60.00–80.00)[b]	10.932	**0.004**[t]
	DeepSeek	0.00 (0.00–60.00)[a]	80.00 (60.00–83.33)[b]	80.00 (0,00–83.33)[ab]	7.541	**0.023**[t]
	Test statistic	2.000	5.765	0.400		
	p-value[x]	0.368	0.056	0.819		

[x]Friedman test, [y]Repeated-measures analysis of variance, [z]One-way ANOVA, [t]Kruskal-Wallis test.
Data are presented as median (minimum-maximum) or mean ± standard deviation.
[ab]Within each AI tool, question groups sharing the same lowercase letter do not differ significantly.
[AB]Within each question group, AI tools sharing the same uppercase letter do not differ significantly.
SMA: Spinal muscular atrophy, AI: Artificial intelligence, ANOVA: Analysis of variance

readability across LLMs indicate that linguistic style and sentence structure influence patients' comprehension of concepts. Variability in LLM performance across SMA FAQ categories likely reflects differences in underlying genetic complexity: disease definition questions are largely descriptive, whereas screening, diagnostic, and treatment-related questions require integration of inheritance, test interpretation, risk assessment, and evolving therapeutic options (21-23). Actionability similarly depends more on question intent than on model characteristics, as definition-focused questions inherently offer limited guidance, while screening and treatment related questions align more closely with clinical decision making and follow-up processes (18,21,23).

Implications for Clinical Genetic Counseling Practice in the Era of Large Language Models

Genetic counseling is one of the core components of clinical genetics practice, and its role becomes increasingly critical as patients seek genetic information and test result evaluation from AI-based tools. Counseling often involves supporting patients through uncertainty, facilitating informed decision making by addressing expectations, and emotional responses unique to each individual or family rather than directing patients toward a single predefined choice (24,25). Such individualized, bidirectional communication remains inherently human and may not be consistently achieved by LLM-based tools, which generate standardized responses without access to personal context or psychosocial cues. Clinical geneticists are now more likely to encounter patients who present after consulting LLMs and who may have acquired false, incomplete, or oversimplified information. In such cases, counseling should begin with identifying and correcting misinformation and clarifying unmet informational needs. This reinforces the importance of structured pretest and posttest genetic counseling and a strong patient-physician relationship grounded in clear communication and trust.

In addition, genetic risk assessment is highly individualized and cannot always be generalized using population-based information. For example, genetic screening tests are designed to detect the most common pathogenic variants within a population to maximize efficiency and cost-effectiveness, but they do not capture all disease-causing mechanisms. In SMA, rare scenarios such as silent carrier status, pathogenic single-nucleotide variants not included in standard screening panels, or *de novo* mutations may not be identified through routine carrier screening (26). Consequently, identical screening results may carry different residual risks across individuals and families. Communicating these individualized risks is a core component of genetic counseling, and it extends beyond the current capabilities of AI-generated responses.

Effective genetic counseling in this context requires more than a classical verbal explanation. New supportive tools should be integrated into counseling practice, including disease-specific visual materials to explain inheritance patterns, carrier states such as 2 + 0 silent carriers, and the distinctions between screening and diagnostic tests. These tools may improve understanding of complex genetic concepts that are frequently misinterpreted in AI-generated responses. Genetic counseling should also be recognized as a longitudinal and repetitive process, requiring sufficient time and multiple sessions to allow patients and families to process information, consider options, and make informed decisions. As external information sources become more prevalent, counseling may need to be more detailed and iterative than previously, continuing until patients and families demonstrate sufficient understanding of their genetic risks and choices.

An additional consideration concerns the governance of LLM platforms , particularly those used for training and deployment. For content addressing genetic screening, diagnosis, or treatment, LLM outputs could be regulated at the platform level to include standardized prompts encouraging users to seek clinician input for individualized risk assessment and final interpretation. Incorporating safeguards such as explicit statements directing users to consult healthcare professionals for definitive results may help reduce the risks of misinterpretation.

Several limitations should be considered when interpreting these findings. First, the analysis was based on a limited set of SMA-related FAQs obtained from Google search results at a single time point, which may not capture the full spectrum of information needs encountered in clinical genetics practice. Second, all questions and LLM-generated responses were evaluated in Turkish, and the findings may reflect language-specific characteristics of both the models and the assessment tools, as the LLMs are mostly trained in English. Third, LLM outputs are dynamic and subject to change with model updates, retraining, and prompt sensitivity, which may affect reproducibility over time. Fourth, the evaluation relied on expert-based assessment instruments that, while validated for health information appraisal, were not specifically designed to assess disease-specific genetic accuracy or the adequacy of counseling. Finally, the study focused on static written responses and did not account for interactive dialogue or individualized context, which are essential components of real-world genetic counseling encounters.

CONCLUSION

LLMs can generate generally consistent responses to SMA-related FAQs; however, their performance varies with readability, actionability, and question category. While overall information quality and understandability were similar across models, differences in linguistic accessibility and topic-specific response characteristics show important limitations for direct clinical use. These findings suggest that LLM-generated SMA information may serve as a supplementary educational resource today, but should not replace clinician-led genetic counseling, particularly in contexts requiring individualized risk assessment, interpretation of genetic test results, and shared decision-making.

Ethics

Ethics Committee Approval: Ethical approval was not required for this study, as it involved no human participants, patient data, or identifiable personal information.

Informed Consent: No informed consent is required for this study as we analysed publicly available online content and responses generated by AI systems.

Footnotes

Authorship Contributions

Concept: D.D.E., Design: D.D.E., M.A., Data Collection or Processing: D.D.E., Analysis or Interpretation: D.D.E., Literature Search: D.D.E., Writing: D.D.E., M.A.

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

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

Supplementary Table 1: https://d2v96fxpocvxx.cloudfront.net/c81cd83d-6f66-4047-b292-9f60f894af26/documents/Supplementary%20Table%201.xlsx

REFERENCES

1. Keinath MC, Prior DE, Prior TW. Spinal Muscular Atrophy: Mutations, Testing, and Clinical Relevance. Appl Clin Genet. 2021; 14: 11-25.
2. Prior TW. Carrier screening for spinal muscular atrophy. Genet Med. 2008; 10: 840-2.
3. Verhaart IEC, Robertson A, Wilson IJ, Aartsma-Rus A, Cameron S, Jones CC, et al. Prevalence, incidence and carrier frequency of 5q-linked spinal muscular atrophy - a literature review. Orphanet J Rare Dis. 2017; 12: 124.
4. Cooper K, Nalbant G, Sutton A, Harnan S, Thokala P, Chilcott J, et al. Systematic review of presymptomatic treatment for spinal muscular atrophy. Int J Neonatal Screen. 2024; 10: 56
5. Prior TW, Leach ME, Finanger EL. Spinal muscular atrophy. In: Adam MP, Bick S, Mirzaa GM, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-. 2000 Feb 24 [updated 2026 Feb 12]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1352/
6. Milligan JN, Blasco-Pérez L, Costa-Roger M, Codina-Solà M, Tizzano EF. Recommendations for interpreting and reporting silent carrier and disease-modifying variants in SMA testing workflows. Genes (Basel). 2022; 13: 1657.
7. Shahsavar Y, Choudhury A. User intentions to use ChatGPT for self-diagnosis and health-related purposes: cross-sectional survey study. JMIR Hum Factors. 2023; 10: e47564.
8. Kamatani Y, Kaname T. Artificial intelligence in medical genomics. J Hum Genet. 2024; 69: 475.
9. Yun HS, Bickmore T. Online health information-seeking in the era of large language models: cross-sectional web-based survey study. J Med Internet Res. 2025; 27: e68560.
10. Ayik G, Kolac UC, Aksoy T, Yilmaz A, Sili MV, Tokgozoglu M, et al. Exploring the role of artificial intelligence in Turkish orthopedic progression exams. Acta Orthop Traumatol Turc. 2025; 59: 18-26.
11. Kolac UC, Karademir OM, Ayik G, Kaymakoglu M, Familiari F, Huri G. Can popular AI large language models provide reliable answers to frequently asked questions about rotator cuff tears? JSES Int. 2025; 9: 390-7.
12. Mehmet S, et al. Comparative evaluation of large language models in addressing autism-related information queries: insights from ChatGPT, Gemini, and Copilot. Gazi Med J. 2025; 36: 407-16.
13. Volk SC, Schäfer MS, Lombardi D, Mahl D, Yan X. How generative artificial intelligence portrays science: Interviewing ChatGPT from the perspective of different audience segments. Public Underst Sci. 2025; 34: 132-53.
14. Küçüksabanoğlu ZK. Yapay Zeka Politikaları Derneği (AIPA) Gelecek Araştırması: Toplumda yapay zeka algısı. Yapay Zeka Politikaları Derneği (AIPA); 2025. Available from: https://aipaturkey.org/media/pdfs/YapayZeka_AlgiArastirmasi_2025_Toplum.pdf
15. Likert R. A technique for the measurement of attitudes. Arch Psychol. 1932;140:1-55. Available from: https://legacy.voteview.com/pdf/Likert_1932.pdf
16. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health. 1999; 53: 105-11.
17. Kincaid JP, Fishburne RP Jr, Rogers RL, Chissom BS. Derivation of new readability formulas: automated Readability Index, Fog Count and Flesch Reading Ease Formula for Navy enlisted personnel. Millington (TN): Naval Technical Training Command, Research Branch; 1975 Feb. Available from: https://stars.library.ucf.edu/istlibrary/56/.
18. Shoemaker SJ, Wolf MS, Brach C. Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. Patient Educ Couns. 2014. 96: 395-403.
19. Jeon S, Lee SA, Chung HS, Yun JY, Park EA, So MK, et al. Evaluating the use of generative artificial intelligence to support genetic counseling for rare diseases. Diagnostics (Basel). 2025; 15: 672.
20. Ahimaz P, Bergner AL, Florido ME, Harkavy N, Bhattacharyya S. Genetic counselors' utilization of ChatGPT in professional practice: a cross-sectional study. Am J Med Genet A. 2024; 194: e63493.
21. Brett GR, Ward A, Bouffler SE, Palmer EE, Boggs K, Lynch F, et al. Co-design, implementation, and evaluation of plain language genomic test reports. NPJ Genom Med. 2022; 7: 61.
22. Farmer GD, Gray H, Chandratillake G, Raymond FL, Freeman ALJ. Recommendations for designing genetic test reports to be understood by patients and non-specialists. Eur J Hum Genet. 2020; 28: 885-95.
23. Haga SB, Mills R, Pollak KI, Rehder C, Buchanan AH, Lipkus IM, et al. Developing patient-friendly genetic and genomic test reports: formats to promote patient engagement and understanding. Genome Med. 2014; 6: 58.
24. Coen E, Del Fiol G, Kaphingst KA, Borsato E, Shannon J, Smith H, et al. Chatbot for the return of positive genetic screening results for hereditary cancer syndromes: prompt engineering project. JMIR Cancer. 2025; 11: e65848.
25. Meekins-Doherty L, Dive L, McEwen A, Sexton A. Generative AI and the profession of genetic counseling. J Genet Couns. 2025; 34: e2009.
26. Tıbbi Genetik Derneği. Spinal Musküler Atrofi El Kitabı [Internet]. Ankara: Tıbbi Genetik Derneği; 2024. 20 p. Available from: https://yonetim.citius.technology/files/kurum/kurum75/menu/spinal-muskuler-atrofi-el-kitabi--bitmis-hali-15x23-olacak---1-.pdf

Gazi Med J 2026;37(2):238-243

DOI: http://dx.doi.org/10.12996/gmj.2026.4602

Investigation of HFE Mutations Associated with Hemochromatosis the Case of Ordu Province

Hemokromatozis ile İlişkili HFE Mutasyonlarının Araştırılması: Ordu İli Örneği

Çağrı Doğan[1], Aslı Odabaşı Giden[2], Muhammet Özbilen[3]

[1]Department of Medical Genetics, Ordu University, Faculty of Medicine, Ordu, Türkiye
[2]Department of Hematology, Ordu State Hospital, Türkiye
[3]Department of Internal Medicine, Ordu University, Faculty of Medicine, Ordu, Türkiye

ABSTRACT

Objective: Hereditary hemochromatosis (HH) is a common autosomal recessive disorder characterised by increased intestinal iron absorption and progressive systemic iron overload. Although the C282Y variant is the predominant cause of HH in Northern European populations, the distribution of *HFE* variants varies geographically, and data from different regions of Türkiye remain limited. This study aimed to evaluate the molecular spectrum of *HFE* gene variants and their biochemical correlates in patients investigated for HH in Ordu province, located in the Black Sea region of Türkiye.

Methods: A retrospective analysis was conducted of 100 patients who underwent *HFE* gene sequencing between January 2024 and October 2025. Demographic characteristics, hemoglobin, serum ferritin, serum iron, transferrin saturation (TS), and *HFE* genotypes were obtained from medical records. PCR-based amplification followed by next-generation sequencing on the Illumina MiSeq platform was performed. Variants were classified using international databases, including ClinVar, Franklin, VarSome, HGMD Public®, gnomAD, and dbSNP. Comparisons between variant carriers and non-carriers were analysed using the Mann–Whitney U and chi-square tests.

Results: Among the 1000 patients (53 males, 47 females; mean age 47.2 ± 17.1 years), 64 (64.0%) exhibited a wild-type *HFE* genotype, while 36 (36.0%) carried at least one variant. The c.187C>G (H63D) variant was the most common, with 30 heterozygotes and 3 homozygotes, followed by one case each of C282Y homozygosity, c.76 + 2 dup heterozygosity, and H63D + c.76+2 dup compound heterozygosity. Ferritin, serum iron, and hemoglobin levels did not differ significantly between variant

ÖZ

Amaç: Herediter hemokromatozis (HH), artmış intestinal demir emilimi ve ilerleyici sistemik demir yüklenmesi ile karakterize yaygın bir otozomal resesif hastalıktır. C282Y varyantı Kuzey Avrupa popülasyonlarında HH'nin baskın nedeni olmakla birlikte, HFE varyantlarının dağılımı coğrafi olarak değişkenlik göstermektedir ve Türkiye'nin farklı bölgelerine ait veriler halen sınırlıdır. Bu çalışma, Türkiye'nin Karadeniz Bölgesi'nde yer alan Ordu ilinde HH açısından araştırılan hastalarda *HFE* geni varyantlarının moleküler spektrumunu ve bunların biyokimyasal korelatlarını değerlendirmeyi amaçlamıştır.

Yöntemler: Ocak 2024 ile Ekim 2025 tarihleri arasında *HFE* gen dizilemesi yapılan 100 hastanın retrospektif analizi gerçekleştirildi. Demografik özellikler, hemoglobin, serum ferritin, serum demiri, transferrin satürasyonu (TS) ve *HFE* genotipleri tıbbi kayıtlardan elde edildi. PCR temelli amplifikasyonun ardından Illumina MiSeq platformunda yeni nesil dizileme uygulandı. Varyantlar; ClinVar, Franklin, VarSome, HGMD Public®, gnomAD ve dbSNP dâhil uluslararası veri tabanları kullanılarak sınıflandırıldı. Varyant taşıyıcıları ile taşıyıcı olmayanlar arasındaki karşılaştırmalar Mann-Whitney U ve ki-kare testleri ile analiz edildi.

Bulgular: Yüz hastanın (53 erkek, 47 kadın; ortalama yaş 47,2 ± 17,1 yıl) 64'ünde (%64,0) vahşi tip *HFE* genotipi saptanırken, 36'sında (%36,0) en az bir varyant bulundu. En sık görülen varyant, 30 heterozigot ve 3 homozigot olgu ile c.187C>G (H63D) idi. Bunu birer olgu şeklinde C282Y homozigotluğu, c.76+2dup heterozigotluğu ve H63D + c.76+2dup bileşik heterozigotluğu izledi. Ferritin, serum demiri ve hemoglobin düzeyleri açısından varyant taşıyıcıları ile taşıyıcı olmayanlar arasında

Cite this article as: Doğan Ç, Odabaşı Giden A, Özbilen M. Investigation of HFE mutations associated with hemochromatosis the case of Ordu province. Gazi Med J. 2026;37(2):238-243

Address for Correspondence/Yazışma Adresi: Çağrı Doğan, Department of Medical Genetics, Ordu University Faculty of Medicine, Ordu, Türkiye
E-mail / E-posta: mdgencagri@gmail.com
ORCID ID: orcid.org/0000-0001-5068-0380

Received/Geliş Tarihi: 15.01.2026
Accepted/Kabul Tarihi: 01.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

carriers and non-carriers (p = 0.869, 0.204, and 0.674, respectively). However, elevated TS (> 45%) was observed significantly more often in variant carriers than in wild-type individuals (41.7% vs. 20.3%, p = 0.040), indicating an approximately 2.8-fold higher likelihood of high TS among carriers. The proportion of participants with ferritin > 400 µg/L did not differ significantly between groups. The highest ferritin level (6104 µg/L) was observed in a heterozygous H63D female patient with severe anemia, while the highest TS (99%) was detected in a patient without any detectable variants.

Conclusion: In this Black Sea cohort, H63D was the predominant *HFE* variant, whereas C282Y was rare, consistent with regional genetic patterns in Türkiye. Although *HFE* variants were not associated with significant differences in ferritin, their strong association with elevated TS suggests a measurable impact on iron handling. These findings highlight the importance of TS in the diagnostic evaluation of HH in regions where C282Y-related HH is uncommon. Larger multicenter studies incorporating imaging-based iron quantification are needed to better define genotype–phenotype relationships in the Turkish population.

Keywords: Hereditary hemochromatosis, *HFE* gene, iron overload, transferrin saturation

ÖZ

anlamlı fark saptanmadı (sırasıyla p = 0,869, 0,204 ve 0.674). Bununla birlikte, yüksek TS düzeyi (> %45) varyant taşıyıcılarında vahşi tip bireylere göre anlamlı olarak daha sık gözlendi (%41,7'ye karşı %20,3, p = 0,040); bu durum, taşıyıcılarda yüksek TS görülme olasılığının yaklaşık 2.8 kat daha fazla olduğunu göstermektedir. Ferritin düzeyinin > 400 µg/L olma oranı açısından gruplar arasında anlamlı fark bulunmadı. En yüksek ferritin düzeyi (6104 µg/L), ağır anemisi olan H63D heterozigot kadın bir hastada saptanırken; en yüksek TS değeri (%99), saptanabilir herhangi bir varyantı olmayan bir hastada gözlendi.

Sonuç: Bu Karadeniz kohortunda H63D başlıca HFE varyantı olarak saptanırken, C282Y'nin nadir olduğu görüldü; bu bulgu Türkiye'deki bölgesel genetik örüntülerle uyumludur. HFE varyantları ferritin düzeylerinde anlamlı farklılıklarla ilişkili olmasa da, yüksek TS ile gösterdikleri güçlü ilişki demir metabolizması üzerinde ölçülebilir bir etkilerine işaret etmektedir. Bu bulgular, C282Y ile ilişkili HH'nin nadir olduğu bölgelerde HH'nin tanısal değerlendirmesinde TS'nin önemini vurgulamaktadır. Türk popülasyonunda genotip-fenotip ilişkilerini daha iyi tanımlayabilmek için görüntüleme temelli demir kantifikasyonunu içeren daha geniş, çok merkezli çalışmalara ihtiyaç vardır.

Anahtar Sözcükler: Herediter hemokromatozis, *HFE* geni, demir birikimi, transferrin satürasyonu

INTRODUCTION

Hereditary hemochromatosis (HH) is one of the most prevalent autosomal recessive metabolic disorders worldwide. It is characterised by excessive intestinal iron absorption and progressive parenchymal iron deposition in multiple organs, primarily the liver, pancreas, heart, joints, and pituitary gland (1). The resultant iron overload leads to tissue injury and, if untreated, may cause cirrhosis, diabetes mellitus, cardiomyopathy, hypogonadism, arthropathy, and even hepatocellular carcinoma (HCC) (2). Although environmental and secondary causes of iron overload exist, HH represents the major genetic form, with its pathogenesis predominantly linked to pathogenic variants in the *HFE* gene (2).

The *HFE* gene, located on chromosome 6p21.3 within the major histocompatibility complex region, spans approximately 12 kb and contains six exons encoding a 343-amino-acid transmembrane glycoprotein (3). The HFE protein has been demonstrated to interact with the transferrin receptor 1 (TfR1), thereby playing a pivotal role in modulating hepcidin expression. This, in turn, is the central regulator of systemic iron homeostasis (4). Under normal physiological conditions, the HFE–TfR1 complex regulates iron uptake by hepatocytes and enterocytes through feedback signalling that limits dietary iron absorption. Disruption of this pathway through *HFE* variants leads to insufficient hepcidin synthesis, unchecked intestinal iron absorption, and progressive systemic iron overload (5).

The most prevalent pathogenic *HFE* variants are *c.845G>A(p. Cys282Tyr, C282Y) and c.187C>G(p.His63Asp, H63D),* which account for the vast majority of cases, especially in populations of Northern European descent (6). The C282Y variant has been demonstrated to disrupt a critical disulfide bond that is necessary for the proper folding and cell-surface expression of the *HFE* protein. This results in impaired TfR1 binding and dysregulated iron sensing (7). Homozygosity for the *C282Y* allele has been demonstrated to be responsible for over 80–90% of clinically diagnosed HH cases in Western populations, whereas compound heterozygosity for C282Y/H63D has been shown to be associated with a milder, often incomplete phenotype (8). The *H63D* variant, while more common globally, is generally considered a low-penetrance variant that contributes to iron overload only in the presence of other genetic or environmental risk factors such as alcohol intake, viral hepatitis, or metabolic syndrome (9).

In addition to these classical variants, a number of rare or population-specific variants in *HFE* and non-*HFE* genes, including *TFR2, HAMP, HJV, and SLC40A1,* have been reported to cause distinct subtypes of HH with variable clinical penetrance (10). Studies from Southern Europe, the Mediterranean basin, and the Middle East have revealed a more diverse variant spectrum, underscoring the importance of regional genetic screening and local genotype–phenotype correlation analyses (10).

The prevalence of HH varies considerably across populations and ethnic groups. The C282Y allele frequency has been estimated at approximately 6–10% in Northern European populations, 2–4% in Southern European populations, and less than 1% in Asian and African populations, reflecting its well-established Northwestern European origin and founder effect. In contrast, studies conducted in Türkiye have consistently demonstrated that the H63D variant is the predominant *HFE* variant, whereas C282Y remains relatively rare. This distribution pattern likely reflects the complex population genetics, historical migration routes, and admixture events that have shaped the genetic landscape of Anatolia. Nevertheless, comprehensive data regarding the molecular spectrum and regional variability of *HFE* variants within different Turkish subpopulations remain limited. These limitations underscores the importance of region-specific genetic studies aimed at clarifying both carrier frequencies and the clinical relevance of *HFE* variants in the Turkish population. Clinically, HH displays substantial variability in onset age, biochemical profile, and organ involvement, even among

individuals harbouring identical genotypes (11). This heterogeneity reflects the interplay among genetic background, dietary factors, alcohol use, and comorbid conditions. Early recognition through genetic screening is therefore essential, as therapeutic phlebotomy when initiated before irreversible organ damage has occurred can effectively normalize iron levels and prevent complications (2,11). In recent years, next-generation sequencing and expanded molecular panels have greatly improved the detection of both common and rare *HFE* variants, facilitating earlier diagnosis and more precise classification of hereditary iron overload syndromes (10).

In view of these observations, a systematic evaluation of *HFE* gene variants within distinct geographic regions has the potential to enhance understanding of local variant profiles and their clinical correlates. The present study aims to determine the variant detection rate of the *HFE* gene am*ong* patients referred for suspected HH in Ordu Province and its surrounding region, to analyse their demographic and clinical characteristics, and to contribute data to the scientific literature.

MATERIALS AND METHODS

A retrospective analysis was conducted of the medical records of patients who attended the Ordu University outpatient clinic between January 1, 2024, and October 1, 2025, with a preliminary diagnosis of hemochromatosis. All patients for whom *HFE* gene sequencing was requested were included in the study. This cohort represents a referral-based screening population rather than clinically confirmed HH cases. All consecutive patients who were referred with a preliminary clinical suspicion of HH and for whom *HFE* gene analysis was requested were included, regardless of whether they ultimately fulfilled established diagnostic criteria for HH based on biochemical or clinical parameters. Therefore, the study reflects the real-world diagnostic spectrum of patients evaluated for suspected HH in our region. However, the referral-based inclusion strategy may introduce selection bias and should be considered a methodological limitation when interpreting genotype–phenotype associations. The primary objective was to evaluate the frequency and distribution of *HFE* variants and to assess their association with iron-related biochemical parameters in a real-world referral cohort. The clinical findings and the variants identified in the patients were described. In the present study, patients' ages, sex, hemoglobin levels, ferritin levels, serum iron levels, transferrin saturation (TS), and *HFE* gene *anal*ysis results were obtained by retrospective review of their medical records.

A total of 100 patients who were referred to the Department of Medical Genetics at Ordu University Hospital with a preliminary diagnosis of hemochromatosis were evaluated. Peripheral blood samples were collected in EDTA tubes for DNA extraction.

The genomic DNA was isolated using the DiaRex® Whole Blood Genomic DNA Extraction Kit (Cat. No. BLD-5295, Diagen, Ankara). This process involved lysis, proteinase K digestion, ethanol precipitation, and column-based purification to obtain high-quality DNA. The measurement of nucleic acid concentrations was conducted utilising a Colibri Microvolume Spectrometer (Titertek-Berthold, Germany). The exons of the HFE *gene* were amplified via PCR and subjected to sequencing on the Illumina MiSeq platform using the Nextera XT DNA Library Prep Kit (Illumina, San Diego, CA). The variants were visualised using the Integrative Genomics Viewer (IGV) and classified according to the Franklin, VarSome, HGMD Public®, ClinVar, gnomAD, ExAC, KGP, and dbSNP databases.

The present study was conducted with the approval of the Non-Interventional Clinical Research Ethics Committee of Ordu University (decision number: 2025/227, date: 09.07.2025). This retrospective study was conducted using data obtained from medical records in accordance with institutional ethical standards.

RESULTS

A total of 100 patients (53 males and 47 females) who underwent *HFE* gene analysis were included in the study. The mean age of the patients was 47.2 ± 17.1 years (range: 2–81 years). Regarding biochemical parameters, serum ferritin levels were available for 97 patients, with a mean ferritin concentration of 447.9 ± 726.7 µg/L (min: 9.27; max: 6104). Serum iron levels were available in 93 patients, with an average of 105.2 ± 49.6 µg/dL (min: 27.8; max: 268.3). The measurement of hemoglobin values was conducted in all patients, yielding a mean of 14.3 ± 2.2 g/dL (min: 8.4; max: 19.1). TS data were available for 65 patients (65% of the cohort), and all TS-based statistical analyses were performed within this subset, with a mean TS of 41.7 ± 24.4% (min: 5.67%; max: 99%). The baseline demographic and biochemical characteristics of the study population are summarized in Table 1.

Genotype analysis revealed that 64 of the 100 patients (64.0%) had no HFE variants (wild-type). Among the 36 patients carrying at least one variant, diverse genotypic patterns were observed:

- 30 patients (30.0%) were identified as heterozygous for c.187C>G (p.His63Asp, H63D).
- patients (3.0%) were identified as homozygous for H63D.
- 1 patient (1.0%) was identified as homozygous for c.845G>A (p.Cys282Tyr, C282Y).
- 1 patient (1.0%) was identified as heterozygous for c.76+2 dup.
- 1 patient (1.0%) was found to carry both c.187C>G and c.76+2 dup as a compound heterozygote.

At the allele level, a total of 41 mutant alleles were identified: 37 c.187C>G (H63D), 2 c.845G>A (C282Y), and 2 c.76+2 dup.

When ferritin levels were assessed the upper reference limit of 400 µg/L, 34 patients (34.0%) had ferritin >400 µg/L. Using a TS threshold of 45%, elevated TS values were observed in 28 patients (28.0%).

Patients were then categorised into two main groups based on the presence of any *HFE* variant: those without variants (n = 64) and those carrying at least one *HFE* variant (n = 36). The comparative biochemical findings of variant carriers and non-carriers are shown in Table 2. The mean age of the participants did not differ significantly between the groups (47.3 ± 16.4 vs. 46.9 ± 18.7 years). Ferritin, serum iron, and hemoglobin levels were compared between variant carriers and non-carriers using the Mann–Whitney U test. This comparison revealed no statistically significant differences.

- ferritin ($p = 0.869$),
- serum iron ($p = 0.204$),
- hemoglobin ($p = 0.674$).

TS levels were found to be comparable between the groups, with mean values of 39.6 ± 23.6% recorded in the wild-type group and 44.7 ± 25.6% observed among variant carriers ($p = 0.453$).

Table 1. Baseline demographic and biochemical characteristics of the patients.

	n	Mean ± SD	Median (IQR)	Min–Max
Age	100	47.2 ± 17.1	49 (32–62)	2–81
Sex (male/female)	100	53 / 47	–	–
Hemoglobin (g/dL)	100	14.3 ± 2.2	14.4 (12.9–15.7)	8.4–19.1
Ferritin (µg/L)	97	447.9 ± 726.7	205 (95–405)	9.27–6104
Serum iron (µg/dL)	93	105.2 ± 49.6	97 (73–130)	27.8–268.3
Transferrin saturation (%)	65	41.7 ± 24.4	37 (22–54)	5.67–99

Min: Minimum, Max: Maximum, SD: Standard deviation, IQR: Interquartile range.

Table 2. Comparison of biochemical parameters between patients with and without HFE variants.

Parameter	No variant (n = 64)	≥1 variant (n = 36)	p–value
Ferritin (µg/L)	462 ± 841	420 ± 498	0.869
Serum iron (µg/dL)	110 ± 49	96 ± 48	0.204
Hemoglobin (g/dL)	14.3 ± 2.3	14.2 ± 1.9	0.674
Transferrin saturation (%)**	39.6 ± 23.6	44.7 ± 25.6	0.453
TS > 45% (n%)**	13 / 64 (20.3%)	15 / 36 (41.7%)	0.040*
Ferritin > 400 µg/L (n%)	23 / 64 (35.9%)	11 / 36 (30.6%)	0.745

*Statistically significant.
**Transferrin saturation analyses were performed only in patients with available TS data (n = 65); percentages are calculated within this subgroup.
TS: Transferrin saturation.

However, the frequency of elevated TS (TS > 45%) was notably higher in patients carrying *HFE* variants. Among the 65 patients with available TS data, 15 of 36 variant carriers (41.7%) and 13 of 64 non-carriers (20.3%) exhibited TS > 45%. Chi-square analysis demonstrated a statistically significant association between the presence of *HFE* variants and elevated TS ($\chi^2 = 4.21$, $p = 0.040$). On the basis of this finding, the likelihood of having TS > 45% was approximately 2.8 times higher in variant carriers compared non-carriers (odds ratio ≈ 2.80; 95% confidence interval: 1.14–6.89).

In contrast, no significant difference was observed in the proportion with ferritin > 400 µg/L between variant carriers (11/36; 30.6%) and non-carriers (23/64; 35.9%) ($p = 0.745$).

A detailed genotypic subgroup review showed that the highest ferritin value (6104 µg/L) occurred in a 78-year-old female patient who was heterozygous for H63D and had severe anemia (Hb: 8.4 g/dL, TS: 79.91%). The highest TS value, 99%, was observed in a male patient without any detectable *HFE* variant; ferritin was not available for this patient.

In summary, the most frequently observed *HFE* variant in this cohort was c.187C>G (H63D), predominantly in the heterozygous form. While no significant differences were observed between variant carriers and non-carriers in ferritin, serum iron, or hemoglobin levels, a significant association was detected between carriage of *HFE* variants and elevated TS. These findings suggest that *HFE* variants may influence iron loading primarily through TS, whereas ferritin levels in this cohort appear to be strongly influenced by additional modifying factors.

DISCUSSION

The present study provides a comprehensive evaluation of *HFE* gene variants and iron-related biochemical parameters in a clinically heterogeneous cohort from the Black Sea region of Türkiye. The present findings corroborate the predominance of the c.187C>G (H63D) variant and the rarity of the classical c.845G>A (C282Y) variant in accordance with population-based studies conducted in Mediterranean, Middle Eastern, and Central Asian populations (12,13). The low prevalence of C282Y in Türkiye has been repeatedly highlighted, suggesting that classical Northern European–type HH (HH type 1) is uncommon in this region, whereas H63D represents the major allele influencing iron parameters (14,15).

The absence of significant differences in ferritin and serum iron levels between variant carriers and non-carriers in this study aligns with the widely accepted notion that H63D has low penetrance and limited phenotypic expression in isolation (16). Numerous studies have demonstrated that even *H63D* homozygotes frequently do not develop clinically significant iron overload unless additional genetic or environmental modifiers are present, including metabolic syndrome, chronic liver disease, alcohol use or concurrent inflammatory disorders (17,18).

In contrast, the significantly higher TS values observed among variant carriers in our cohort provide biologically meaningful insight. TS has been shown to rise earlier than ferritin in the natural history of HH, reflecting increased ferroportin-mediated iron efflux and decreased hepcidin activity. These phenomena may be subtly influenced by H63D variants (19,20). It has been demonstrated through experimental studies that mutant *HFE* proteins exhibit altered interactions with TfR1 and transferrin receptor 2 (TfR2)

which suggests the potential to modify hepcidin regulation even in the absence of severe phenotypic expression (21,22). Therefore, the elevated TS among *HFE* variant carriers in our cohort may represent an early biochemical footprint of altered iron homeostasis.

Although mean TS values did not differ significantly between variant carriers and non-carriers, the categorical analysis using the clinically established threshold of TS > 45% revealed a statistically significant association with *HFE* variant carriage. This discrepancy may be explained by the clinical relevance of threshold-based interpretation in HH, where TS values exceeding 45% are widely accepted as early biochemical indicators of altered iron metabolism. Comparisons of continuous means may obscure clinically meaningful distributional shifts around diagnostic cut-off points, whereas dichotomized analyses better reflect real-world diagnostic decision-making. Nevertheless, the calculated odds ratio should be interpreted with caution given the retrospective design, the relatively limited sample size, and the availability of TS measurements in only a subset of patients.

However, ferritin levels did not differ significantly across genotypes. In view of ferritin's well-established role as an acute-phase reactant influenced by hepatic inflammation, obesity, and systemic metabolic conditions, this finding is not surprising (23,24). In this cohort, several individuals with markedly elevated ferritin levels exhibited concomitant anemia or other comorbidities. These findings support the hypothesis that ferritin lacks specificity in distinguishing genetic iron overload from secondary hyperferritinemia.

Beyond the primary biochemical findings, the genotype distribution in our cohort reflects a characteristic pattern observed in Southern European and Middle Eastern countries. Genetic epidemiology studies suggest that the C282Y allele originated in Northwestern Europe and underwent a founder expansion. However, H63D exhibits a much older and geographically widespread evolutionary history (12,13,15,25). The predominance of H63D in our population may therefore be attributed to ancient migration patterns and long-term allelic equilibrium, rather than to disease-driven selection.

The single case of C282Y homozygosity in the present dataset indicates the rarity of classical *HFE*-associated hemochromatosis genotypes in Türkiye. While C282Y homozygotes demonstrate the highest clinical penetrance and strongest association with progressive iron overload, even among C282Y individuals the penetrance varies widely across ethnic groups and is strongly influenced by sex, dietary iron intake, alcohol consumption, viral hepatitis, metabolic syndrome and hormonal status (12,13,26). Therefore, the low prevalence of C282Y in our cohort is consistent with the limited number of clinically confirmed HH type 1 cases reported by hepatology centers across Türkiye.

In addition to the classical variants, one patient in our cohort was identified as carrying the c.76+2dup splice-site variant. This alteration affects the highly conserved +2 position of the donor splice site, a region that is critical for proper mRNA splicing. Variants occurring at canonical splice-site positions are generally considered likely to disrupt normal transcript processing and may result in exon skipping or aberrant protein products. However, functional validation data for c.76+2dup are currently limited, and its precise impact on HFE protein expression and hepcidin regulation remains unclear. Therefore, although the variant may have potential pathogenic relevance based on its predicted molecular consequence, cautious interpretation is warranted in the absence of experimental confirmation or segregation analysis.

A notable observation in the present study was the discordance between ferritin and TS across genotypes. This finding is biologically plausible: while TS reflects circulating iron that is immediately available for cellular uptake, ferritin reflects stored iron and is heavily influenced by inflammatory pathways. Systemic inflammation has been observed to stimulate ferritin synthesis via IL-6-mediated pathways and to suppress ferroportin, resulting in the paradoxical coexistence of high ferritin and low TS—commonly observed in anemia of chronic disease and non-alcoholic fatty liver disease related hyperferritinemia (23,24). In contrast, early HH is classically characterised by disproportionately elevated TS with normal or mildly elevated ferritin (20). The present findings support this model: variant carriers exhibited higher TS despite having ferritin levels comparable to those of wild-type individuals. Such discrepancies emphasise the clinical importance of TS as an early diagnostic parameter, particularly in regions where H63D predominates and C282Y-based HH is rare.

The mild phenotype of H63D carriers observed in the present dataset is consistent with molecular functional studies demonstrating that the H63D variant results in minimal disruption of the *HFE–TfR1–TfR2–hepcidin* axis compared with C282Y, H63D mutant proteins have been observed to retain a partial ability to form functional complexes, resulting in a subtle reduction rather than complete absence of hepcidin signaling (21,22). Consequently, many H63D carriers may present with mild biochemical alterations without clinically significant iron deposition. However, environmental interactions, such as obesity, alcohol consumption, or metabolic syndrome have been demonstrated to amplify these effects and unmask subclinical dysregulation (17,18).

This study has several limitations. First, its referral-based design meant that all participants were included based on a preliminary clinical suspicion of HH rather than on confirmed diagnostic criteria, resulting in a heterogeneous diagnostic spectrum. This may limit the generalizability of genotype–phenotype associations and partly explain the relatively high proportion of wild-type individuals and the modest biochemical differences observed. Second, TS measurements were available for only 65% of the cohort and TS-based analyses were therefore restricted to this subgroup, potentially reducing statistical power and influencing the interpretation of the results. Finally, the relatively small sample size (n = 100) and the restriction of genetic analysis to the *HFE* gene—without sequencing other hemochromatosis-related genes—limit broader conclusions, and non-*HFE* genetic causes of iron overload cannot be excluded.

CONCLUSION

This study confirms that the *HFE* variant profile in Türkiye is characterised by a high prevalence of the H63D variant and a very low frequency of C282Y. While H63D does not significantly influence ferritin levels, it is associated with a measurable increase in TS, suggesting a mild effect on iron transport rather than substantial iron accumulation. These findings highlight the diagnostic value of TS in populations where classical C282Y-related hemochromatosis is rare.

Further multicenter and prospective studies using comprehensive clinical and imaging-based assessments are required to clarify genotype–phenotype relationships and to optimise diagnostic strategies for iron metabolism disorders in the Turkish population.

Ethics

Ethics Committee Approval: The present study was conducted with the approval of the Non-Interventional Clinical Research Ethics Committee of Ordu University (decision number: 2025/227, date: 09.07.2025).

Informed Consent: This retrospective study was conducted using data obtained from medical records in accordance with institutional ethical standards.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ç.D., A.O.G., M.Ö., Concept: Ç.D., A.O.G., M.Ö., Design: Ç.D., A.O.G., M.Ö., Data Collection or Processing: Ç.D., A.O.G., M.Ö., Analysis or Interpretation: Ç.D., A.O.G., M.Ö., Literature Search: Ç.D., A.O.G., M.Ö., Writing: Ç.D., A.O.G., M.Ö.

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

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

Acknowledgments

We extend our sincere thanks to Dr. Elif Doğan for her contributions to language editing and final revisions in this study.

REFERENCES

1. Pietrangelo A. Hereditary hemochromatosis: pathogenesis, diagnosis, and treatment. Gastroenterology. 2010; 139: 393-408.
2. Allen KJ, Gurrin LC, Constantine CC, Osborne NJ, Delatycki MB, Nicoll AJ, et al. Iron-overload-related disease in HFE hereditary hemochromatosis. N Engl J Med. 2008; 358: 221-30.
3. Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A, et al. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. Nat Genet. 1996; 13: 399-408.
4. Nemeth E, Ganz T. Regulation of iron metabolism by hepcidin. Annu Rev Nutr. 2006; 26: 323-42.
5. Fleming RE, Sly WS. Mechanisms of iron accumulation in hereditary hemochromatosis. Annu Rev Physiol. 2002; 64: 663-80.
6. Kowdley KV, Brown KE, Ahn J, Sundaram V. ACG Clinical Guideline: Hereditary Hemochromatosis. Am J Gastroenterol. 2019; 114: 1202-18.
7. Waheed A, Parkkila S, Zhou XY, Tomatsu S, Tsuchihashi Z, Feder JN, et al. Hereditary hemochromatosis: effects of C282Y and H63D mutations on association with beta2-microglobulin, intracellular processing, and cell surface expression of the HFE protein in COS-7 cells. Proc Natl Acad Sci U S A. 1997; 94: 12384-9.
8. Hasan SMM, Farrell J, Borgaonkar M. C282Y/H63D Compound Heterozygosity Is a Low Penetrance Genotype for Iron Overload-related Disease. J Can Assoc Gastroenterol. 2022; 5: 240-7.
9. Anderson GJ, Bardou-Jacquet E. Revisiting hemochromatosis: genetic vs. phenotypic manifestations. Ann Transl Med. 2021; 9: 731.
10. Bardou-Jacquet E, Brissot P. Diagnostic evaluation of hereditary hemochromatosis (HFE and non-HFE). Hematol Oncol Clin North Am. 2014; 28: 625-35.
11. Powell LW, Seckington RC, Deugnier Y. Haemochromatosis. Lancet. 2016; 388: 706-16.
12. Zarifian Yeganeh R, Akbari Kelishomi M, Ahmadpour Jenaghard A, Salmani B, Vahidi Z, Makvand M, et al. HFE and Non-HFE hereditary hemochromatosis based on screening of 854 individuals: 12 years of an Iranian experience. Genet Test Mol Biomarkers. 2024; 28: 289-96.
13. Wallace DF, Subramaniam VN. The global prevalence of HFE and non-HFE hemochromatosis estimated from analysis of next-generation sequencing data. Genet Med. 2016; 18: 618-26.
14. P S. Hereditary Hemochromatosis Gene Mutations in the South Eastern Region of Turkey. Artuklu Int J Health Sci. 2021.
15. Simsek H, Sumer H, Yilmaz E, Balaban YH, Ozcebe O, Hascelik G, et al. Frequency of HFE mutations among Turkish blood donors according to transferrin saturation: genotype screening for hereditary hemochromatosis among voluntary blood donors in Turkey. J Clin Gastroenterol. 2004; 38: 671-5.
16. Brissot P, Pietrangelo A, Adams PC, de Graaff B, McLaren CE, Loréal O. Haemochromatosis. Nat Rev Dis Primers. 2018; 4: 18016.
17. Sun Z, Pan X, Tian A, Surakka I, Wang T, Jiao X, et al. Genetic variants in HFE are associated with non-alcoholic fatty liver disease in lean individuals. JHEP Rep. 2023; 5: 100744.
18. Kowdley KV, Belt P, Wilson LA, Yeh MM, Neuschwander-Tetri BA, Chalasani N, et al. Serum ferritin is an independent predictor of histologic severity and advanced fibrosis in patients with nonalcoholic fatty liver disease. Hepatology. 2012; 55: 77-85.
19. Baddam S, Chen RJ. Iron Overload and Toxicity. [Updated 2025 Jul 7]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2026 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK526131/
20. Bacon BR, Adams PC, Kowdley KV, Powell LW, Tavill AS; American Association for the Study of Liver Diseases. Diagnosis and management of hemochromatosis: 2011 practice guideline by the American Association for the Study of Liver Diseases. Hepatology. 2011; 54: 328-43.
21. Schmidt PJ, Toran PT, Giannetti AM, Bjorkman PJ, Andrews NC. The transferrin receptor modulates Hfe-dependent regulation of hepcidin expression. Cell Metab. 2008; 7: 205-14.
22. Babitt JL, Lin HY. Molecular mechanisms of hepcidin regulation: implications for the anemia of CKD. Am J Kidney Dis. 2010; 55: 726-41.
23. Kell DB, Pretorius E. Serum ferritin is an important inflammatory disease marker, as it is mainly a leakage product from damaged cells. Metallomics. 2014; 6: 748-73.
24. Ganz T, Nemeth E. Hepcidin and iron homeostasis. Biochim Biophys Acta. 2012; 1823: 1434-43.
25. Barton JC, Parker CJ. HFE-Related Hemochromatosis. 2000 Apr 3 [updated 2024 Apr 11]. In: Adam MP, Bick S, Mirzaa GM, Pagon RA, Wallace SE, Amemiya A, editors. GeneReviews®[Internet]. Seattle (WA): University of Washington, Seattle; 1993–2026.
26. Rossi E, Jeffrey GP. Clinical penetrance of C282Y homozygous HFE haemochromatosis. Clin Biochem Rev. 2004; 25: 183-90.

Gazi Med J 2026;37(2):244-247

DOI: http://dx.doi.org/10.12996/gmj.2026.4585

Thoracic Disc Herniation Presenting As Visceral Pain: A Case Report

Visseral Ağrı ile Prezente Olan Torakal Disk Herniasyonu: Olgu Sunumu

Zeynep Balaban, Cansu Sönmez, Nurullah Gök, Aydemir Kale

Department of Neurosurgery, Gazi University, Faculty of Medicine, Ankara, Türkiye

ABSTRACT

Thoracic disc herniation (TDH) is a rare clinical entity, accounting for less than 1% of all intervertebral disc herniations. Although it most commonly presents with axial back pain, myelopathy, or radiculopathy, TDH can occasionally manifest with atypical and misleading symptoms such as visceral or abdominal pain, significantly complicating diagnosis and potentially delaying appropriate neurosurgical intervention. We report the case of a 48-year-old woman who presented with chronic abdominal and back pain, accompanied by intermittent numbness in the lower extremities. Despite extensive prior investigations, no definitive diagnosis was established. Thoracic spine magnetic resonance imaging demonstrated a right paracentral disc herniation at the T9–10 level, causing compression of the spinal cord and adjacent nerve roots. The patient underwent a right transfacet discectomy, which resulted in complete resolution of symptoms without postoperative complications. This case highlights the diagnostic challenge of TDH presenting with atypical abdominal symptoms. Spinal imaging should be considered in patients with unexplained visceral pain, especially when neurological findings are present. Early detection can prevent unnecessary interventions and allow timely surgical treatment.

Keywords: Thoracic disc herniation, abdominal pain, atypical presentation, transfacet discectomy, radiculopathy, spinal cord compression

ÖZ

Torakal disk herniasyonu (TDH), tüm intervertebral disk herniasyonlarının %1'inden azını oluşturan nadir bir klinik durumdur. En sık aksiyel sırt ağrısı, miyelopati veya radikülopati ile prezente olmakla birlikte, bazı durumlarda visseral veya abdominal ağrı gibi atipik ve yanıltıcı semptomlarla ortaya çıkabilir. Bu durum tanıyı güçleştirmekte ve uygun nöroşirürjik müdahalenin gecikmesine neden olabilmektedir.

Kronik abdominal ve sırt ağrısı ile başvuran, alt ekstremitelerde aralıklı uyuşma şikayeti bulunan 48 yaşında kadın hasta sunulmuştur. Daha önce yapılan kapsamlı tetkiklere rağmen kesin bir tanı konulamamıştır. Torakal omurga manyetik rezonans görüntülemede T9–10 seviyesinde sağ parasantral disk herniasyonu saptanmış olup, spinal kord ve komşu sinir köklerine bası yaptığı izlenmiştir. Hastaya sağ transfacet diskektomi uygulanmış ve postoperatif dönemde herhangi bir komplikasyon gelişmeden semptomlarda tam düzelme sağlanmıştır.

Bu olgu, TDH'nin atipik abdominal semptomlarla prezente olabileceğini ve tanısal zorluklar oluşturabileceğini göstermektedir. Açıklanamayan visseral ağrı varlığında, özellikle nörolojik bulgular eşlik ediyorsa, spinal görüntüleme mutlaka düşünülmelidir. Erken tanı, gereksiz girişimlerin önlenmesini ve zamanında cerrahi tedaviyi mümkün kılar.

Anahtar Sözcükler: Torakal disk herniasyonu, abdominal ağrı, atipik prezentasyon, transfacet diskektomi, radikülopati, spinal kord basısı

Cite this article as: Balaban Z, Sönmez C, Gök N, Kale A. Thoracic disc herniation presenting as visceral pain: a case report. Gazi Med J. 2026;37(2):244-247

Address for Correspondence/Yazışma Adresi: Zeynep Balaban, Department of Neurosurgery, Gazi University, Faculty of Medicine, Ankara, Türkiye
E-mail / E-posta: zeynep.balaban@yahoo.com
ORCID ID: orcid.org/0009-0002-1995-3218

Received/Geliş Tarihi: 07.11.2025
Accepted/Kabul Tarihi: 04.01.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

INTRODUCTION

Thoracic disc herniation (TDH) is a rare spinal pathology, accounting for less than 1% of all intervertebral disc herniations (1). This rarity is largely due to the relative immobility of the thoracic spine (1,2). TDH most frequently affects individuals in middle-to-late adulthood and shows no clear gender predominance. The typical clinical presentation includes axial back pain, sensory deficits, signs of myelopathy, and varying degrees of lower extremity weakness (3-5). However, atypical and misleading presentations are increasingly reported in the literature.

Among these presentations, visceral pain syndromes–such as abdominal, testicular, or chest pain–pose a significant diagnostic challenge and frequently result in delayed diagnosis (6). The underlying pathophysiological mechanism is thought to involve irritation of the thoracic nerve roots, producing referred pain patterns that mimic intra-abdominal or cardiopulmonary disorders (7). For this reason, TDH should be considered in the differential diagnosis of unexplained visceral pain, especially when conventional investigations fail to identify a clear etiology.

We present a rare case of TDH in a patient who initially had chronic abdominal pain accompanied by dorsal back pain and sensory disturbances, leading to a delayed diagnosis.

CASE REPORT

A 48-year-old woman was admitted to our clinic with a one-year history of persistent back pain, abdominal discomfort, and intermittent numbness in her lower extremities. She had previously attended several healthcare centers and had been treated for various preliminary diagnoses without significant clinical improvement. Initial suspicion of *Helicobacter pylori* infection led to eradication therapy, which resulted in partial relief of her abdominal symptoms; however, her back pain remained unchanged.

As part of the differential diagnostic work-up, magnetic resonance imaging (MRI) of the thoracic spine was performed. Imaging revealed a right paracentral disc herniation at the T9–10 level, and the patient was subsequently referred to our neurosurgery department (Figure 1). Following evaluation by a multidisciplinary neurosurgical council, surgical intervention was considered the most appropriate treatment strategy.

Neurological examination revealed globally hypoactive deep tendon reflexes, except for a relatively preserved left patellar reflex. No focal motor deficits or pathological reflexes were observed. Systemic examination revealed no additional abnormalities. The patient had no history of trauma, prior surgery, or chronic medication use. Conservative treatment, including physical therapy and rehabilitation, had previously been attempted without success.

Thoracic MRI demonstrated preserved vertebral alignment and localised degenerative signal changes at multiple thoracic levels. At the T9–10 level, a right paracentral subarticular disc extrusion with inferior migration, accompanied by a diffuse annular protrusion, was identified.

Under general anesthesia, the patient was placed in a modified prone position. A midline incision and a unilateral subperiosteal dissection were performed. The right facet joint at T9–10 was removed using Kerrison rongeurs and a high-speed drill to gain access to the lateral spinal canal. A right transfacet approach was used to reach the disc space. Partial inferior laminectomy at T9, and resection of the ligamentum flavum, allowed identification and decompression of the right T10 nerve root (Figure 2). A migrated disc fragment compressing the thecal sac was removed, resulting in adequate neural decompression.

The patient experienced no postoperative neurological deficits, reported complete symptom resolution, and was discharged without complications on postoperative day two.

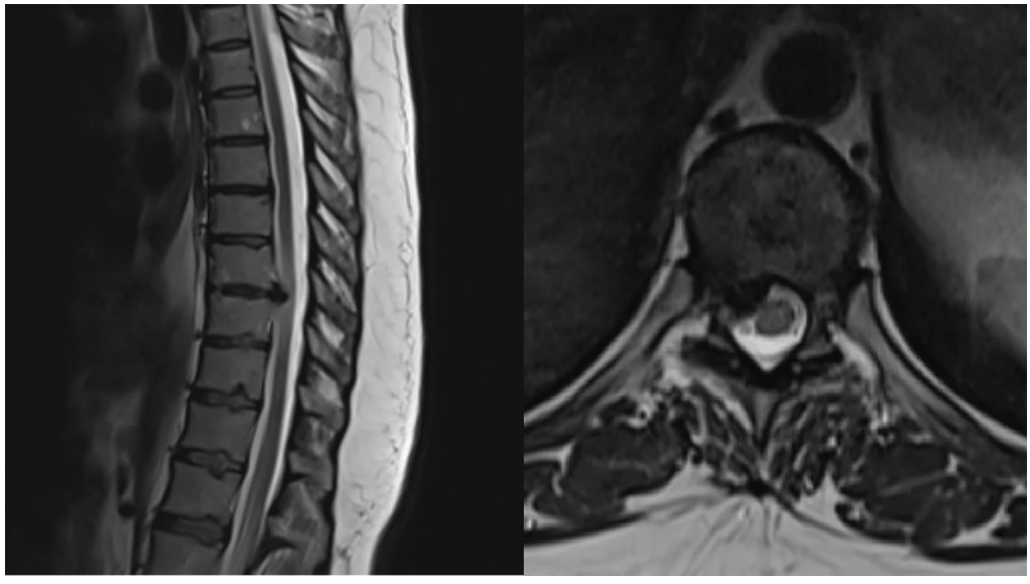

Figure 1. Sagittal and axial T2-weighted thoracic MRI demonstrating a right paracentral T9–10 disc herniation with inferior migration, resulting in anterior and anterolateral compression of the spinal cord and indentation of the descending right T10 nerve root.

MRI: Magnetic resonance imaging.

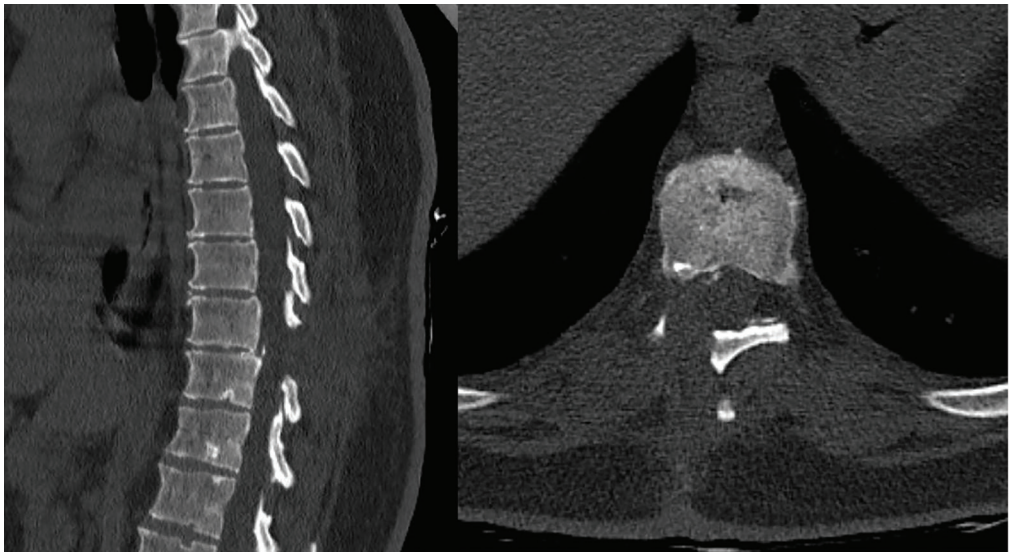

Figure 2. Postoperative sagittal and axial computed tomography images showing decompression at the T9–10 level after right transfacet discectomy.

DISCUSSION

Thoracic disc herniation is a rare spinal pathology, accounting for less than 1% of all intervertebral disc herniations (1,2). Although pain is the most common presenting symptom, its character and distribution can vary widely, ranging from localized axial back pain to radicular pain radiating to the chest or abdomen (1). While radiculopathy is frequently observed in TDH, its occurrence as an isolated symptom is extremely rare (1,4). According to Nishimura et al., only a limited number of cases have been reported in which TDH presented exclusively with radicular symptoms, such as abdominal pain or abdominal muscle weakness (8).

Atypical presentations of TDH – including abdominal, groin, testicular, or cardiac-like chest pain – can mimic gastrointestinal, genitourinary, cardiopulmonary, or even psychiatric conditions (9). This often leads to extensive and sometimes invasive diagnostic investigations before a spinal etiology is considered. Some patients have undergone unnecessary procedures, including diagnostic laparoscopy, due to misattribution of symptoms. It has been suggested that up to 10% of patients presenting with chronic idiopathic abdominal pain may have an undiagnosed thoracic disc herniation (10). Pérez Lara et al. reported that approximately 66% of patients with chronic unexplained abdominal pain were found to have a herniated disc on MRI (11).

In this case, the patient was initially misdiagnosed and treated for *Helicobacter pylori* infection, resulting in only partial symptom relief. This emphasises the importance of including thoracic spinal pathology in the differential diagnosis of chronic abdominal pain, particularly when routine gastrointestinal evaluations are inconclusive. Differential diagnoses should include gastrointestinal disorders, cardiopulmonary diseases, and genitourinary pathologies.

Radiologically, thoracic disc herniations can be classified into five types according to size and location: Type 0 (small, <40% canal compromise, without neural compression); Type 1 (small, paracentral), Type 2 (small, central), Type 3 (large, >40%, paracentral), and Type 4 (large, central) herniations (5). This classification assists in determining the most appropriate surgical approach. Posterolateral approaches–including transfacet, transpedicular, and costotransversectomy techniques–are generally preferred for lateralised or paracentral soft herniations, whereas anterior or transthoracic approaches are more suitable for central or calcified lesions (12-14).

In this case, a right transfacet approach was preferred due to the paracentral and lateralized nature of the soft disc herniation without significant calcification. Alternative posterior approaches, such as transpedicular or costotransversectomy techniques, were not selected because they would have required more extensive bone removal and would have carried a higher risk of postoperative spinal instability. Anterior or transthoracic approaches, although effective for large central or calcified thoracic disc herniations, were considered unnecessarily invasive for this lesion and were associated with increased cardiopulmonary morbidity. Therefore, the transfacet approach provided sufficient exposure and safe decompression, with minimal tissue disruption.

Surgical intervention should be reserved for patients with persistent pain, progressive neurological deficits, or myelopathy refractory to conservative treatment (15). Conservative management remains the first-line treatment for most patients with uncomplicated thoracic disc herniation.

CONCLUSION

Thoracic disc herniation is a rare but important cause of atypical visceral pain that can easily be overlooked. This case highlights that thoracic spinal pathology should be considered in patients with persistent, unexplained abdominal pain, particularly when subtle neurological findings are present. Early recognition and appropriate spinal imaging are crucial to avoid unnecessary diagnostic procedures and delays in treatment. Increased awareness of such atypical presentations allows for timely surgical intervention, leading to complete resolution of symptom and the prevention of potentially irreversible neurological deficits.

Footnotes

Authorship Contributions

Design: Z.B., Data Collection or Processing: C.S., N.G., Analysis or Interpretation: Z.B., A.K., Literature Search: Z.B., Writing: Z.B.

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

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

REFERENCES

1. Court C, Mansour E, Bouthors C. Thoracic disc herniation: Surgical treatment. Orthop Traumatol Surg Res. 2018; 104: S31-40.
2. Browne JE, Branch TP. Surgical alternatives for treatment of articular cartilage lesions. J Am Acad Orthop Surg. 2000; 8: 180-9.
3. Stillerman CB, Chen TC, Couldwell WT, Zhang W, Weiss MH. Experience in the surgical management of 82 symptomatic herniated thoracic discs and review of the literature. J Neurosurg. 1998; 88: 623-33.
4. Brown CW, Deffer PA Jr, Akmakjian J, Donaldson DH, Brugman JL. The natural history of thoracic disc herniation. Spine (Phila Pa 1976). 1992; 17: S97-102.
5. Hott JS, Feiz-Erfan I, Kenny K, Dickman CA. Surgical management of giant herniated thoracic discs: Analysis of 20 cases. J Neurosurg Spine. 2005; 3: 191-7.
6. Shirzadi A, Drazin D, Jeswani S, Lovely L, Liu J. Atypical presentation of thoracic disc herniation: Case series and review of the literature. Case Rep Orthop. 2013; 2013: 621476.
7. Xiong Y, Lachmann E, Marini S, Nagler W. Thoracic disk herniation presenting as abdominal and pelvic pain: A case report. Arch Phys Med Rehabil. 2001; 82: 1142-4.
8. Ishii M, Nishimura Y, Hara M, Eguchi K, Nagashima Y, Awaya T, et al. Thoracic disc herniation manifesting as abdominal pain alone associated with thoracic radiculopathy. NMC Case Rep J. 2020; 7: 161-5.
9. Brown NJ, Kuo C, Pennington Z, Zhang A, Choi AE, Chan AK, et al. Diagnosis and management of thoracolumbar spinal disorders presenting as cardiac, gastrointestinal, and other false pain syndromes. Clin Spine Surg. 2025; 38: 114-22.
10. Wilke A, Wolf U, Lageard P, Griss P. Thoracic disc herniation: A diagnostic challenge. Man Ther. 2000; 5: 181-4.
11. Lara FJ, Quesada JQ, Ramiro JA, Toledo RB, Del Rey Moreno A, Muñoz HO. Chronic abdominal syndrome due to nervous compression. Study of 100 cases and proposed diagnostic-therapeutic algorithm. J Gastrointest Surg. 2015; 19: 1059-71.
12. Quraishi NA, Khurana A, Tsegaye MM, Boszczyk BM, Mehdian SM. Calcified giant thoracic disc herniations: considerations and treatment strategies. Eur Spine J. 2014; 23 Suppl 1: S76-83.
13. Quraishi NA, Khurana A, Tsegaye MM, Boszczyk BM, Mehdian SM. Calcified giant thoracic disc herniations: Considerations and treatment strategies. Eur Spine J. 2014; 23 Suppl 1: S76-83.
14. Strom RG, Mathur V, Givans H, Kondziolka DS, Perin NI. Technical modifications and decision-making to reduce morbidity in thoracic disc surgery: An institutional experience and treatment algorithm. Clin Neurol Neurosurg. 2015; 133: 75-82.
15. Russo A, Balamurali G, Nowicki R, Boszczyk BM. Anterior thoracic foraminotomy through mini-thoracotomy for the treatment of giant thoracic disc herniations. Eur Spine J. 2012; 21 Suppl 2: S212-20.
16. Vanichkachorn JS, Vaccaro AR. Thoracic disk disease: Diagnosis and treatment. J Am Acad Orthop Surg. 2000; 8: 159-69.

Gazi Med J 2026;37(2):248-251

DOI: http://dx.doi.org/10.12996/gmj.2025.4573

Gastric Schwannoma Masquerading as a Gastrointestinal Stromal Tumor: A Case Report and Diagnostic Challenges

Gastrointestinal Stromal Tümörü Taklit Eden Gastrik Schwannoma: Bir Olgu Sunumu ve Tanısal Zorluklar

Mehmet Akif Türkoğlu[1], Lorena Remenar[2], Seda Alabaş[3], Yiğit Kağan Zeren[1], Emine Türkmen Şamdancı[3]

[1]Department of General Surgery, Gazi University, Faculty of Medicine, Ankara, Türkiye
[2]General Practitioner, Health Center Zagreb – West, Zagreb, Croatia
[3]Department of Pathology, Gazi University, Faculty of Medicine, Ankara, Türkiye

ABSTRACT

Gastric schwannomas (GSs) are rare, benign mesenchymal tumors originating from Schwann cells of the nerve sheath, accounting for approximately 0.2% of all gastric tumors. These tumors pose diagnostic challenges due to their resemblance to gastrointestinal stromal tumors (GISTs) and other submucosal neoplasms. We present the case of a 47-year-old male with a history of GI bleeding who was initially diagnosed with suspected GIST based on an enhanced abdominal computed tomography scan. The scan revealed a 7.4 × 5.6-cm exophytic mass with homogeneous enhancement in the gastric antrum. Laparoscopic distal gastrectomy with Roux-en-Y gastrojejunostomy was performed, revealing a tumor with purple-red discoloration, serosal invasion, omental involvement, and suspected metastatic infra-pyloric nodules. Histopathological examination confirmed a GS, characterized by spindle-shaped cells with Antoni A and B patterns, Verocay bodies, and strong S-100 protein immunoreactivity, while negative for CD117, DOG-1, CD34, SMA, and desmin. The patient recovered uneventfully and was discharged on postoperative day 6. This case underscores the diagnostic complexity of GSs, emphasizing the critical role of histopathological and immunohistochemical analysis in distinguishing them from other mesenchymal tumors. Minimally invasive surgical resection offers both diagnostic confirmation and effective treatment, with favorable oncologic and functional outcomes.

Keywords: Gastric schwannoma, schwann cells, s-100 protein, gastric submucosal tumors, laparoscopic distal gastrectomy, immunohistochemical analysis

ÖZ

Gastrik schwannomalar (GSs), nöral kılıfın Schwann hücrelerinden köken alan, tüm gastrik tümörlerin yaklaşık %0,2'sini oluşturan nadir, benign mezenkimal tümörlerdir. Bu tümörler, gastrointestinal stromal tümörlere (GISTs) ve diğer submukozal neoplazilere benzerlik göstermeleri nedeniyle tanısal zorluklar oluşturmaktadır. GI kanama öyküsü olan 47 yaşında erkek hastada, kontrastlı abdominal bilgisayarlı tomografi taramasına dayanarak başlangıçta GIST şüphesiyle tanı konulan bir olguyu sunuyoruz. Taramada gastrik antrumda homojen kontrastlanma gösteren 7,4 × 5,6 cm'lik ekzofitik bir kitle saptandı. Laparoskopik distal gastrektomi ve Roux-en-Y gastrojejunostomi uygulandı; mor-kırmızı renkli, serozal invazyon, omental tutulum ve metastatik infra-pilorik nodül şüphesi olan bir tümör görüldü. Histopatolojik inceleme, Antoni A ve B paternleri olan iğsi hücreler, Verocay cisimcikleri ve güçlü S-100 protein immünreaktivitesi ile karakterize, CD117, DOG-1, CD34, SMA ve desmin için negatif olan gastrik schwannomayı doğruladı. Hasta sorunsuz iyileşti ve postoperatif 6. günde taburcu edildi. Bu olgu, GS tanısal karmaşıklığını vurgulamakta ve bunları diğer mezenkimal tümörlerden ayırt etmede histopatolojik ve immünohistokimyasal analizin kritik rolünü öne çıkarmaktadır. Minimal invaziv cerrahi rezeksiyon hem tanısal doğrulama hem de etkili tedavi sağlar ve olumlu onkolojik ve fonksiyonel sonuçlar sunar.

Anahtar Sözcükler: Gastrik schwannoma, schwann hücreleri, S-100 proteini, gastrik submukozal tümörler, laparoskopik distal gastrektomi, immünohistokimyasal analiz

Cite this article as: Türkoğlu MA, Remenar L, Zeren YK, Alabaş S, Türkmen Şamdancı E. Gastric schwannoma masquerading as a gastrointestinal stromal tumor: a case report and diagnostic challenges. Gazi Med J. 2026;37(2):248-251

Address for Correspondence/Yazışma Adresi: Mehmet Akif Türkoğlu, Department of General Surgery, Gazi University Faculty of Medicine, Ankara
E-mail / E-posta: makturko@gmail.com
ORCID ID: orcid.org/0000-0002-7511-8201

Received/Geliş Tarihi: 25.10.2025
Accepted/Kabul Tarihi: 18.12.2025
Epub: 10.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

INTRODUCTION

Schwannomas (neurilemmomas) are rare tumors that originate from Schwann cells of peripheral nerves, which are surrounded by a fibrous sheath. Schwannomas are uncommon in the gastrointestinal (GI) tract, accounting for only 2–7% of mesenchymal GI tumors, 0.2% of all gastric tumors, and 4% of benign gastric neoplasms (1). Patients typically present with non-specific symptoms such as upper abdominal pain, dyspepsia, bleeding, and abdominal masses. These symptoms generally correlate with the tumor's size and location. Gastric schwannomas (GSs) characteristically present as submucosal or intramuscular lesions, posing considerable difficulty in distinguishing them from GISTs, leiomyomas, and leiomyosarcomas (1,2).

We report a case of a 47-year-old male who presented with GI bleeding. Preoperative diagnosis suggested a GI stromal tumor (GIST); however, postoperative histopathological examination definitively identified the lesion as a GS.

CASE REPORT

A 47-year-old male patient presented to an external medical facility with hematemesis. Upper GI endoscopy was performed but yielded inadequate visualization due to the presence of a blood coagulum. Subsequently, an enhanced abdominal computed tomography (CT) scan incidentally identified a gastric tumor with exophytic extension located in the gastric antrum, measuring approximately 7.4 × 5.6 cm in diameter. The lesion was preliminarily diagnosed as a GIST (Figure 1). The scan also identified perigastric lymphadenopathy, with the largest lymph node measuring 3 cm in diameter. Following spontaneous resolution of the tumor-related bleeding and transfusion of two units of packed red blood cells, the patient was referred to our institution for further evaluation and management. On admission, laboratory findings revealed a hemoglobin level of 9.1 g/dL and an elevated C-reactive protein level of 124 mg/L, with all other laboratory parameters within normal limits. Tumor markers, including carcinoembryonic antigen at 2.9 ng/mL and carbohydrate antigen 19-9 (CA19-9) at 15 U/mL, were within normal ranges. Physical examination yielded unremarkable findings. The patient's complex medical history included prior coronary artery bypass grafting necessitating long-term anticoagulation therapy, type 2 diabetes mellitus, hypertension, and morbid obesity (body mass index 43.2 kg/m²).

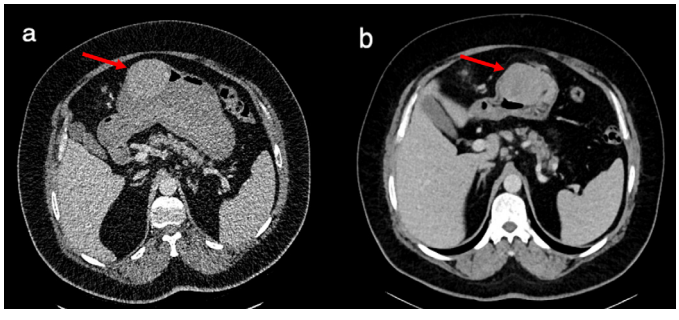

Figure 1. (a and b). An enhanced abdominal computed tomography scan showing a mass with an exophytic growth pattern and homogeneous enhancement located at the gastric antrum, measuring approximately 7.4 × 5.6 cm. (red arrows).

A laparoscopic distal gastrectomy with Roux-en-Y gastrojejunostomy was performed. The surgical technique has been previously described in detail (3). Upon gross examination, an endophytic-exophytic mass measuring 7 × 5 cm was identified, characterized by irregular margins and an ulcerated, nodular surface. The solid lesion displayed purple-red discoloration, suggestive of areas of necrosis and hemorrhage. Serosal invasion was evident, with apparent involvement of the omentum and the presence of suspected metastatic nodules (Figure 2).

Postoperatively, the patient was initiated on a clear-liquid diet and was gradually advanced to a regular diet by postoperative day 3. The surgical drain was removed on postoperative day 6, and the patient was discharged without complications.

Pathology

Macroscopically, the tumor measured 6.5 × 5 × 4.5 cm and exhibited an ulceration on the gastric mucosal surface. The cut surface of the tumor was off-white and fibrous. Tissue samples subjected to routine follow-up were stained with hematoxylin and eosin and examined by light microscopy. Histopathologically, the lesion is typically well-circumscribed and encapsulated. The gastric mucosa was ulcerated. Two characteristic histological patterns, known as Antoni A and Antoni B areas, are observed. Antoni A areas are hypercellular, consisting of spindle-shaped Schwann cells arranged in interlacing fascicles with prominent nuclear palisading and forming Verocay bodies. In contrast, Antoni B areas are hypocellular, exhibiting a loose, myxoid stroma with microcystic degeneration and less organized cellular architecture (Figure 3). Mitosis and necrosis were not detected. Immunohistochemically, tumor cells show strong, diffuse positivity for S-100 protein (Figure 4). The tumor cells were negative for antibodies to CD117, DOG-1, CD34, smooth muscle actin (SMA), and Desmin.

Figure 2. Intraoperative image showing a tumor with purple-red discoloration, serosal involvement, and adjacent infra-pyloric nodules. These nodules were initially suspected to be metastatic intraoperatively but were later confirmed by histopathology to be reactive lymph nodes.

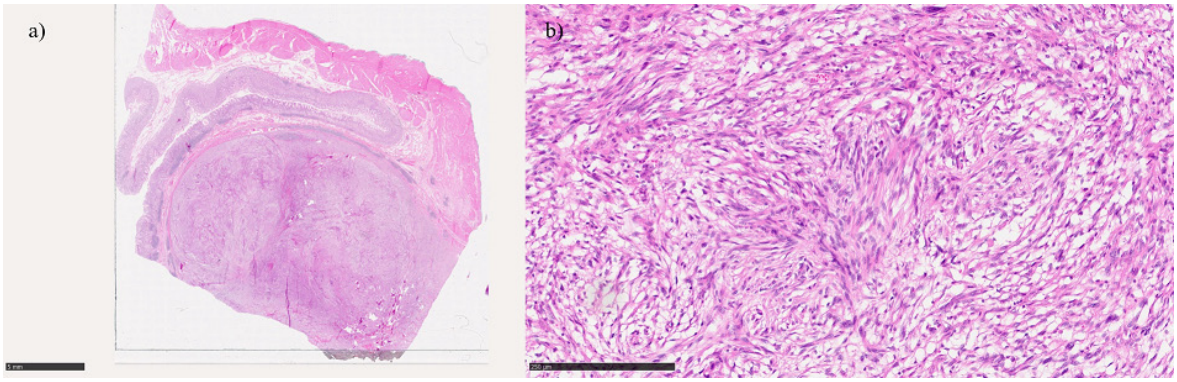

Figure 3. (a) The tumor is well-circumscribed and located under the gastric mucosa. Hematoxylin and eosin (H&E) x1. (b) The tumor cells are narrow, elongated, and wavy, with ill-defined cytoplasm and dense chromatin. H&E x100.

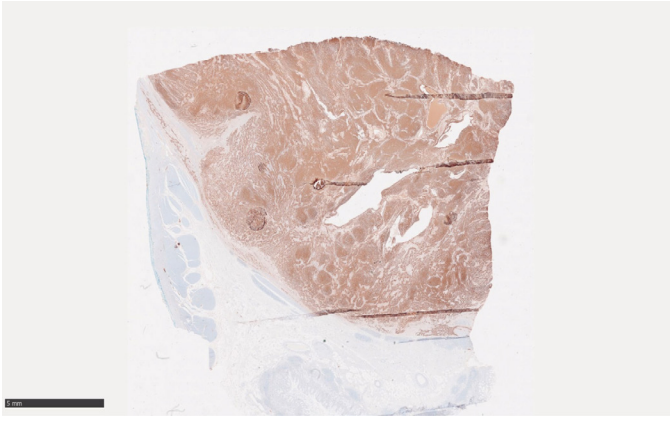

Figure 4. Strong S100 protein immunoreactivity in the tumor cells. Hematoxylin and eosin x1.

DISCUSSION

Gastric submucosal tumors (SMTs) are classified into three categories: myogenic tumors (comprising leiomyomas and leiomyosarcomas), neurogenic tumors (including schwannomas, granular cell tumors, and neurofibromas), and GISTs, each with unique histopathological and immunohistochemical characteristics (1). GSs are uncommon tumors that most frequently arise in the stomach (2,4). Anatomically, GSs primarily occur in the gastric body and less frequently in the antrum and fundus, in that order (1). These GSs originate from the nerve sheath of Auerbach's plexus or, less commonly, Meissner's plexus (5). They are characterized as slow-growing, encapsulated neoplasms composed of Schwann cells embedded within a collagenous matrix. As the tumor expands, it displaces the nerve toward its periphery, thereby preserving neural function (4).

GSs predominantly affect individuals in the fifth to sixth decades of life, with a marked female preponderance. These neoplasms frequently remain clinically silent and may be discovered incidentally during radiological investigations or exploratory laparotomy. In symptomatic cases, GI hemorrhage is the most common clinical manifestation, followed by abdominal pain or discomfort (2,4,5). Hemorrhage may occur in the setting of deep ulceration, and a mass may be palpable in the epigastrium when exophytic tumor growth is present. Our patient's presentation with GI bleeding corroborates findings documented in the existing literature.

Preoperative differential diagnosis remains challenging; the primary difficulty is accurately distinguishing GSs, GISTs, and smooth muscle neoplasms (1). On CT imaging, GSs appear as intraluminal masses with exophytic or mixed growth patterns; they typically lack hemorrhage, necrosis, cystic degeneration, or calcification and demonstrate homogeneous enhancement on contrast-enhanced studies (1,2,6). Notably, smaller GISTs may manifest as hypervascular masses with pronounced enhancement on CT imaging, complicating their distinction from GSs. Gastric leiomyomas typically present on CT as homogeneous, hypoattenuating masses characterized by an endoluminal growth pattern and mild-to-moderate contrast enhancement. A distinguishing feature of leiomyomas is their predilection for the gastric cardia and esophagogastric junction.

Fluorodeoxyglucose positron emission tomography may show moderate uptake in GS, but it is not specific and cannot reliably distinguish GS from GIST, which often demonstrates higher metabolic activity (4). Endoscopic examination reveals a submucosal, elevated lesion with smooth, intact mucosa; central ulceration, due to ischemic changes, occurs in 25–50% of cases. Nevertheless, standard endoscopic biopsy techniques frequently fail to yield definitive histological diagnoses. This diagnostic limitation stems from the fact that SMTs are covered by normal epithelium, resulting in superficial biopsy specimens that demonstrate only normal mucosal tissue rather than the underlying pathology. endoscopic ultrasound–guided fine needle aspiration biopsy is established as a reliable and suitable method for obtaining histological diagnoses of SMTs, such as GISTs (7). Despite this utility, the time-intensive nature of the procedure meant it could not be performed on our patient, who presented with tumor-related bleeding.

Schwannomas present as distinct, encapsulated, nodular masses originating from the nerve sheath. Microscopically, they are spindle cell tumors characterized by an admixture of two classic patterns: compact, cellular Antoni A areas, often containing Verocay bodies (nuclear palisading), interspersed with hypocellular Antoni B areas (8). Features such as focal nuclear atypia and mitotic activity may be observed. The frequent presence of prominent thick-walled, hyalinized blood vessels is another diagnostic clue. By immunohistochemistry, schwannoma cells exhibit strong,

diffuse positivity for S100 protein. A negative immunoprofile for CD117, DOG1, CD34, SMA, and desmin is critical for distinguishing schwannomas from other common GI mesenchymal tumors. This negative immunophenotype readily differentiates schwannomas from GISTs, which express the KIT protein (CD117) and DOG-1, and from leiomyomas, which are typically positive for the smooth muscle markers SMA and desmin (9).

Malignant transformation of GI schwannomas is exceedingly rare, with only isolated cases of metastasis documented in the literature (10,11). Therefore, GSs are treated definitively by either en bloc resection or partial resection. A literature analysis showed satisfactory long-term outcomes and no statistically significant differences between patients undergoing local resection and those undergoing extended surgery (subtotal or total gastrectomy) (1). Small tumors can also be removed endoscopically. Cai et al. (12) reported 12 successful endoscopic cases, most of which were located in the gastric body, with a mean tumor size of 1.73 ± 1.10 cm. During a mean follow-up of 4 years, no recurrence or metastasis was detected (12). In our case, the intraoperative finding of suspected metastatic infrapyloric nodules highlights a diagnostic pitfall. While GSs are overwhelmingly benign, they can incite a significant peri-tumoral inflammatory or desmoplastic reaction. This can manifest as enlarged, firm perigastric lymph nodes or as adherent omental tissue, mimicking metastatic disease on gross inspection. Histopathological examination is crucial to differentiate these reactive changes from true malignancy.

Surgical resection remains the cornerstone of treatment for symptomatic or large GSs. Given their benign nature, organ-preserving and minimally invasive approaches are preferred when technically feasible. Laparoscopic resection of benign gastric tumors, including schwannomas, has been shown to be safe and effective, offering advantages of reduced postoperative pain, shorter hospital stays, and improved cosmetic outcomes compared with open surgery (13,14). For our patient, a laparoscopic distal gastrectomy was selected due to the tumor's large size and antral location, which precluded local excision. This approach also accommodated the need for a Roux-en-Y reconstruction in a patient with morbid obesity, providing a durable surgical solution while minimizing laparotomy-related morbidity in a high-risk patient with multiple comorbidities.

CONCLUSION

Consequently, definitive preoperative diagnosis of GSs remains challenging. Surgical resection can be performed safely via a minimally invasive approach, providing acceptable postoperative GI function while ensuring both a definitive histopathological diagnosis and curative treatment.

Ethics

Informed Consent: It was obtained from all patients.

Footnotes

Authorship Contributions: Surgical and Medical Practices: M.A.T., L.R., Y.K.Z., Concept: M.A.T., E.T.Ş., Design: M.A.T., Data Collection or Processing: M.A.T., L.R., S.A., Y.K.Z., E.T.Ş., Analysis or Interpretation: M.A.T., Literature Search: M.A.T., L.R., S.A., Y.K.Z., E.T.Ş., Writing: M.A.T., L.R., E.T.Ş.

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

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

REFERENCES

1. Lauricella S, Valeri S, Mascianà G, Gallo IF, Mazzotta E, Pagnoni C, et al. What about gastric schwannoma? A review article. J Gastrointest Cancer. 2021; 52: 57-67.
2. Yoon HY, Kim CB, Lee YH, Kim HG. Gastric schwannoma. Yonsei Med J. 2008; 49: 1052-4.
3. Türkoğlu MA, Ekinci M, Sert A, Alyaz BE, Gocun FPU. Is laparoscopic distal gastrectomy a justified approach for adult hypertrophic pyloric stenosis, a rare cause of gastric outlet obstruction? (with video). Gazi Medical Journal. 2025; 36: 221-5.
4. Fujiwara S, Nakajima K, Nishida T, Takahashi T, Kurokawa Y, Yamasaki M, et al. Gastric schwannomas revisited: Has precise preoperative diagnosis become feasible? Gastric Cancer. 2013; 16: 318-23.
5. Pu C, Zhang K. Gastric schwannoma: A case report and literature review. J Int Med Res. 2020; 4: 300060520957828.
6. Singh A, Mittal A, Garg B, Sood N. Schwannoma of the stomach: A case report. J Med Case Rep. 2016; 10: 4.
7. Hoda KM, Rodriguez SA, Faigel DO. EUS-guided sampling of suspected GI stromal tumors. Gastrointest Endosc. 2009; 69: 1218-23.
8. Wang G, Chen P, Zong L, Shi L, Zhao W. Cellular schwannoma arising from the gastric wall misdiagnosed as a gastric stromal tumor: A case report. Oncol Lett. 2014; 7: 415-8.
9. Yang JH, Zhang M, Zhao ZH, Shu Y, Hong J, Cao YJ. Gastroduodenal intussusception due to gastric schwannoma treated by Billroth II distal gastrectomy: One case report. World J Gastroenterol. 2015; 21: 2225-8.
10. Bees NR, Ng CS, Dicks-Mireaux C, Kiely EM. Gastric malignant schwannoma in a child. Br J Radiol. 1997; 70: 952-5.
11. Zheng L, Wu X, Kreis ME, Yu Z, Feng L, Chen C, et al. Clinicopathological and immunohistochemical characterisation of gastric schwannomas in 29 cases. Gastroenterol Res Pract. 2014; 2014: 202960.
12. Cai MY, Xu JX, Zhou PH, Xu MD, Chen SY, Hou J, et al. Endoscopic resection for gastric schwannoma with long-term outcomes. Surg Endosc. 2016; 30: 3994-4000.
13. Lee CM, Kim HH. Minimally invasive surgery for submucosal (subepithelial) tumors of the stomach. World J Gastroenterol. 2014; 20: 13035-43.
14. Takahashi K, Kanehira E, Kamei A, Tanida T, Sasaki K. Laparoscopic surgery for large gastric submucosal tumors. Surg Laparosc Endosc Percutan Tech. 2017; 27: 465-9.

Gazi Med J 2026;37(2):252-253

DOI: http://dx.doi.org/10.12996/gmj.2026.4599

Letter to the Editor: Comprehensive Prediction of FBN1 Targeting miRNAs: A Systems Biology Approach for Marfan Syndrome

Editöre Mektup: FBN1 Hedefli miRNA'ların Kapsamlı Tahmini: Marfan Sendromu için Sistem Biyolojisi Yaklaşımı

Burcu Baba, Ayşegül Akbay

Department of Medical Biochemistry, Yüksek İhtisas University, Faculty of Medicine, Ankara, Türkiye

Keywords: Aortopathy, biomarker, long non-coding RNA, microRNA, Marfan syndrome

Anahtar Sözcükler: Aortopati, biyobelirteç, uzun kodlamayan RNA, mikroRNA, Marfan sendromu

To the Editor,

In recent years, interest in small and long non-coding RNAs as modulators and potential biomarkers of the aortopathy that defines Marfan syndrome (MFS) has increased. Clinical and translational studies indicate that specific microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) contribute to extracellular matrix (ECM) remodeling, vascular smooth muscle cell (VSMC) phenotypic switching, apoptosis, and dysregulated TGF-β signaling, all of which are central processes in MFS. Synthesizing these findings highlights both opportunities and important gaps before RNA-based approaches can be applied clinically (1,2).

The study by Orhan et al. (1) provides a significant advancement in our understanding of MFS by presenting the first systems-level mapping of the miRNA-FBN1 interactome. By screening the entire human miRNome, they identified 251 potential regulators, thereby shifting the focus from candidate-driven research to a comprehensive regulatory landscape. Importantly, the hsa-miR-181 family, as a high-affinity regulator of FBN1, adds a novel layer to the established paradigm of ECM dysregulation in MFS. miR-181–mediated repression of FBN1 may act upstream of miR-29b–driven ECM remodeling, suggesting complementary but mechanistically distinct pathways that jointly compromise aortic wall integrity. Orhan et al. (1) provide a microfibril-specific mechanism by demonstrating multiple high-affinity binding sites on the FBN1 transcript, while established paradigms highlight the "anti-matrix" activity of miRNAs like miR-29b in general ECM degradation (2).

The miR-29 family, particularly miR-29b, is one of the most consistent players in MFS-associated aneurysm formation. Increased aortic miR-29b has been experimentally linked to increased ECM degradation and VSMC apoptosis, whereas antisense inhibition of miR-29b reduced early aneurysm development in murine MFS, which supports the idea that modulation of this miRNA may influence disease course. These preclinical results are supported by subsequent studies showing long-term benefits of systemic

Cite this article as: Baba B, Akbay A. Letter to the editor: comprehensive prediction of FBN1 targeting miRNAs: a systems biology approach for Marfan syndrome. Gazi Med J. 2026;37(2):252-253

Address for Correspondence/Yazışma Adresi: Burcu Baba, Department of Medical Biochemistry, Yüksek İhtisas University, Faculty of Medicine, Ankara, Türkiye
E-mail / E-posta: burcubaba@yiu.edu.tr
ORCID ID: orcid.org/0000-0003-0994-3577

Address for Correspondence/Yazışma Adresi: Ayşegül Akbay, Department of Medical Biochemistry, Yüksek İhtisas University, Faculty of Medicine, Ankara, Türkiye
E-mail / E-posta: aysegulakbay@yiu.edu.tr
ORCID ID: orcid.org/0009-0003-6793-4887

Received/Geliş Tarihi: 05.12.2025
Accepted/Kabul Tarihi: 24.02.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

miR-29b suppression in Marfan models, strengthening the rationale for developing miR-29-targeted therapies or biomarker panels.

Similarly, the miR-143/145 cluster maintains the contractile VSMC phenotype. Numerous studies report downregulation of miR-143/145 in thoracic aneurysmal tissues and VSMC phenotypic switching toward a synthetic, matrix-remodeling state—a cellular program highly relevant to MFS aortopathy. The fact that miR-143/145 regulates cytoskeletal and contractile gene networks and interfaces with TGF-β signaling, its dysregulation plausibly contributes to the weakened aortic media observed in patients (1,2,3).

Circulating miRNA signatures in patients with MFS have translational potential as noninvasive biomarkers. Case-control studies reveal distinct miRNA profiles in MFS patients, particularly in those with dissecting or rapidly enlarging aneurysms, suggesting that serum/plasma miRNA panels could aid in risk stratification or the early detection of complications. However, heterogeneity between studies in sample processing, cohort size, and inconsistent normalization strategies currently limits clinical application. Larger, prospective cohorts are required (4).

Beyond miRNAs, lncRNAs remain less well characterized in MFS, but increasing evidence indicates altered lncRNA landscapes in affected aortic tissue. Microarray and transcriptomic analyses of MFS aortas have demonstrated differentially expressed lncRNAs that correlate with ECM-related mRNAs and with pathways implicated in aneurysm biology. Reviews synthesizing coding and non-coding transcriptomic data report dysregulation of lncRNAs associated with TGF-β signaling, inflammation, and mitochondrial function—all plausible contributors to the variable penetrance and progression of MFS aortopathy. Nevertheless, functional validation and clinical correlation studies (e.g., linking specific lncRNA levels with rates of aortic dilation or dissection in patients) remain limited (5). The findings of Orhan et al. (1) underscore this gap and highlight the need to move beyond single miRNA–mRNA axes toward an integrated non-coding RNA network analysis, in which miRNAs and lncRNAs are evaluated as coordinated regulators of ECM homeostasis and vascular smooth muscle cell function. The systems-level strategy of Orhan et al. (1) provide a conceptual framework that could be extended to include lncRNAs, enabling reconstruction of higher-order non-coding RNA networks governing ECM integrity in MFS.

Taken together, the evidence supports that miRNAs and lncRNAs are central to molecular circuits that modulate aortic wall integrity in MFS. miR-29b and the miR-143/145 cluster remain strong candidates for translational development: miR-29b for therapeutic antagonism, and miR-143/145 for restoration or as biomarkers of VSMC health. Circulating miRNA panels show promise for noninvasive monitoring, whereas lncRNAs represent an intriguing but still nascent field with a substantial unmet need for mechanistic and longitudinal clinical studies.

To move this field forward, we need (1) standardized protocols for miRNA/lncRNA measurement in blood and tissue, (2) multicenter prospective cohorts linking RNA signatures to imaging-based aortic outcomes, and (3) careful safety assessment of any RNA-modulating therapy, given the systemic roles of these molecules. If these hurdles are addressed, non-coding RNAs could become an important adjunct to genetic testing and imaging in the personalized management of patients with Marfan syndrome.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Orhan ME, Demirci YM, Demirci MDS. Comprehensive prediction of FBN1 targeting miRNAs: a systems biology approach for Marfan syndrome. Gazi Med J. 2025; 36: 401-6.
2. Udugampolage NS, Frolova S, Taurino J, Pini A, Martelli F, Voellenkle C. Coding and non-coding transcriptomic landscape of aortic complications in Marfan syndrome. Int J Mol Sci. 2024; 25: 7367.
3. D'Amico F, Doldo E, Pisano C, Scioli MG, Centofanti F, Proietti G, Falconi M, Sangiuolo F, Ferlosio A, Ruvolo G, Orlandi A. Specific miRNA and gene deregulation characterize the increased angiogenic remodeling of thoracic aneurysmatic aortopathy in Marfan syndrome. Int J Mol Sci. 2020; 21: 6886.
4. Abu-Halima M, Kahraman M, Henn D, Rädle-Hurst T, Keller A, Abdul-Khaliq H, Meese E. Deregulated microRNA and mRNA expression profiles in the peripheral blood of patients with Marfan syndrome. J Transl Med. 2018; 16: 60.
5. Gu L, Ni J, Sheng S, Zhao K, Sun C, Wang J. Microarray analysis of long non-coding RNA expression profiles in Marfan syndrome. Exp Ther Med. 2020; 20: 3615-24.

Gazi Med J 2026;37(2):254-255

DOI: http://dx.doi.org/10.12996/gmj.2025.4547

Resistance to Thyroid Hormone Complicating the Management of Papillary Thyroid Carcinoma

Papiller Tiroid Karsinomu Yönetimini Komplike Eden Tiroid Hormon Direnci

Ragıp Fatih Kural[1], Güzide Gonca Oruk[2]

[1]Division of Immunology and Allergy, Department of Internal Medicine, Ege University, Faculty of Medicine, İzmir, Türkiye
[2]Division of Endocrinology and Metabolism, Department of Internal Medicine, Katip Çelebi University, Faculty of Medicine, İzmir, Türkiye

Keywords: Resistance to thyroid hormone, papillary thyroid carcinoma, TSH suppression, cervical lymph node recurrence, family screening

Anahtar Sözcükler: Tiroid hormonu direnci, papiller tiroid karsinomu, TSH baskılanması, servikal lenf nodu nüksü, aile taraması

To the Editor,

In this letter, we aim to draw attention to the clinical challenges encountered when papillary thyroid carcinoma (PTC) coexists with resistance to thyroid hormone (RTH). RTH is a rare autosomal-dominant disorder (≈1 in 40,000 births), usually due to mutations in the thyroid hormone receptor beta. It presents with elevated thyroid hormone levels and a non-suppressed thyroid-stimulating hormone (TSH) (1). Its coexistence with PTC is exceedingly rare but important, as persistently high TSH may promote tumorigenesis (2).

We present a case of a 38-year-old woman with anterior neck swelling. She was clinically euthyroid, with no family history of malignancy. Tests revealed elevated free T4 (2.11 ng/mL) and free T3 (6.25 pg/mL) with a non-suppressed TSH (3.28 µIU/mL). Autoantibodies were negative, and pituitary magnetic resonance imaging (MRI) was normal, thereby excluding TSHoma. Family screening revealed similar abnormalities in her two children, consistent with familial RTH. Neck ultrasound revealed multiple hypoechoic nodules; the largest, in the left lobe (13 × 9 × 7 mm), was cytologically classified as Bethesda V. The patient underwent total thyroidectomy with central lymph node dissection, and histopathology confirmed classical PTC (T1bN1bM0) with metastasis in the central cervical lymph nodes (Figure 1). Postoperatively, levothyroxine was initiated for TSH suppression, but despite dose escalation from 150 to 350 µg/day, TSH rose to 176.3 µIU/mL at 6 months. At 12 months, cervical recurrence was detected, requiring repeat dissection and radioiodine ablation. Following completion of surgery, the levothyroxine dose was further increased to 425 µg/day to achieve greater TSH suppression; however, due to the development of hyperthyroid symptoms, the dose was reduced to 350 µg/day, which remained the highest tolerated dose. Over the subsequent three years, TSH gradually declined (176.3 → 92.4 → 55.7 → 21.8 µIU/mL) as medication adherence improved, while free T3 and T4 levels remained within reference ranges. At her latest evaluation, serum thyroglobulin was undetectable (<0.20 ng/mL); she had no hypothyroid symptoms; and neck ultrasound revealed no pathological lymphadenopathy.

Cite this article as: Kural RF, Oruk GG. Resistance to thyroid hormone complicating the management of papillary thyroid carcinoma. Gazi Med J. 2026;37(2):254-255

Address for Correspondence/Yazışma Adresi: Ragıp Fatih Kural, Division of Immunology and Allergy, Department of Internal Medicine, Ege University, Faculty of Medicine, İzmir, Türkiye
E-mail / E-posta: ragip.fatih.kural@ege.edu.tr
ORCID ID: orcid.org/0000-0001-5690-4090

Received/Geliş Tarihi: 14.09.2025
Accepted/Kabul Tarihi: 07.12.2025
Publication Date/Yayınlanma Tarihi: 31.03.2026

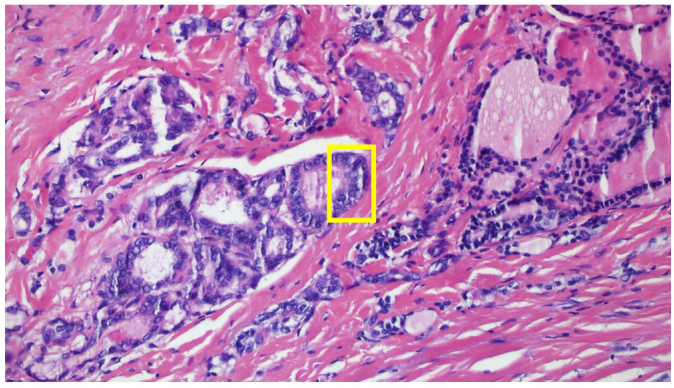

Figure 1. Histopatology of thyroid gland showing papillary thyroid carcinoma, classical variant: characterized by enlarged and overlapping nuclei (yellow box).

This case illustrates the difficulties in managing PTC in the setting of RTH. When free T3/T4 levels are elevated but TSH is not suppressed, the main differential diagnosis is a TSH-secreting pituitary adenoma. In our patient, the pituitary MRI was normal, supporting the diagnosis of RTH. Although genetic confirmation is recommended (3), it was not available at our center; instead, family screening supported the diagnosis.

Persistent TSH elevation despite high-dose levothyroxine is a hallmark of RTH and complicates cancer management. In our case, TSH stayed high despite >300 µg/day of levothyroxine; this is consistent with reports that suppression is rarely achievable (4), and no secondary cause, such as malabsorption or drug interactions, could explain this finding. Occasional noncompliance during the early postoperative period may also have contributed to the excessive elevation in TSH. This is clinically important, as elevated TSH may stimulate tumor growth and recurrence (2). Our patient developed regional recurrence within one year, underlining this risk. Alternative therapies such as 3,5,3'-triiodothyroacetic acid (TRIAC) can suppress TSH in RTH without causing thyrotoxic symptoms (5). However, TRIAC was not used in our case due to limited availability in our country and because biochemical improvement was gradually achieved with the maximally tolerated dose of levothyroxine alone.

The coexistence of RTH and thyroid carcinoma is exceedingly rare, with only a small number of cases described worldwide (1-3) and very few reported from Türkiye (4). Our case illustrates the practical challenges of this association, including early recurrence, difficulty achieving TSH suppression despite high-dose therapy, and the value of family screening in reaching the diagnosis. These features highlight the need for individualized care and long-term follow-up, with close monitoring for recurrence and careful adjustment of thyroid hormone therapy to balance cancer control with overall metabolic health.

Ethics

Informed Consent: Written informed consent was obtained from the patient for publication.

Footnotes

Authorship Contributions

Surgical and Medical Practices: R.F.K., G.G.O., Concept: R.F.K., G.G.O., Design: R.F.K., Data Collection or Processing: R.F.K., Analysis or Interpretation: R.F.K., G.G.O., Literature Search: R.F.K., Writing: R.F.K.

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

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

Acknowledgments

We thank the Department of Pathology, Faculty of Medicine, İzmir Katip Çelebi University for their support in providing histopathological images.

REFERENCES

1. Igata M, Tsuruzoe K, Kawashima J, Kukidome D, Kondo T, Motoshima H, et al. Coexistence of resistance to thyroid hormone and papillary thyroid carcinoma. Endocrinol Diabetes Metab Case Rep. 2016; 2016: 160003.
2. Paragliola RM, Lovicu RM, Locantore P, Senes P, Concolino P, Capoluongo E, et al. Differentiated thyroid cancer in two patients with resistance to thyroid hormone. Thyroid. 2011; 21: 793-7.
3. Fang Y, Liu T, Hou H, Wang Z, Shan Z, Cao Y, et al. Resistance to thyroid hormone beta coexisting with papillary thyroid carcinoma-two case reports of a thyroid hormone receptor beta gene mutation and a literature review. Front Genet. 2022; 13: 1014323.
4. Karakose M, Caliskan M, Arslan MS, Cakal E, Yesilyurt A, Delibasi T. Thyroid hormone resistance in two patients with papillary thyroid microcarcinoma and their BRAFV600E mutation status. Arch Endocrinol Metab. 2015; 59: 364-6.
5. Ramos-Prol A, Antonia Pérez-Lázaro M, Isabel del Olmo-García M, León-de Zayas B, Moreno-Macián F, Navas-de Solis S, et al. Differentiated thyroid carcinoma in a girl with resistance to thyroid hormone management with triiodothyroacetic acid. J Pediatr Endocrinol Metab. 2013; 26: 133-6.

Gazi Med J 2026;37(2):256-257

DOI: http://dx.doi.org/10.12996/gmj.2025.4511

Comments on "Transanal Specimen Extraction After Laparoscopic Sigmoidectomy for Sigmoid Volvulus"

"Sigmoid Volvulusta Laparoskopik Sigmoidektomi Sonrası Transanal Yolla Spesmen Çıkarılması" Üzerine Tartışmalar

Sabri Selçuk Atamanalp

Department of General Surgery, Atatürk University, Faculty of Medicine, Erzurum, Türkiye

ABSTRACT

Sigmoid volvulus (SV) arises from the twisting of the sigmoid colon around itself and tends to recur. The optimal treatment is an elective sigmoidectomy following endoscopic detorsion. In this field, laparoscopic sigmoid colectomy (LSC) with natural orifice specimen extraction (NOSE) is the most recent approach. In this paper, my comments pertain to LSC with NOSE for SV.

Keywords: Sigmoid volvulus, laparoscopic sigmoidectomy, transanal specimen extraction

ÖZ

Sigmoid kolonun kendi etrafında dönmesi sonucu ortaya çıkan nadir bir kapalı lup kolon tıkanıklığı şekli olan sigmoid volvulus (SV), nüks etmeye meyillidir. En iyi tedavi seçeneği endoskopik detorsiyon sonrası elektif sigmoidektomidir. Bu alanda, laparoskopik sigmoid kolektomi (LSK) ile birlikte doğal yolla spesmen çıkarılması (DYSÇ), seçilmiş olgularda en güncel yaklaşımdır. Bu yazıda tartışmalarım, SV'de LSK ve DYSÇ ile ilgilidir.

Anahtar Sözcükler: Sigmoid volvulus, laparoskopik sigmoidektomi, doğal yolla spesmen çıkarılması

Dear Sir,

I read with great pleasure the article by Uylaş et al. (1) on laparoscopic sigmoid colectomy (LSC) with natural orifice specimen extraction (NOSE) for Sigmoid volvulus (SV); it reports one of the largest series in this field. Although SV is uncommon in Western populations, Eastern Anatolia, our practice area, is an endemic region for SV (2). As a consequence, our SV series, comprising 1,096 patients treated over 59 years (from June 1966 to July 2025), is the largest published SV series in the world (3). Among 763 patients (93.8%) with viable bowel in our series, 638 cases (83.6%) underwent successful endoscopic detorsion. Elective sigmoid colectomy was performed in 124 patients (19.4%; 95 open and 29 laparoscopic procedures), with no mortality, a morbidity rate of 1.3%, and no recurrences. However, we have no cases treated with LSC and NOSE in this series.

LSC with NOSE represents the most recent advancement in the treatment of SV. However, a database search of the literature in Web of Science (3), under the heading "sigmoid volvulus" and covering the last 80 years (1945–2025), revealed only a few (ten) SV cases treated with LSC with NOSE. Among those, the largest series, presenting 16 patients, was reported by Chen et al. (4). Due to the rarity of such cases, some controversy remains regarding this subject.

First, all reports, including the discussed paper, describe elective cases treated with endoscopic detorsion, and the results of LSC with NOSE are uncertain in emergency SV cases. Similarly, the outcomes of this procedure in children, elderly individuals, pregnant women, and morbidly obese individuals are unclear (1,4,5).

Cite this article as: Atamanalp SS. Comments on "transanal specimen extraction after laparoscopic sigmoidectomy for sigmoid volvulus". Gazi Med J. 2026;37(2):256-257

Address for Correspondence/Yazışma Adresi: Sabri Selçuk Atamanalp, Department of General Surgery, Atatürk University, Faculty of Medicine, Erzurum, Türkiye
E-mail / E-posta: ssa@atauni.edu.tr
ORCID ID: orcid.org/0000-0003-2561-6461

Received/Geliş Tarihi: 29.07.2025
Accepted/Kabul Tarihi: 07.12.2025
Epub: 05.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

Second, there are some controversies regarding the technical approach. Some practitioners use laparoscopic procedures alone, while others prefer robotic surgery. Although the authors used four 5- and 12-mm trocars in the present paper, instruments reported in the literature are generally heterogeneous, including different numbers and sizes of trocars. Similarly, the authors used the transanal extraction route for small specimens; however, they suggested the transvaginal route for larger ones. Although the authors preferred splitting the specimens, total specimen extraction remains an alternative technique (1,4,5).

Third, following the resection, the choice of anastomotic technique remains controversial. Although the authors generally used side-to-end anastomosis, with one case employing a side-to-side procedure, other anastomotic techniques, including end-to-end and end-to-side procedures are also possible. On the other hand, the authors placed a rectal device for three days. Although rectal tubes may discharge the gas and preserve the anastomosis, their use is widely debated (1,4,5).

Finally, as demonstrated by the authors, both relatively prolonged operative time and relatively increased cost remain controversial issues in LSC with NOSE for SV (1,4,5).

I appreciate the authors' didactic presentation, and I wonder about the authors' opinion on my comments. It is clear that prospective studies are needed to standardize LSC with NOSE for SV.

Footnotes

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

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

REFERENCES

1. Uylaş U, Gündoğdu R, Bağ YM, Aktaş A, Sümer F, Kayaalp C. Transanal specimen extraction after laparoscopic sigmoidectomy for sigmoid volvulus. Gazi Med J. 2025; 36: 245-9.
2. Atamanalp SS, Disci E, Peksoz R, Atamanalp RS, Tatar Atamanalp C. Management of sigmoid volvulus: a literature review. Ibnosina J Med Biomed Sci. 2024; 16: 5-9.
3. Sigmoid volvulus. Web of Science [Internet]. Available from: https://www.webofscience.com/wos/woscc/summary/3964f12d-4de8-476d-bc61-62fb708ac9aa-016ce4d39e/relevance/1.
4. Chen MZ, Cartmill J, Gilmore A. Natural orifice specimen extraction for colorectal surgery: early adoption in a Western population. Colorectal Dis. 2021; 23: 937-43.
5. Atamanalp SS. Laparoscopic sigmoid colectomy with natural orifice specimen extraction in sigmoid volvulus. Eurasian J Med. 2024; 56: 142–5.

Gazi Med J 2026;37(2):258-264

DOI: http://dx.doi.org/10.12996/gmj.2026.4587

Recent Advances in Obesity Biology: From Genetics to Transgenerational Effects

Obezite Biyolojisindeki Son Gelişmeler: Genetikten Nesiller Arası Etkilere

Khalida I. Noel, Sameh S. Akkila, Nibras H. Khamees

[1]Department of Human Anatomy, Histology and Embryology, College of Medicine, Mustansiriyah University, Baghdad, Iraq

ABSTRACT

Obesity, which is a global problem regarded as a complex disorder, is influenced by environmental, genetic, and epigenetic factors. Genetic discoveries, epigenetic alterations, nutritional roles, hormonal effects, inflammation, and the precise problem of middle-aged abdominal fat development are prominent factors. This article provides a summary of current theories related to the biology of obesity. The distinct role of neuroestrogens in the development of obesity has also been described. This review also explores how maternal obesity affects the development of the fetal liver and subsequent childhood obesity, emphasizing the long-term metabolic effects of maternal overnutrition. The article also underlines the latest data concerning adolescent obesity and its influence on the subsequent development of obesity in offspring and the impact of paternal obesity, not only maternal obesity, on offspring. New prospects for clinical study and medication development are illustrated by the newly identified role of neuroestrogens in appetite regulation and energy balance. In addition, creating efficient management and prevention measures for obesity requires an understanding of the mechanisms that cause increases in abdominal fat in middle-aged individuals, such as hormonal changes, metabolic alterations, and lifestyle factors. A conceptual shift from late-stage obesity care to early, preventive, and personalized interventions is supported by the clinical application of mechanistic insights from developmental biology, genetics, and epigenetics.

Keywords: Adipose progenitors, neuroestrogen, epigenetics, FGF19, CP-As

ÖZ

Karmaşık bir bozukluk olarak kabul edilen küresel bir sorun olan obezite, çevresel, genetik ve epigenetik faktörlerden etkilenir. Genetik keşifler, epigenetik değişiklikler, beslenme rolleri, hormonal etkiler, inflamasyon ve orta yaşta karın bölgesinde yağlanmanın kesin sorunu öne çıkan faktörlerdir. Bu makale, obezitenin biyolojisiyle ilgili güncel teorilerin bir özetini sunmaktadır. Nöroöstrojenlerin obezite gelişimindeki belirgin rolü de açıklanmıştır. Bu derleme ayrıca, anne obezitesinin fetal karaciğer gelişimini ve daha sonraki çocukluk çağı obezitesini nasıl etkilediğini, anne aşırı beslenmesinin uzun vadeli metabolik etkilerini vurgulayarak incelemektedir. Makale ayrıca, ergenlik obezitesi ve bunun yavruda obezite gelişimine etkisi ile ilgili en son verilerin ve sadece anne obezitesinin değil, baba obezitesinin de yavru üzerindeki etkisinin altını çizmektedir. Nöroöstrojenlerin iştah düzenlemesi ve enerji dengesindeki yeni tanımlanmış rolü, klinik çalışma ve ilaç geliştirme için yeni perspektifler göstermektedir. Ayrıca, obezite için etkili yönetim ve önleme tedbirleri oluşturmak, orta yaşlı bireylerde karın bölgesindeki yağ artışına neden olan mekanizmaları (hormonal değişiklikler, metabolik değişimler ve yaşam tarzı faktörleri gibi) anlamayı gerektirir. Gelişim biyolojisi, genetik ve epigenetikten elde edilen mekanistik bilgilerin klinik uygulaması, obezite bakımında geç aşamadan erken, önleyici ve kişiselleştirilmiş müdahalelere doğru kavramsal bir geçişi desteklemektedir.

Anahtar Sözcükler: Yağ dokusu öncü hücreleri, nöroöstrojen, epigenetik, FGF19, CP-As

Cite this article as: Noel KI, Akkila SS, Khamees NH. Recent advances in obesity biology: from genetics to transgenerational effects. Gazi Med J. 2026;37(2):258-264

Address for Correspondence/Yazışma Adresi: Khalida I. Noel, Department of Human Anatomy, Histology and Embryology, College of Medicine, Mustansiriyah University, Baghdad, Iraq
E-mail / E-posta: dr.khalidanoel@uomustansiriyah.edu.iq
ORCID ID: orcid.org/0000-0001-6135-7087

Received/Geliş Tarihi: 08.11.2025
Accepted/Kabul Tarihi: 22.01.2026
Epub: 11.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

INTRODUCTION

Obesity is a global health concern that increases the risk of type 2 diabetes, heart disease, and several types of cancer, and is associated with other metabolic health concerns (1). The genetic and epigenetic basis of obesity, as well as the role of nutrition and lifestyle in its development and treatment, has been clarified by recent advances in obesity biology (2). This article discusses recent research on the various pathways involved in obesity and highlights important subjects linked to adolescent and maternal obesity.

Genetic Insights into Obesity

Recent genetic studies have identified many loci linked to obesity, offering insights into its genetic architecture. More than 97 loci associated with body mass index (BMI) were identified in a study by Locke et al. (3), who demonstrated the polygenic nature of obesity. These genetic correlations suggest the molecular processes underlying adipogenesis and energy homeostasis. For example, variations in the *FTO* gene affect caloric intake and expenditure, and this gene has been repeatedly associated with obesity (4). However, according to a recent study by Künzel et al. (5), morbid obesity is not linked to genes, whereas only milder forms of obesity are. Lonky's (6) explanation of the heredity of obesity, which is nearly entirely epigenetic, represents an insightful contribution. Therefore, how genes are turned on or off in response to environmental factors such as pollutants, diet, sleep, and prenatal exposure than the DNA sequence does (6).

Epigenetic Modifications and Obesity

Environmental influences can affect epigenetic alterations, including DNA methylation and histone modifications, which are important in controlling gene expression (7). Macartney-Coxson et al. (8)'s genome-wide DNA methylation analysis revealed that people with obesity have distinct methylation patterns across adipose tissue types. These epigenetic modifications may affect metabolic processes and adipogenesis, potentially leading to obesity. Moreover, famine exposure during pregnancy leads to long-lasting epigenetic alterations associated with metabolic diseases (8).

The Role of Diet in Obesity and Metabolic Health

Obesity and metabolic health are influenced by the type of diet. A balanced diet is essential for preserving metabolic health, a conclusion consistent with a recent analysis by Wali et al. (9) that explored the effects of macronutrients on obesity and IR. In animal models, high-carb diets have been shown to cause obesity and alter genes linked to inflammation and eating habits. A holistic approach to dietary studies that considers the interconnections between macronutrients may provide more thorough insights into obesity, according to the geometric context for nutrition.

Hormonal Effects and Obesity

The development of obesity is considerably affected by hormonal regulation, with new research highlighting the significance of brain estrogen (neuroestrogen) in energy balance and appetite control.

Brain Oestrogen (Neuroestrogen) and Obesity

A new role of neuroestrogen, a form of estrogen produced in the brain, in regulating hunger and body weight. According to a study by Hayashi et al. (10), neurogenesis increases the expression of the hypothalamic melanocortin-4 receptor *(MC4R)*, a crucial receptor implicated in appetite regulation. Compared with healthy control mice, mice lacking ovaries or the aromatase enzyme, which is required for the generation of neuroestrogen, presented greater food intake and body weight. However, *MC4R* expression increased, and food intake decreased when the aromatase gene was specifically reactivated in the brains of these mice (Figure 1). This finding makes neuroestrogen a viable target for the development of novel therapies for obesity (10).

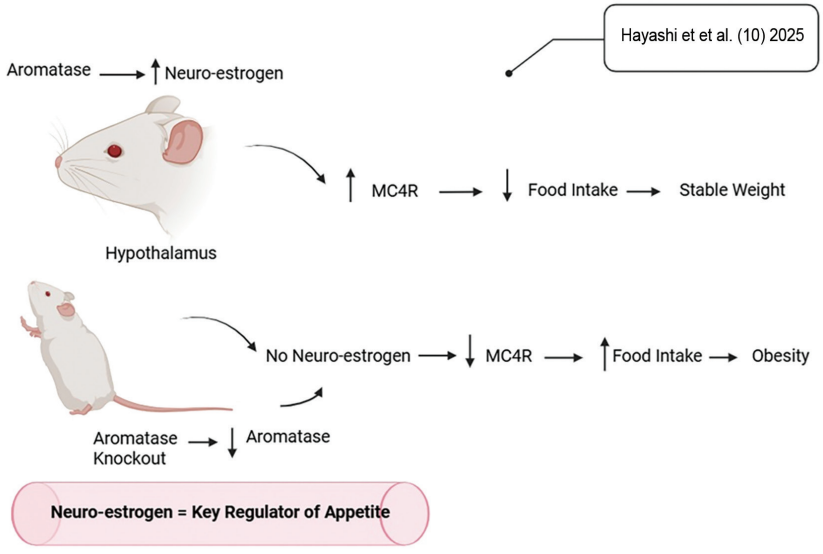

Figure 1. Role of neuroestrogen in regulating appetite and body weight.
MC4R: Melanocortin-4 receptor.

Hormones have been shown to speed up fat burning and aid in weight loss in obese rats. In an experiment, the intestinal production of fibroblast growth factor 19 (FGF19), affects brain regions, causing increased energy expenditure for heat production. This finding opens the door to novel medications. In obese animals, FGF19 activates pathways that increase energy expenditure, promote fatty acid oxidation, and support regulation of blood glucose and body weight. These outcomes are linked to the activity of FGF19 in the hypothalamus, a specific brain region that integrates ambient and peripheral cues to regulate energy metabolism. The authors reported that increased thermogenic adipocyte activity, i.e., the activity of fat cells that burn energy to produce heat, results from FGF19 signalling in the hypothalamus (Figure 2) (11).

Hypoxia and Inflammation in Obesity

A key factor in the pathophysiology of obesity-related insulin resistance (IR) is the interaction among hypoxia, inflammation, and metabolic dysregulation. Adipose tissue, especially visceral adipose tissue, secretes a variety of bioactive compounds that affect systemic metabolic processes, making it an active endocrine organ (12). Hypoxia is a common feature of expanding adipose tissue in obesity, initiating a series of molecular processes that exacerbate IR and inflammation. One important factor in these processes is hypoxia-induced dysregulation of miRNAs and adipokines. Proinflammatory cytokines and adipokines, including vascular endothelial growth factor, adiponectin, and leptin, are responsible for the metabolic dysregulation and inflammatory environment observed in obese individuals. Furthermore, miRNAs alter gene expression, impacting inflammatory and insulin receptor signalling pathways. Although our understanding of these systems has advanced considerably, many unresolved questions remain. Determining the sequence of the molecular events that trigger IR, the primary contributing variables, and the interactions among various signalling pathways remains necessary (13). Thus, hypoxia, or low oxygen levels in adipose tissue, may inhibit weight loss and cause inflammation (6).

Why Does Belly Fat Expand in Middle Age?

Fat accumulation in middle-aged individuals is common and is influenced by several factors, including lifestyle choices, hormonal changes, and metabolic adaptations (14).

Hormonal Changes

Decreasing Estrogen Levels: One of the main reasons for increased abdominal fat in women is the decline in estrogen levels after menopause. The distribution and storage of fat are considerably affected by estrogen; when estrogen levels decline, fat is redistributed from the hips and thighs to the belly (15).

Stress and Cortisol: Cortisol, the stress hormone, promotes the accumulation of abdominal fat. Increased visceral fat storage can result from elevated cortisol levels, which are commonly produced in response to chronic stress. A lack of sleep contributes to this worsening by altering the balance of appetite-regulating hormones, such as ghrelin and leptin (16).

Metabolic Shifts

Slowing Metabolism: People's metabolism slows with age, making it more difficult to burn calories effectively. This metabolic slowdown, associated with a natural decrease in muscle mass, contributes to increased fat storage, particularly in the abdominal area (17).

Cellular Alterations: Recent studies indicate that age-related cellular alterations can increase adipogenesis. As people age, adipocyte progenitor cells in adipose tissue become more active, promoting adipogenesis and increasing abdominal adiposity. Age-enriched

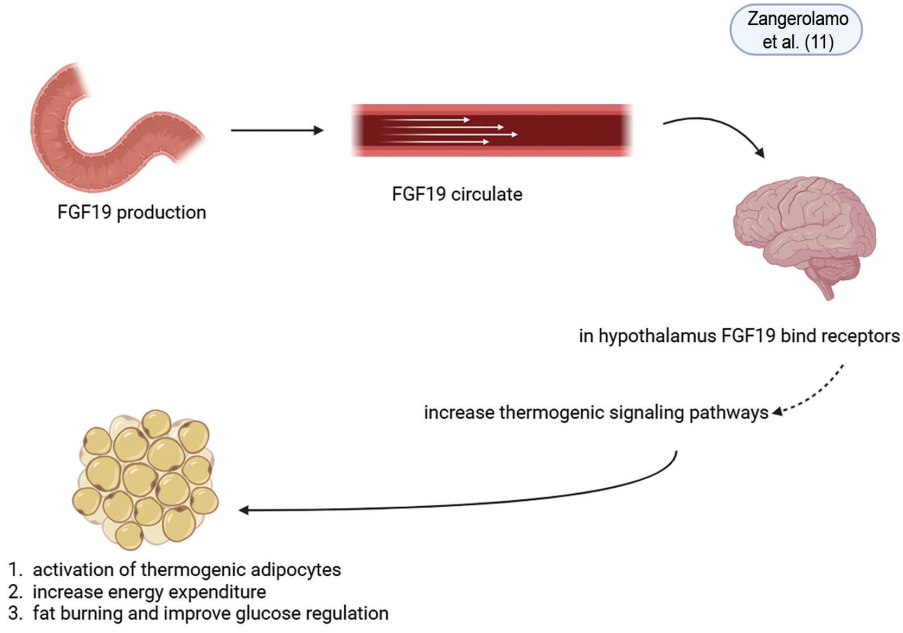

Figure 2. FGF19 promotes fat burning and weight loss via the hypothalamic action.

FGF19: Fibroblast growth factor 19.

committed preadipocytes (CP-As) constitute a novel population of fat-cell precursors that arise particularly during middle age and contribute to the rapid accumulation of visceral fat, according to a study published in science (18).

Emergence of CP-As: These special progenitor cells are almost non-existent in children, but become prevalent in middle-aged adults, particularly in individuals with excess abdominal fat. The activation of CP-As to generate new fat cells depends on the leukemia inhibitory factor receptor (LIFR) pathway. Visceral fat growth in mice was inhibited by blocking this route (19). Similarly, human tissue samples demonstrated the presence and activity of CP-As, indicating a comparable process. This study advances our understanding of age-related fat accumulation, moving beyond lifestyle influences to encompass specific cellular changes. There may be novel ways to treat or prevent visceral obesity and related health concerns by targeting CP-As or the LIFR pathway (Figure 3) (18).

Lifestyle Factors

Diet and Exercise: High-calorie diets and sedentary lifestyles are major contributors to weight gain in middle-aged individuals. A balanced diet rich in fruits, vegetables, and lean meats, combined with regular exercise, can help counteract this effect (20).

Stress and Sleep Management: Hormonal balance can be disrupted by excessive stress and inadequate sleep, resulting in increased appetite and cravings for high-calorie foods. Abdominal fat can be reduced through relaxation techniques to manage stress and by ensuring adequate sleep (21).

Maternal Obesity and Its Effects on the Fetal Liver and Future Child Obesity

In addition to increasing the risk of obesity and metabolic diseases in progeny, maternal obesity and overnutrition during pregnancy can have major long-term effects on fetal liver development.

Fetal Metabolic Disorders Associated with Maternal Obesity

Maternal diet may have a persistent impact on fetal gene expression through epigenetic mechanisms that lead to metabolic disorders. Because of transgenerational inheritance, epigenetic changes that occur during crucial stages of fetal development may have long-term effects. Determining the relationship between epigenetic changes and clinical and molecular outcomes in children associated with maternal obesity is crucial (22).

Maternal Obesity, High-Fat Diet, and Inflammation

Hyperlipidemia, systemic IR, and inflammation of adipose tissue are linked to maternal obesity. Proinflammatory cytokines can be activated during pregnancy by high-fat diets (FFTs) (23). Maternal IR and inflammation cause increased adipose tissue lipolysis and uptake of free fatty acids (FFAs). Maternal during pregnancy HFD significantly increased fetal FFA levels. Early-life obesity is exacerbated by both maternal IR and HFD (24).

Systemic inflammation, which includes elevated levels of tumor necrosis factor-alpha and monocyte chemoattractant protein-1, is positively associated with maternal BMI. An HFD has been shown to increase placental production of proinflammatory cytokines. Toll-like receptor 4 activation via FFAs may trigger the c-Jun N-terminal kinase and nuclear factor kappa B inflammatory signalling pathways in obese animals. Maternal HFDs result in IR through inflammatory alterations in fetal adipose tissue (25).

Epigenetic Mechanisms: Fetal metabolic disorders associated with maternal obesity.

An HFD during pregnancy caused the fetal liver to exhibit hyperacetylation of histones H3K14, H3K9, and H3K18, increased *DNMT1* expression, decreased *HDAC1* expression, and elevated hepatic triglyceride levels. Recent research indicates that maternal HFD induces metabolic programming of the fetal liver and heart

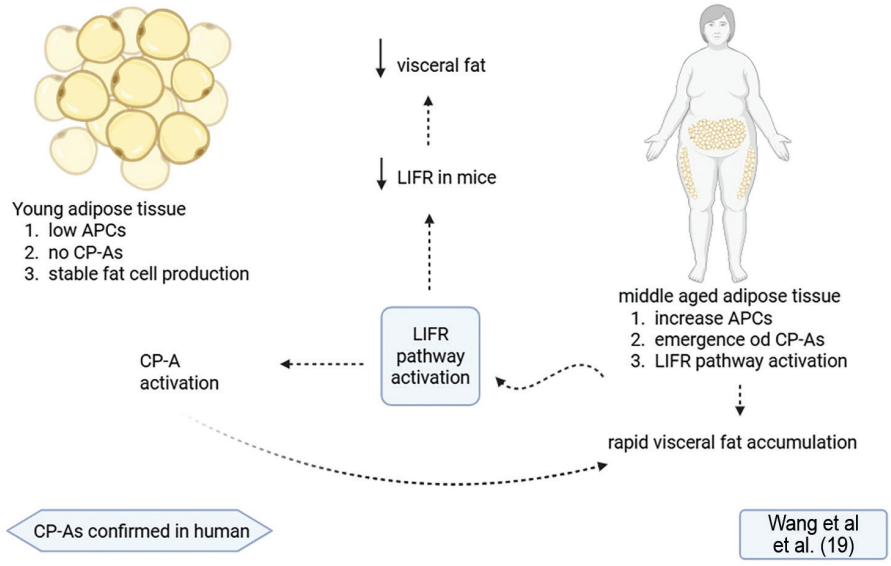

Figure 3. Age-related cellular changes promote visceral fat accumulation via CP-As.

CP-As: Committed preadipocytes, LIFR: Leukemia inhibitory factor receptor, APCs: Adipocytes precursor cells.

by decreasing *SIRT1* expression and increasing histone H3K14 acetylation (26).

Kupffer Cell Programming by Maternal Obesity

According to a recent study, maternal obesity causes HIF1α-dependent metabolic reprogramming of yolk sac-derived Kupffer cells (KCs) that persists into adulthood and leads to fatty liver disease (FLD). Using multiomic profiling together with fate mapping, depletion, and conditional knockout models, researchers discovered that KCs act as intergenerational messengers converting maternal dietary signals into chronic liver dysfunction (27).

The inflammatory response of KCs induced by maternal obesity alters their metabolic state at the transcriptional level. Through paracrine signalling, this reprogramming causes FLD and reduces the metabolic potential of progeny KCs (28).

A novel pathway linking maternal obesity to metabolic disease in offspring, independent of postnatal lifestyle, has recently been described in Nature. During pregnancy, fetal KCs undergo epigenetic reprogramming (27).

Cellular metabolism changes from oxidative phosphorylation to glycolysis as a result of oxidative phosphorylation to glycolysis (27).

Even in the absence of a "bad" diet, metabolic changes cause KCs to encourage fat buildup and inflammation in the livern. The immune system of the liver retains memories of early encounters. In addition to affecting delivery outcomes, maternal obesity leaves biological imprints that may lead to chronic illness. Because more than 25% of American adults suffer from NAFLD, this study highlights that immediate maternal health measures are essential. Diagnosis is not the first step in prevention. It begins prior to birth (28).

Impact on Fetal Liver Development

Significant alterations in the fetal liver, such as elevated lipid accumulation and oxidative stress, can result from maternal obesity and an HFD. Maternal HFD/obesity was associated with increased expression of lipid metabolism-related genes in the fetal liver, including ACC isoform 1 and LPL. Furthermore, the development of fatty liver is linked to elevated indicators of oxidative stress and decreased levels of antioxidant enzymes in the fetal liver (29,30).

Long-Term Consequences for Offspring

Offspring exposed to FFTs and maternal obesity during pregnancy experience long-term metabolic effects, such as a greater chance of becoming obese, developing IR, and suffering from FLD than adults do. The significance of maternal nutrition during pregnancy for long-term health consequences is highlighted by the fact that these effects are mediated by metabolic reprogramming of hepatic cells and epigenetic alterations (24). The following actions must be taken:

• Healthcare executives should incorporate weight management and prenatal nutrition into standard care as lifestyle and disease-prevention measures.

• Employers and insurers can fund maternal wellness initiatives that lower the risk of chronic diseases in the next generation and yield long-term returns on investment.

• Public health groups should expand campaigns that demonstrate how prenatal exposures affect our children's health well into adulthood.

• Researchers and clinicians need to look for biomarkers to identify KC dysregulation before symptoms appear.

• The prevention of liver disease may begin in the womb.

Effect of Teenage Obesity on Child's DNA

In a recent epigenome-wide association study published in Nature Communications Biology, 739 participants were tracked across two generations. This study examined the relationships between adolescent body-shape changes, especially during voice breaks, and DNA methylation patterns in offspring, focusing on 339 father-child pairs (31).

Fathers who gained additional weight throughout puberty exhibited the most pronounced effects. More than 2,000 methylation differences—chemical indicators that control gene activity—were detected in their offspring. Many of these variations (31) are located in genes involved in inflammation, lung function, and fat metabolism.

Methylation alterations were detected in imprinted genes, such as *VTRNA21* and *BLCAP*, which are sensitive to environmental influences, and in genes important for metabolic regulation, such as *KCNJ10*, *NCK2*, and *ATP5B*. These alterations are particularly evident in daughters, who exhibit sex-specific patterns of inheritance (32).

The timing of puberty is crucial. The biological effects that occur during this brief embryonic stage could last into subsequent generations.

This, together with maternal programming highlights that the health of both parents during critical periods—puberty and pregnancy—contributes to the transgenerational inheritance of obesity risk.

Epigenetic mechanisms alter gene expression without changing the DNA sequence. Not only are we transferring genes, but we are also transferring their behavior.

Adolescent obesity is a chronic issue. This may leave the following generation with epigenetic fingerprints.

Human Research Equivalent to Animal Research

Many studies in humans have taken into account findings from experimental animal research regarding various topics related to the biology of obesity. One of these findings was explained by Heijmans et al. (33), who reported that adults exposed to famine in utero exhibited persistent DNA methylation in *IGF2* and other metabolism-related genes for more than 60 years. Jaddoe et al. (34) emphasized that maternal BMI prior to pregnancy was associated with offspring's fat mass and the development of cardiometabolic problems at 6–10 years of age. Another study by Sharp et al. (35) found DNA methylation in children of obese mothers. Conversely, Soubry et al. (36) reported altered DNA methylation at the *IGF2* and *MEG3* genes in the offspring of obese fathers. Other researchers, such as Berenson et al. (38), have previously linked childhood obesity to adolescent and adult obesity, and to the development of metabolic syndrome and cardiovascular disease through their lifelong cohort studies (37).

Clinical Implications

The increasing number of studies linking developmental programming, genetics, and epigenetics to obesity is of substantial clinical significance. First, polygenic risk scores and newly discovered

epigenetic biomarkers enable early risk stratification and may permit prenatal diagnosis of high-risk offspring. Integrating these tools into clinical practice could guide personalized counselling and prompt treatments.

Improving maternal health before and during pregnancy is the second important objective. Preconception weight control, dietary support, and reduction of inflammation may reduce the risk of intergenerational transmission of obesity. Furthermore, human data from cohorts of women who underwent bariatric surgery indicate that enhanced maternal metabolism before conception could substantially influence offspring health. Importantly, current research suggests that paternal obesity and pubertal weight status might affect sperm epigenetics and consequent adverse child health outcomes, indicating that fathers' health should not be neglected.

Third, this is more likely to occur during puberty and adolescence. By intervening early, we can reduce the effects of obesity in future generations and prevent its progression into adulthood. This highlights the need to clinically assess pubertal development in obese children and to implement adolescent-centered public health programs.

Finally, new findings in hormones and neuroendocrine systems, such as leptin, ghrelin, neuroestrogen, and FGF19, are ushering in a new era of pharmacological treatments. Future therapies might combine these substances with dietary and lifestyle strategies based on each patient's unique hormonal and epigenetic characteristics. Concurrently, public health frameworks ought to adopt family-centered, multigenerational approaches more consistently, recognizing obesity as a disorder that affects people of all ages and has long-term consequences.

In decision-making, a paradigm shift from late-stage obesity care to early, preventive, and personalized interventions is strengthened by the clinical application of mechanistic information from developmental biology, genetics, and epigenetics.

CONCLUSION

Our knowledge of the genetic, epigenetic, nutritional, and hormonal organization underlying obesity has increased owing to recent developments in the biology of this condition. Numerous loci linked to obesity have been identified through genetic studies, whereas epigenetic studies have focused on the impact of environmental factors on gene expression. A balanced and broad approach to nutrition affects obesity and metabolic health. The newly identified role of neuroestrogens in appetite regulation and energy balance offers new opportunities for clinical studies and drug development. In addition, developing effective prevention and management strategies for obesity requires an understanding of the mechanisms that promote fat accumulation in middle-aged individuals, such as hormonal changes, metabolic alterations, and lifestyle factors. Additionally, the long-term metabolic effects of maternal overnutrition are underscored by the influence of maternal obesity on fetal liver development and on subsequent childhood obesity. Obesity in teenagers may increase the risk of obesity in their offspring and influence their DNA. These complex relationships should be investigated in future studies to develop more effective obesity management and prevention strategies. Future research should emphasize integrating multi-omics data with longitudinal cohort studies to inform the development of targeted preventive guidelines across the lifespan.

Footnotes

Authorship Contributions

Surgical and Medical Practices: K.I.N., Concept: K.I.N., S.S.A., N.H.K., Design: K.I.N., Data Collection or Processing: K.I.N., Analysis or Interpretation: K.I.N., S.S.A., Literature Search: K.I.N., N.H.K., Writing: K.I.N.

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

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

REFERENCES

1. Scully T, Ettela A, LeRoith D, Gallagher EJ. Obesity, type 2 diabetes, and cancer risk. Front Oncol. 2021; 10: 615375.
2. Mahmoud AM. An overview of epigenetics in obesity: The role of lifestyle and therapeutic interventions. Int J Mol Sci. 2022; 23: 1341.
3. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015; 518: 197-206.
4. Huang C, Chen W, Wang X. Studies on the fat mass and obesity-associated (FTO) gene and its impact on obesity-associated diseases. Genes Dis. 2022; 10: 2351-65.
5. Künzel R, Faust H, Bundalian L, Blüher M, Jasaszwili M, Kirstein A, et al. Detecting monogenic obesity: A systematic exome-wide workup of over 500 individuals. Int J Obes (Lond). 2025.
6. Lonky S. Outsmarting obesity: A doctor reveals why we gain weight, why it matters, and what we can do about It. Greenleaf Book Group; 2024. Available from: [URL].
7. Mijač S, Banić I, Genc AM, Lipej M, Turkalj M. The effects of environmental exposure on epigenetic modifications in allergic diseases. Medicina (Kaunas). 2024; 60: 110.
8. Macartney-Coxson D, Benton MC, Blick R, Stubbs RS, Hagan RD, Langston MA. Genome-wide DNA methylation analysis reveals loci that distinguish different types of adipose tissue in obese individuals. Clin Epigenetics. 2017; 9: 48.
9. Wali JA, Solon-Biet SM, Freire T, Brandon AE. Macronutrient determinants of obesity, insulin resistance and metabolic health. Biology (Basel). 2021; 10: 336.
10. Hayashi T, Kumamoto K, Kobayashi T, Hou X, Nagao S, Harada N, et al. Estrogen synthesized in the central nervous system enhances MC4R expression and reduces food intake. FEBS J. 2025; 292: 3900-9.
11. Zangerolamo L, Carvalho M, Solon C, Sidarta-Oliveira D, Soares GM, Marmentini C, et al. Central FGF19 signaling enhances energy homeostasis and adipose tissue thermogenesis through sympathetic activation in obese mice. Am J Physiol Endocrinol Metab. 2025; 328: E524-42.
12. Kawai T, Autieri MV, Scalia R. Adipose tissue inflammation and metabolic dysfunction in obesity. Am J Physiol Cell Physiol. 2021; 320: C375-91.
13. Mirabelli M, Misiti R, Sicilia L, Brunetti FS, Chiefari E, Brunetti A, et al. Hypoxia in human obesity: new insights from inflammation towards insulin resistance-a narrative review. Int J Mol Sci. 2024; 25: 9802.
14. Palmer AK, Jensen MD. Metabolic changes in aging humans: current evidence and therapeutic strategies. J Clin Invest. 2022; 132: e158451.

15. Kodoth V, Scaccia S, Aggarwal B. Adverse changes in body composition during the menopausal transition and relation to cardiovascular risk: A contemporary review. Womens Health Rep (New Rochelle). 2022; 3: 573-81.
16. van der Valk ES, Savas M, van Rossum EFC. Stress and obesity: Are there more susceptible individuals? Curr Obes Rep. 2018; 7: 193-203.
17. Farhana A, Rehman A. Metabolic consequences of weight reduction. [Updated 2023 Jul 10]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025.
18. Wang G, Li G, Song A, Zhao Y, Yu J, Wang Y, et al. Distinct adipose progenitor cells emerging with age drive active adipogenesis. Science. 2025; 388: eadj0430.
19. Wang J, Chang CY, Yang X, Zhou F, Liu J, Feng Z, et al. Leukemia inhibitory factor, a double-edged sword with therapeutic implications in human diseases. Mol Ther. 2023; 31: 331-43.
20. Kim HJ, Kwon O. Nutrition and exercise: cornerstones of health with emphasis on obesity and type 2 diabetes management-a narrative review. Obes Rev. 2024; 25: e13762.
21. Chao AM, Jastreboff AM, White MA, Grilo CM, Sinha R. Stress, cortisol, and other appetite-related hormones: Prospective prediction of 6-month changes in food cravings and weight. Obesity (Silver Spring). 2017 25: 713-20.
22. Amorín R, Liu L, Moriel P, DiLorenzo N, Lancaster PA, Peñagaricano F. Maternal diet induces persistent DNA methylation changes in the muscle of beef calves. Sci Rep. 2023;13:1587. doi: 10.1038/s41598-023-28896-3.
23. Akkila SS, Noel KI, Ibraheem MM. Adipose tissue elastography, anthropometric parameters and non-alcoholic fatty liver disease in obese adults: A cross-sectional study. Al-Rafidain J Med Sci. 2025 Apr 5;8(2):30-4.
24. Harmancıoğlu B, Kabaran S. Maternal high fat diets: impacts on offspring obesity and epigenetic hypothalamic programming. Front Genet. 2023; 14: 1158089.
25. St-Germain LE, Castellana B, Baltayeva J, Beristain AG. Maternal Obesity and the Uterine Immune Cell Landscape: The Shaping Role of Inflammation. Int J Mol Sci. 2020; 21: 3776.
26. Suter MA, Ma J, Vuguin PM, Hartil K, Fiallo A, Harris RA, Charron MJ, Aagaard KM. In utero exposure to a maternal high-fat diet alters the epigenetic histone code in a murine model. Am J Obstet Gynecol. 2014; 210: 463.e1-11.
27. Huang H, Balzer NR, Seep L, Splichalova I, Blank-Stein N, Viola MF, et al. Kupffer cell programming by maternal obesity triggers fatty liver disease. Nature. 2025; 644: 790-8.
28. Huang H, Splichalova I, Balzer N, Splichalova I, Blank-Stein N, Viola MF, et al. Developmental programming of Kupffer cells by maternal obesity causes fatty liver disease in the offspring. 10 August 2023, PREPRINT (Version 1) available at Research Square.
29. Wang YW, Yu HR, Tiao MM, Tain YL, Lin IC, Sheen JM, et al. Maternal obesity related to high fat diet induces placenta remodeling and gut microbiome shaping that are responsible for fetal liver lipid dysmetabolism. Front Nutr. 2021; 8: 736944.
30. Noel KI. A reciprocal relationship between oxidative stress, antioxidants, and cancer: A review. Siriraj Med J. 2024; 76: 550-6.
31. Kitaba NT, Østergaard TM, Lønnebotn M, Accordini S, Real FG, Malinovschi A, et al. Father's adolescent body silhouette is associated with offspring asthma, lung function and BMI through DNA methylation. Commun Biol. 2025; 8: 796.
32. Rajić S, Delerue T, Ronkainen J, Zhang R, Ciantar J, Kostiniuk D, et al. Regulation of nc886 (vtRNA2-1) RNAs is associated with cardiometabolic risk factors and diseases. Clin Epigenetics. 2025; 17: 68.
33. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A. 2008;105: 17046-9.
34. Jaddoe VW, de Jonge LL, Hofman A, Franco OH, Steegers EA, Gaillard R. First trimester fetal growth restriction and cardiovascular risk factors in school age children: Population based cohort study. BMJ. 2014; 348: g14.
35. Sharp GC, Alfano R, Ghantous A, Urquiza J, Rifas-Shiman SL, Page CM, et al. Paternal body mass index and offspring DNA methylation: Findings from the PACE consortium. Int J Epidemiol. 2021; 50: 1297-315.
36. Soubry A, Schildkraut JM, Murtha A, Wang F, Huang Z, Bernal A, et al. Paternal obesity is associated with IGF2 hypomethylation in newborns: Results from a Newborn Epigenetics Study (NEST) cohort. BMC Med. 2013; 11: 29.
37. Al-Rubai A, Ibraheem MM, Hameed AF, Noel KI, Eleiwi SA. Comparison of placental expression of basic fibroblast growth factor and insulin-like growth factor-1 in placentae of normal, pregnancy-induced hypertension, and preeclamptic pregnancies in Iraqi mothers. Medical Journal of Babylon. 2023; 20: 681-8.
38. Berenson GS, Srinivasan SR, Bao W, Newman WP 3rd, Tracy RE, Wattigney WA. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. The Bogalusa Heart Study. N Engl J Med. 1998; 338: 1650-6.

Gazi Med J 2026;37(2):265-270

DOI: http://dx.doi.org/10.12996/gmj.2026.4530

Diagnostic Errors in Clinical Reasoning: A Comprehensive Literature Review on Cognitive Processes, Causes, and Error Reduction Strategies

Klinik Akıl Yürütmede Tanı Hataları: Bilişsel Süreçler, Nedenler ve Hata Azaltma Stratejilerine İlişkin Kapsamlı Bir Literatür İncelemesi

Neşe Pekcan[1], Özlem Coşkun[2]

[1]Education Specialist, Ministry of National Education, Ankara, Türkiye
[2]Department of Medical Education, Gazi University Faculty of Medicine, Ankara, Türkiye

ABSTRACT

Diagnostic decision-making in complex and uncertain clinical settings relies heavily on diagnostic thinking as a core component of clinical reasoning. As healthcare complexity increases, understanding how physicians and medical students reason is essential for ensuring patient safety and quality of care. This review explores the evolution of diagnostic models from early hypothetico-deductive frameworks to contemporary knowledge-based and dual-process perspectives. A narrative synthesis was conducted by drawing on the literature from major biomedical and educational databases, including PubMed, Web of Science, and Scopus. The review focused on conceptual and empirical discussions regarding cognitive processes, diagnostic errors, and error-reduction strategies in clinical practice. Diagnostic errors, which often arise from incomplete data collection or misinterpretation of findings, significantly contribute to patient harm and increased healthcare costs. Notably, interpretation errors are resistant to improvement through clinical experience alone, emphasizing the role of cognitive biases and underdeveloped mental representations (illness scripts). Contextual factors, such as time pressure and physician density, further influence diagnostic accuracy. Reducing diagnostic errors requires targeted educational efforts to enhance pattern recognition, metacognitive awareness, and the systematic use of debiasing strategies. Integrating artificial intelligence as a supportive tool and adopting constructive terminology, such as "missed diagnostic opportunities," may foster a more reflective and safe diagnostic environment.

ÖZ

Karmaşık ve belirsizlik içeren klinik ortamlarda tanısal karar verme süreçleri, klinik akıl yürütmenin temel bileşenlerinden biri olan tanısal düşünmeye büyük ölçüde dayanmaktadır. Sağlık hizmetlerinin giderek daha karmaşık hâle gelmesiyle birlikte, hekimlerin ve tıp öğrencilerinin nasıl akıl yürüttüklerinin anlaşılması, hasta güvenliğinin ve bakım kalitesinin sağlanması açısından kritik önem taşımaktadır. Bu derleme, tanısal modellere ilişkin yaklaşımların erken dönem varsayımsal-tümdengelimli (hipotetiko-dedüktif) çerçevelerden, günümüzde öne çıkan bilgi temelli ve ikili süreç (dual-process) perspektiflerine uzanan gelişimini incelemektedir. Bu çalışma, PubMed, Web of Science ve Scopus başta olmak üzere temel biyomedikal ve eğitim veri tabanlarında yer alan literatüre dayalı olarak gerçekleştirilen anlatısal bir sentez niteliğindedir. Derleme kapsamında, klinik uygulamada bilişsel süreçler, tanı hataları ve hata azaltma stratejilerine ilişkin kavramsal ve ampirik tartışmalara odaklanılmıştır. Çoğu zaman eksik veri toplama ya da bulguların yanlış yorumlanmasından kaynaklanan tanı hataları, hasta zararına ve sağlık hizmeti maliyetlerinin artmasına önemli ölçüde katkıda bulunmaktadır. Özellikle yorumlama hatalarının, yalnızca klinik deneyim yoluyla iyileştirilmesinin zor olduğu görülmektedir. Bu durum, bilişsel yanlılıkların ve yeterince gelişmemiş zihinsel temsillerin (hastalık şemaları/illness scripts) tanısal süreçlerdeki rolünü vurgulamaktadır. Ayrıca zaman baskısı ve hekim yoğunluğu gibi bağlamsal faktörlerin de tanısal doğruluğu etkilediği belirlenmiştir. Tanı hatalarının azaltılması; örüntü tanıma becerilerinin geliştirilmesi, üstbilişsel farkındalığın artırılması ve yanlılık azaltma (debiasing) stratejilerinin sistematik

Cite this article as: Pekcan N, Coşkun Ö. Diagnostic errors in clinical reasoning: a comprehensive literature review on cognitive processes, causes, and error reduction strategies. Gazi Med J. 2026;37(2):265-270

Address for Correspondence/Yazışma Adresi: Neşe Pekcan, Ministry of National Education, Ankara, Türkiye
E-mail / E-posta: npekcannn@gmail.com
ORCID ID: orcid.org/0000-0001-9687-9504

Received/Geliş Tarihi: 14.08.2025
Accepted/Kabul Tarihi: 13.01.2026
Epub: 09.03.2026
Publication Date/Yayınlanma Tarihi: 31.03.2026

ABSTRACT

Keywords: Diagnostic thinking, clinical reasoning, diagnostic error, cognitive bias, metacognition

ÖZ

biçimde kullanılması gibi hedefe yönelik eğitimsel yaklaşımları gerektirmektedir. Ayrıca, yapay zekânın destekleyici bir araç olarak entegrasyonu ve "tanı fırsatının kaçırılması" (missed diagnostic opportunities) gibi daha yapıcı terminolojilerin benimsenmesi, daha yansıtıcı ve güvenli bir tanı ortamının oluşmasına katkı sağlayabilir.

Anahtar Sözcükler: Tanısal düşünme, klinik akıl yürütme, tanısal hata, bilişsel hatalar, metabiliş

INTRODUCTION

The decision-making process in healthcare often involves complex situations characterized by uncertainty and time pressure. Although treatment-related decisions can also be complex, clinicians typically face the highest degree of uncertainty during the diagnostic phase (1). The diagnostic process encompasses a range of cognitive skills, such as analytical thinking, non-analytical pattern recognition, metacognitive monitoring, and the use of knowledge structures. In this respect, diagnostic thinking is positioned as an integral component of clinical reasoning (2). In recent years, the thought processes of physicians and medical students have become a subject of interest. The development of these processes will enable more accurate and efficient healthcare. To understand and foster the development of diagnostic thinking, it is necessary to examine the cognitive processes of physicians and medical students.

Research on medical thinking processes began in the 1980s. One of the earliest approaches in this field was the hypothetico-deductive diagnostic model developed by Elstein et al. (3). Subsequent research moved toward knowledge-based models, focusing on how medical information is structured and retrieved from memory (4-8). The hypothetico-deductive model is an iterative process comprising the stages of cue acquisition, hypothesis generation, data interpretation, and hypothesis testing (3). In contrast, the knowledge-based model focuses on how medical knowledge is organized in memory and how it is accessed; its core components include the recognition of meaningful information, the structuring of clinical knowledge, and the access to stored knowledge frameworks (9).

In summary, as cognitive psychology has advanced, various strategies have been employed to conceptualize medical thinking processes. These processes have been examined in terms of the ways the mind operates, the retrieval of knowledge, pattern recognition, and metacognition. This narrative review aims to synthesize the current understanding of diagnostic thinking processes, examine the cognitive origins of diagnostic errors, and discuss evidence-based strategies to mitigate such errors in clinical practice.

Diagnostic Thinking

Diagnostic thinking is regarded as one of the most critical cognitive skills of physicians, and the enhancement of this process constitutes a primary objective in improving patient safety. Contemporary frameworks for understanding the diagnostic process often employ Dual Process Theory, which categorizes clinical reasoning into two cognitive modes: intuitive (System 1) and analytical (System 2) reasoning. While earlier models focused on hypothetico-deductive methods, this dual-process perspective provides a robust lens for examining how clinicians balance heuristic-based speed with deliberate logic (2). The intuitive approach relies largely on the clinician's experience and is characterized by inductive reasoning. Experienced physicians often recognize patient information within general patterns (gestalt) and make rapid, largely unconscious decisions. These decisions are frequently made under conditions of time pressure, incomplete information, and uncertainty. In intuitive processes, mental shortcuts such as "thin slicing", which rely on instinctive first impressions, are commonly employed. However, these rapid decisions have limitations, as they are often made without access to the full range of data.

In contrast, analytical reasoning is grounded in a more systematic and logical process. When decision-makers have access to more information and resources, they follow a deliberate, step-by-step process to reduce uncertainty. This approach is more closely aligned with normative reasoning and rationality. Most robust clinical decisions are made under conditions in which analytical processes are active. In recent years, these two approaches have been conceptualized within the psychology literature as dual-process theory. System 1 produces rapid, automatic, and context-sensitive decisions, whereas System 2 is engaged in situations requiring slower, deliberate, and logical analysis. Physicians often draw on both systems simultaneously when making diagnoses. Clinical decisions are influenced not only by patient data but also by contextual factors such as environmental conditions, workload, resource limitations, and prior experiences (10,11).

Croskerry (10) argues that System 1 and System 2 are not entirely independent structures but operate along a cognitive continuum, allowing transitions between them. These transitions can influence the accuracy and efficiency of decisions. When conflict arises between the two systems, activating System 2 to review and challenge a System 1 intuitive judgment may provide a safer approach for the patient. Royce et al. (12 note that System 1 relies on pattern recognition and heuristic shortcuts for rapid decision-making, whereas System 2 is more likely to be engaged in the presence of unfamiliar problems or contradictory evidence, operating in a slower and more analytical manner. Dual-process theory suggests that these two systems interact dynamically. While System 1 often generates rapid, intuitive responses, System 2 can be engaged to critically analyze the situation, monitor the outputs of System 1, or take over the reasoning process when high levels of complexity or novelty are encountered (10,13) Norman and Eva (14) contend that when experts deliberately become more analytical, their likelihood of making errors can increase. Their research indicates that training aimed at reducing cognitive biases is often ineffective due to limited

transferability, whereas interventions supporting both intuitive and analytical reasoning can yield small but consistent improvements in accuracy.

Coughlan et al. (15) argue that improving medical decision-making requires physicians to recognize their own thinking patterns and to learn strategies for avoiding cognitive errors. They emphasize that such training should incorporate metacognitive awareness and epistemological understanding. Metacognition involves monitoring and regulating one's own thought processes and consciously engaging in problem-solving, particularly via System 2. Reflection within this process enables decisions to be evaluated more analytically (12).

According to foundational cognitive models, the diagnostic process is structured around several key elements: accurate data collection (gathering high-quality information during history-taking and physical examination), identification of meaningful information (selecting relevant cues), information integration (combining disparate data within a clinical context), and interpretation (evaluating findings against diagnostic hypotheses) (16). Emphasizing the content-specific nature of diagnosis highlights the context sensitivity of diagnostic thinking: recognizing the role of content in identifying errors increases the need for probabilistic or hypothetical reasoning. This aspect parallels the processes of trial-and-error and probabilistic hypothesis generation seen in physiotherapy (17). While these thinking strategies are well-represented in the literature, they warrant further discussion. In particular, the integration of metacognitive concepts into diagnostic thinking processes may enhance awareness among physicians and medical students. Thinking errors arising from these processes are addressed in the subsequent section of this study.

Diagnostic Thinking Errors

Diagnosis plays a central role in patient care, providing meaning to the disease and shaping treatment decisions. The term diagnostic error is used as an umbrella concept, describing a broad spectrum of unintentional errors, including delays in diagnosis, misdiagnoses, and missed diagnoses. However, the terminology used in this field varies considerably among authors (1). Diagnostic errors most often arise from insufficient knowledge, inaccurate data collection, and flawed integration or interpretation of data. While clinical experience can mitigate knowledge gaps and reduce data collection errors, errors in data interpretation do not improve to the same extent. These findings (18-20) suggest that interpretive skill is related not only to the breadth of knowledge but also to the quality of clinical reasoning Moreover, experience alone does not guarantee expertise, as the ability to recognize key clinical features does not necessarily prevent faulty inferences (18).

Research indicates that the majority of diagnostic errors occur during physician–patient encounters, particularly in core stages such as history-taking, physical examination, selection of diagnostic tests, and interpretation of test results. This underscores the multifaceted nature of errors in the diagnostic process. Clinical competence requires the integrated development of effective communication, examination skills, data analysis and synthesis, and reasoning abilities. Nonetheless, even the most experienced physicians occasionally commit diagnostic errors (1).

Although diagnostic errors are difficult to measure, studies have demonstrated their association with suboptimal patient outcomes, clinician stress, and unnecessary healthcare costs (1). Errors in diagnosis are most often linked to clinical reasoning and generally result from inadequate knowledge, poor data collection, or errors in data integration and interpretation. While knowledge deficits and data collection errors may diminish with clinical experience, interpretive errors tend not to improve proportionally. This suggests that diagnostic reasoning is related not only to the amount of knowledge but also to the way in which information is processed. Interpretation skills, in particular, are considered competencies independent of experience (18). For this reason, it is important to approach diagnostic thinking as a holistic process and to remain aware of the principles of psychological functioning; otherwise, incorrect decisions, misdiagnoses, and diminished quality of care are likely to occur (21).

For example, according to the comprehensive analysis by Berner and Graber (22), error rates are very low (less than 5%) in "perceptual" specialties such as radiology and pathology, whereas in high-intensity settings such as emergency departments, the rate rises to 10–15%. In the United States, diagnostic errors are estimated to contribute to approximately 40,000–80,000 hospital deaths annually, with the combined number of deaths and disabilities related to these errors ranging from 80,000 to 160,000 (23). The actual figures may be even higher. While the most serious errors often occur in emergency departments, the overall burden is largely borne by primary care. Certain groups—such as women, ethnic minorities, and younger patients—are at greater risk of diagnostic errors (1). Therefore, these high-risk populations and contexts should be prioritized in both error-reduction strategies and clinical education.

Diagnostic errors can arise not only from individual cognitive processes but also from systemic failures. Examples of process errors include overlooking vital signs, failing to order necessary tests, or delaying communication of laboratory results (24). In addition, mislabeling of symptoms with an incorrect diagnosis (e.g., interpreting pyelonephritis as musculoskeletal pain) or the absence of any diagnosis are examples of diagnostic labeling errors (25). Process and labeling errors can occur independently or together, and pose threats to patient safety. Overuse of diagnostic labels and overdiagnosis are also considered diagnostic errors. For example, in musculoskeletal disorders, pathoanatomic diagnoses are frequently overused, leading to increased imaging, surgical referrals, overtreatment, and patient dissatisfaction (25). Numerous studies have demonstrated that the results of diagnostic tests are not necessarily associated with patient outcomes (27-30).

How Can Diagnostic Errors Be Reduced?

The most fundamental way to prevent diagnostic errors is to increase awareness of one's own thinking processes. This requires individuals to critically examine their own mental functioning and develop metacognitive awareness—the ability to "think about thinking." In diagnostic decision-making, it is particularly important to avoid becoming overly anchored to the initial diagnostic hypothesis. Clinicians should be aware that first impressions can be misleading and that premature closure may lead to diagnostic errors. Therefore, the validity of the initial hypothesis should be continuously questioned, and alternative diagnoses should be actively considered.

Adopting a systematic approach to diagnostic reasoning also contributes to error reduction. Strategies such as hypothetical–deductive reasoning and pattern recognition can support the structured operation of cognitive processes; however, these strategies should be applied flexibly, tailored to the specific context and patient. At this point, cognitive flexibility emerges as a critical skill that enhances diagnostic accuracy. The ability to recognize both automatic and analytical reasoning processes—and consciously transition between them—helps mitigate diagnostic biases.

To reduce the influence of cognitive biases in clinical decision-making, individuals should base their decisions not only on their own knowledge and intuition, but also on objective data, diagnostic test results, and established clinical guidelines. In particular, accurate interpretation of diagnostic test sensitivity, specificity, and predictive values helps avoid false-positive or false-negative conclusions. It has also been noted that patients with the same diagnosis may present with different phenotypic subgroups—a process known as phenotyping (27,31). Variables such as pain sensitivity, psychological status, body mass index, and muscle strength can influence patient outcomes. Recognizing these differences allows for targeted treatments and more accurate prognoses (27).

In undergraduate medical education, it is also essential to train students in ways that reduce the likelihood of future diagnostic errors. Evidence suggests that medical students' diagnostic thinking skills are often underdeveloped and insufficiently addressed in formal education (32,33). Instructional strategies such as problem-based learning, team-based learning, and case-based learning aim to cultivate skills specific to diagnostic thinking, such as processing information rather than memorizing it, formulating hypotheses, and evaluating alternatives (34). Consequently, these approaches should be integrated into the curriculum.

Diagnostic thinking skills of medical students and residents can be monitored using the Diagnostic Thinking Inventory (DTI) (9,32). Personalized learning strategies can then be developed based on whether the learner relies more on analytical or non-analytical reasoning processes. For learners with low DTI scores, interventions aimed at fostering cognitive flexibility can be designed. Exposure to diverse clinical cases can help students develop adaptable thinking skills. High cognitive-demand environments—such as hospital settings—can be strategically utilized to foster the development of diagnostic thinking (34,35). Systematic instruction in critical thinking skills can also improve diagnostic accuracy, while metacognitive training—teaching individuals to think about their thinking—can increase awareness of diagnostic errors (37).

Finally, efforts to improve diagnostic thinking should be supported not only at the individual level but also through teamwork, a feedback culture, and structured educational programs. Collaborative case analysis fosters alternative perspectives and raises awareness of potential biases (38,39). Receiving feedback enables individuals to recognize their own thinking errors and avoid repeating them in the future. In educational settings, an approach should be adopted in which thinking processes are explicitly articulated and inquiry-based learning methods are encouraged.

In summary, avoiding diagnostic thinking errors requires more than technical knowledge alone; it necessitates understanding the underlying cognitive processes, continuously reviewing these processes, and fostering supportive learning environments. This represents a fundamental mental discipline that strengthens clinical safety and decision-making accuracy over the long term.

DISCUSSION

In medical settings, decision-makers frequently operate under less-than-ideal conditions such as time pressure, distraction, fatigue, sleep deprivation, and resource constraints. Clinical decisions are often made rapidly and intuitively under the influence of factors such as cognitive load, environmental stimuli, and limited resources. This environment compels healthcare providers to develop strategies for maintaining patient care efficiency amidst a dynamic and unpredictable workload (10).

The diagnostic process is not limited to interpreting test results; it encompasses multidimensional tasks such as analytical reasoning, contextual evaluation, and consideration of phenotypic diversity. This approach requires linking the diagnostic process not only to measures such as accuracy or test sensitivity but also to the patient's overall management. For this reason, higher-order thinking skills are critical in clinical decision-making. Moving beyond rote-based approaches, clinicians need structured thinking models for differential diagnosis and patient-centered care (14,27).

Research on malpractice and diagnostic errors highlights the importance of advanced clinical reasoning skills and the necessity of educational strategies focused on critical thinking (12,40). Recommended educational interventions aim to develop metacognitive strategies, increase awareness of cognitive biases, and provide techniques to mitigate these biases (12,14). Furthermore, identifying phases of the diagnostic process where errors are most concentrated and creating targeted educational content for these phases have the potential to reduce error rates.

An analysis of the causes of diagnostic errors reveals that the diagnostic process is shaped not only by individual knowledge but also by contextual factors such as early specialization (which may narrow diagnostic focus) and environmental differences (36). The quality of clinical experience is thus inextricably linked to the context in which it occurs. A critical contextual factor is time pressure, which often stems from high patient volumes. For example, although the number of physicians per 1,000 inhabitants in Turkey has recently increased to 2.2 (41), this figure remains below the OECD average, which may lead to increased workloads and shorter consultation times. Indeed, Monteiro et al. (42) demonstrated that reduced thinking time significantly lowers diagnostic accuracy, particularly when clinicians cannot engage in reflective reasoning. Consequently, an educational process that provides sufficient time and experiential learning opportunities is essential to facilitate more accurate diagnoses (43).

In the clinical diagnostic process, intuitive thinking and cognitive biases significantly affect decision quality. While intuitive approaches can be functional for urgent decisions, they often lead to systematic reasoning errors. Cognitive biases, defined as deviations from rational judgment, are frequently associated with these heuristic shortcuts (15). However, educational strategies that solely focus on identifying biases have proven ineffective to date. This underscores the need to integrate memory-based strategies into training programs (12).

To develop effective mental representations in the diagnostic process, it is important to deliberately expose students to rare or atypical cases. Natural clinical distributions are insufficient for building such representations. Therefore, simulation-based learning techniques and principles of deliberate practice should be integrated early in medical training.

Modern medical education is shifting away from purely knowledge-transmission approaches toward models that emphasize problem-solving and critical thinking. However, for this transformation to be sustained and effective, curricula must include structures that promote analytical thinking and provide strategies for metacognitive awareness. Prioritizing the interpretation and integration of clinical knowledge will support the development of graduates' reasoning skills.

The diagnostic thinking process transcends simple hypothesis generation and testing; it fundamentally relies on knowledge structure, pattern recognition, and rapid information retrieval. While pattern recognition is a hallmark of clinical expertise, over-reliance on this intuitive process without analytical verification can lead to cognitive pitfalls such as premature closure (45). Furthermore, weak, fragmented, or incomplete mental representations—often referred to as underdeveloped illness scripts—render clinicians more susceptible to biases (46). Consequently, medical education must go beyond factual knowledge to systematically teach metacognitive skills and debiasing strategies to safeguard the diagnostic process.

Artificial intelligence (AI)-supported systems hold significant potential for reducing diagnostic errors. However, AI should be positioned as an advisory tool rather than as a replacement for physicians (15,44). AI remains limited in performing cognitive processes, such as analyzing the temporal evolution of symptoms, evaluating contextual information, and prioritizing differential diagnoses. Moreover, overreliance on AI could hinder the development of critical thinking skills and reduce transparency in the diagnostic process. The most effective approach is for the human physician to make the final decision, with AI serving to support the process.

Evaluating the level of diagnostic thinking could be considered a quality management and accreditation metric in healthcare services (45). Assessment activities conducted throughout physicians' training play a critical role in identifying reasoning errors. Additionally, raising awareness of the reliability of the literature and facilitating access to reliable information can help reduce diagnostic errors.

Finally, Royce et al. (12) argue that replacing terms such as "diagnostic error" with more constructive and flexible expressions (e.g., "missed diagnostic opportunity"), or replacing "differential diagnosis" with "diagnostic hypothesis," may make errors easier to acknowledge and discuss. Reframing the language used in diagnostic reasoning could, in turn, contribute to transforming these processes.

CONCLUSION

Diagnostic errors remain a significant threat to patient safety and healthcare quality worldwide. As demonstrated in the literature, these errors arise not only from knowledge deficits but primarily from flaws in cognitive processing, contextual pressures, and underdeveloped illness scripts. Clinical experience alone does not guarantee improved diagnostic accuracy, particularly in relation to interpretive errors. Therefore, structured educational interventions that promote metacognitive awareness, cognitive flexibility, and deliberate practice are essential. In addition, system-level strategies, collaborative diagnostic practices, and supportive integration of artificial intelligence may further enhance diagnostic reliability. Ultimately, improving diagnostic thinking requires a comprehensive approach that integrates cognitive science, medical education, and healthcare system reform.

Footnotes

Authorship Contributions

Concept: N.P., Ö.C., Design: N.P., Ö.C., Data Collection or Processing: N.P., Ö.C., Analysis or Interpretation: N.P., Ö.C., Literature Search: N.P., Ö.C., Writing: N.P., Ö.C.

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

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

REFERENCES

1. Benishek LE, Weaver SJ, Newman Toker DE. The cognitive psychology of diagnostic errors. Sci Am Neurol. 2015: 1-21.
2. Croskerry P. A universal model of diagnostic reasoning. Acad Med. 2009; 84: 1022-8.
3. Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). Medical problem solving: An analysis of clinical reasoning. Cambridge, MA: Harvard University Press.
4. Gale J, Marsden P. Medical diagnosis: From student to clinician. Oxford: Oxford University Press; 1983.
5. Bordage G, Zacks R. The structure of medical knowledge in the memories of medical students and general practitioners: categories and prototypes. Med Educ. 1984; 18: 406-16.
6. Bordage G, Lemieux M. Semantic structures and diagnostic thinking of experts and novices. Acad Med. 1991; 66: S70-2.
7. Patel VL, Groen GJ. Knowledge based solution strategies in medical reasoning. Cogn Sci. 1986; 10: 91-116.
8. Norman GR. Problem-solving skills, solving problems and problem-based learning. Med Educ. 1988; 22: 279-86.
9. Bordage G, Grant J, Marsden P. Quantitative assessment of diagnostic ability. Med Educ. 1990; 24: 413-25.
10. Croskerry P. A universal model of diagnostic reasoning. Acad Med. 2009; 84: 1022-8.
11. Graber ML, Rusz D, Jones ML, Farm-Franks D, Jones B, Cyr Gluck J, et al. The new diagnostic team. Diagnosis (Berl). 2017; 4: 225-38.
12. Royce CS, Hayes MM, Schwartzstein RM. Teaching critical thinking: A case for instruction in cognitive biases to reduce diagnostic errors and improve Patient Safety. Acad Med. 2019; 94: 187-94.
13. Evans JSB. Dual-processing accounts of reasoning, judgment, and social cognition. Annu Rev Psychol. 2008; 59: 255-78.
14. Norman GR, Eva KW. Diagnostic error and clinical reasoning. Med Educ. 2010; 44: 94-100.
15. Coughlan JJ, Mullins CF, Kiernan TJ. Diagnosing, fast and slow. Postgrad Med J. 2021; 97: 103-109.
16. Bordage G. Elaborated knowledge: a key to successful diagnostic thinking. Acad Med. 1994; 69: 883-5.
17. Leuders T, Loibl K, Sommerhoff D, Herppich S, Praetorius AK. Toward an overarching framework for systematizing research perspectives

on diagnostic thinking and practice. Educ Psychol Rev. 2022; 34: 591-611.

18. Groves M, O'Rourke P, Alexander H. Clinical reasoning: the relative contribution of identification, interpretation and hypothesis errors to misdiagnosis. Med Teach. 2003; 25: 621-5.
19. Eva KW. What every teacher needs to know about clinical reasoning. Med Educ. 2005; 39: 98-106.
20. Schmidt HG, Rikers RM. How expertise develops in medicine: Knowledge encapsulation and illness script formation. Med Educ. 2007; 41: 1133-9.
21. Ivanovich C, Fedorovich AA. The importance of researching general components of diagnostic thinking in modern psychology. J Psychol Clin Psychiatry. 2018; 9: 270-1.
22. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. Am J Med. 2008; 121: S2-23.
23. Newman-Toker DE, McDonald KM, Schaffer AC. Charting the course: Improving the diagnosis of healthcare quality. J Gen Intern Med. 2014; 29: 1579-81.
24. National Academies of Sciences, Engineering, and Medicine. Improving diagnosis in health care. Balogh EP, Miller BT, Ball JR, editors. Washington (DC): National Academies Press (US); 2015. doi: 10.17226/21794. Available from: https://nap.nationalacademies.org/catalog/21794/improving-diagnosis-in-health-care
25. Zwaan L, Thijs A, Wagner C, van der Wal G, Timmermans DR. Design of a retrospective patient record review study of diagnostic error in the internal medicine setting. BMC Health Serv Res. 2010;10:331. doi:10.1186/1472-6963-10-331.
26. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. Lancet. 2018; 391: 2368-83.
27. Cook CE, Décary S. Higher order thinking about differential diagnosis. Braz J Phys Ther. 2020; 24: 1-7.
28. Chou R, Fu R, Carrino JA, Deyo RA. Imaging strategies for low-back pain: Systematic review and meta-analysis. Lancet. 2009; 373: 463-72.
29. Jarvik JG, Hollingworth W, Martin B, Emerson SS, Gray DT, Overman S, et al. Rapid magnetic resonance imaging vs. radiographs for patients with low back pain: a randomized controlled trial. JAMA. 2003; 289: 2810-8.
30. Lurie JD. Review: routine spinal imaging for low back pain does not improve outcomes. Evid Based Med. 2003;8(6):191. doi:10.1136/ebm.8.6.191.
31. O'Sullivan PB, Caneiro JP, Kiernan M, Lin I. Unraveling the complexity of low back pain. J Orthop Sports Phys Ther. 2016; 46: 932-7.
32. Picho K, Maggio LA. The diagnostic thinking inventory: A systematic review of its development and use in medical education. Acad Med. 2020;95(12):1930-9. doi:10.1097/ACM.0000000000003452.
33. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication. Acad Med. 1990; 65: 611-21.
34. Owlia F, Keshmiri F, Kazemipoor M, Rashidi Maybodi F. Assessment of clinical reasoning and diagnostic thinking among dental students. Int J Dent. 2022;2022:1085326. Available from: https://www.hindawi.com/journals/ijd/2022/1085326/
35. Durning SJ, Costanzo ME, Beckman TJ, Artino AR Jr, Roy MJ, Cleary TJ, et al. Functional neuroimaging correlates of thinking flexibility and knowledge structure in memory: exploring the relationships between clinical reasoning and diagnostic thinking. Med Teach. 2016;38(6):570-7. doi:10.3109/0142159X.2015.1047755.
36. Grageda ME, Rotor E. Diagnostic thinking skills of Filipino physical therapists. Physiotherapy. 2015;101(Suppl 1):eS480. doi:10.1016/j.physio.2015.03.3560.
37. Lunney M. Use of critical thinking in the diagnostic process. Int J Nurs Terminol Classif. 2010; 21: 82-8.
38. Graber ML, Kissam S, Payne VL, et al. Cognitive interventions to reduce diagnostic error: A narrative review. BMJ Qual Saf. 2012; 21: 535-45.
39. Mamede S, Schmidt HG, Penaforte JC. Effect of vicarious experience, group discussion, and reflection on students' diagnostic accuracy. Med Educ. 2014; 48: 259-73.
40. Saber Tehrani AS, Lee H, Mathews SC, Shore A, Makary MA, Pronovost PJ, et al. 25-Year summary of US malpractice claims for diagnostic errors 1986-2010: an analysis from the National Practitioner Data Bank. BMJ Qual Saf. 2013; 22: 672-80.
41. OECD. Health at a glance 2023: OECD indicators. Paris: OECD Publishing; 2023. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2023_7a7afb35-en
42. Monteiro SD, Sherbino J, Patel A, Mazzetti I, Norman GR, Howey E. Reflecting on diagnostic errors: taking a second look is not enough. J Gen Intern Med. 2015; 30: 1270-4.
43. Monteiro SM, Norman G. Diagnostic reasoning: Where we've been, where we're going. Teach Learn Med. 2013;25(Suppl 1):S26-32. doi: 10.1080/10401334.2013.842911.
44. Harada T, Shimizu T, Kaji Y, Suyama Y, Matsumoto T, Kosaka C, et al. A Perspective from a case conference on comparing the diagnostic process: human diagnostic thinking vs. artificial intelligence (AI) decision support tools. Int J Environ Res Public Health. 2020; 17: 6110.
45. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med. 2003; 78: 775-80.