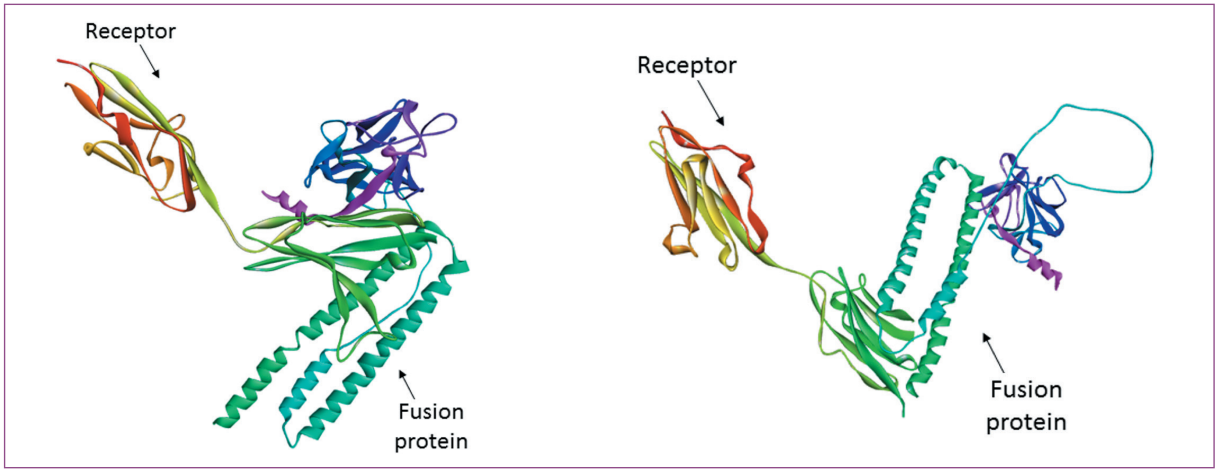
ISSN: 2822-4302
eISSN: 2587-0602

Medical Journal of Western Black Sea

Batı Karadeniz Tıp Dergisi

Cilt / Volume 10
Sayı / Number 1
Nisan / April
2026



In-silico engineering of an antimicrobial growth factor fusion protein to prevent post-laser infection in hemorrhoidal wounds

- Investigation of factors associated with the level of temporomandibular dysfunction in university students
- In-silico engineering of an antimicrobial growth factor fusion protein to prevent post-laser infection in hemorrhoidal wounds
- The correlation between serum YKL-40, proteinuria and inflammation in patients with systemic lupus erythematosus
- Predictors of early readmission in patients awaiting interval cholecystectomy after conservative treatment of acute cholecystitis: A retrospective cohort study
- Investigation of the effects of odontoid bone anatomical location and structure on cervical lordosis angle in cervical CT examinations: A clinical study
- Clinical and radiological response assessment after whole brain radiotherapy for patients with non-small cell lung cancer with brain metastases
- Evaluation of viral respiratory pathogens detected in a tertiary university hospital between 2021 and 2025
- Comparative analysis of pH measurements: Urine dipstick and pH meter
- Applications and research trends of large language models in healthcare: A bibliometric analysis
- Clinical characteristics and hematological inflammatory markers in pediatric peripheral facial paralysis: A retrospective study
- Evaluation of practical application of Tell-Hashomer, Livneh and Yalçınkaya-Özen criteria for diagnosis of Familial Mediterranean Fever in children
- The effect of a surgical nursing course structured with role-playing and flipped learning methods on students' motivation and attitudes toward clinical practice
- Postoperative frequency of normocalcemic elevated parathormone in patients surgically treated for primary hyperparathyroidism: A single-center experience
- Adneksiyal kitlesi olan kadınlarda siberkondri düzeyini etkileyen faktörler: Tek merkezli kesitsel çalışma
- Sensory perception profiles in children with epilepsy: A prospective controlled study
- Effect of intraoperative magnesium sulfate infusion on early postoperative opioid consumption in laparoscopic hysterectomy: A randomized, double-blind, controlled trial
- Effects of preemptive analgesia and preoperative anxiety on emergence delirium after pediatric adenotonsillectomy
- Relationship between demodex infestation and clinical parameters in polycystic ovary syndrome: A prospective observational study
- Clinical and radiological outcome comparison of unilateral vs. bilateral percutaneous vertebroplasty for osteoporotic vertebral compression fractures
- Hemichorea-hemiballismus developing due to non-ketotic hyperglycemia: A rare case report
- Artificial intelligence in pain: A comprehensive review

https://dergipark.org.tr/tr/pub/mjwbs
https://dergipark.org.tr/en/pub/mjwbs

Medical Journal of Western Black Sea

Batı Karadeniz Tıp Dergisi

Zonguldak Bülent Ecevit Üniversitesi Yayın Organıdır.
Official Journal of Zonguldak Bulent Ecevit University.
Yılda üç kez yayınlanır / Published three times per year

ISSN: 2822-4302
eISSN: 2587-0602

Cilt / Volume 10 Sayı / Number 1 Nisan / April 2026

Zonguldak Bülent Ecevit Üniversitesi Adına İmtiyaz Sahibi
Owner on behalf of Zonguldak Bülent Ecevit University
Rektör/Rector Prof. Dr. İsmail Hakkı ÖZÖLÇER

Yönetim Yeri / *Head Office*
Zonguldak Bülent Ecevit University Faculty of Medicine, 67600, Zonguldak, Türkiye

Baş Editör / *Editor-in-Chief*
Doç. Dr. Emrah KESKİN
Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Neurosurgery, Zonguldak, Türkiye
emrah.keskin@beun.edu.tr

Temel Tıp Bilimleri Bölüm Editörleri / *Basic Medical Sciences Section Editors*

Zehra SAFİ ÖZ, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Medical Biology*
Meryem AKPOLAT FERAH, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Histology and Embryology*
Mustafa Çağatay BÜYÜKUYSAL, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Biostatistics*
Salih ERDEM, *Zonguldak Bülent Ecevit University, Ahmet Erdoğan Health Vocational School, Pathology Laboratory Techniques*
Ali Vasfi AĞLARCI, *Kastamonu University, Faculty of Medicine, Department of Biostatistics, Basic Medical Sciences*
Gamze ERGÜLER TANRIÖVER, *Akdeniz University, Faculty of Medicine, Department of Histology and Embryology*

Dahili Tıp Bilimleri Bölüm Editörleri / *Internal Medical Sciences Section Editors*

Özlem ELMAS, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Radiation Oncology*
Esra ACIMAN DEMİREL, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Neurology*
Uğur KÖKTÜRK, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Cardiology*
Halil İbrahim DÜŞÜNCELİ, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Internal Medicine*
Tuğçe KÖKSAL, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Physical Medicine and Rehabilitation*
Emine FİDAN, *Zonguldak Bülent Ecevit University, Faculty of Pharmacy*

Cerrahi Tıp Bilimleri Bölüm Editörleri / *Surgical Medical Sciences Section Editors*

Güldeniz KARADENİZ ÇAKMAK, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of General Surgery*
Mehmet Erhan TÜRKOĞLU, *Hacettepe University, Faculty of Medicine, Department of Neurosurgery*
Deniz BAKLACI, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Otolaryngology (Ear, Nose, and Throat)*
Anıl TURHAN ÇAKIR, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Obstetrics and Gynecology*
Oğuz ARSLANTÜRK, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Cardiovascular Surgery*
Bengü Gülhan KÖKSAL İNCEGÜL, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Anesthesiology and Reanimation*
Şahin HANALİOĞLU, *Hacettepe University, Faculty of Medicine, Department of Neurosurgery*

Uluslararası Editörler / *International Editors*

Ammad Ahmad Farooqi, *Principal Scientist at Institute of Biomedical and Genetic Engineering, Islamabad, Pakistan*
Senta Frol, *Department of Vascular Neurology, University Medical Centre Ljubljana,* Ljubljana, Slovenia
Sabino Luzzi, *Department of Neurosurgery, University of Sassari, Sassari, Italy*

Medical Journal of Western Black Sea

Batı Karadeniz Tıp Dergisi

Zonguldak Bülent Ecevit Üniversitesi Yayın Organıdır.
Official Journal of Zonguldak Bulent Ecevit University.
Yılda üç kez yayınlanır / Published three times per year

ISSN: 2822-4302
eISSN: 2587-0602

Cilt / Volume 10 Sayı / Number 1 Nisan / April 2026

Yardımcı Editörler / *Associate Editors*

Salih ERDEM, *Zonguldak Bülent Ecevit University, Ahmet Erdoğan Health Vocational School, Pathology Laboratory Techniques*
Bengü Gülhan KÖKSAL İNCEGÜL, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Anesthesiology and Reanimation*
Gamze ERGÜLER TANRIÖVER, *Akdeniz University, Faculty of Medicine, Department of Histology and Embryology*

Biyoistatistik Danışmanları / *Consultants in Statistics*

Çağatay BÜYÜKUYSAL, *Zonguldak Bülent Ecevit University, Faculty of Medicine, Department of Biostatistics*
Ali Vasfi AĞLARCI, *Kastamonu University, Faculty of Medicine, Department of Biostatistics, Basic Medical Sciences*
Mustafa Agah TEKİNDAL, *Izmir Katip Celebi University, Faculty of Medicine, Department of Biostatistics*
Harun YONAR, *Konya Selçuk University, Faculty of Veterinary, Department of Biostatistics*

Teknik Destek / *Technical Support*

Salih ERDEM, *Zonguldak Bülent Ecevit University, Ahmet Erdoğan Health Vocational School, Pathology Laboratory Techniques*

Türkçe Redaksiyon Kurulu / *Turkish Redaction Board*

Gül Banu DUMAN, *Zonguldak Bülent Ecevit University, Faculty of Arts and Sciences, Department of Turkish Language and Literature*
Betül MUTLU, *Zonguldak Bülent Ecevit University, Faculty of Arts and Sciences, Department of Turkish Language and Literature*
Muhittin TURAN, *Zonguldak Bülent Ecevit University, Faculty of Arts and Sciences, Department of Turkish Language and Literature*
Hasan ÖZER, *Zonguldak Bülent Ecevit University, Faculty of Arts and Sciences, Department of Turkish Language and Literature*
Osman DARICAN, *Zonguldak Bülent Ecevit University, Faculty of Arts and Sciences, Department of Turkish Language and Literature*
Serdar Deniz ÖZDEMİR, *Zonguldak Bülent Ecevit University, Faculty of Arts and Sciences, Department of Turkish Language and Literature*
Nuray KARAKAYA, *Zonguldak Bülent Ecevit University, Faculty of Arts and Sciences, Department of Turkish Language and Literature*

İngilizce Redaksiyon Kurulu / *English Redaction Board*

Okşan DAĞLI, *Coordinator, Zonguldak Bülent Ecevit University, School of Foreign Languages, Department of Foreign Languages*
İnan TEKİN, *Zonguldak Bülent Ecevit University, School of Foreign Languages, Department of Foreign Languages*
Sevilay YILDIRIM, *Zonguldak Bülent Ecevit University, School of Foreign Languages, Department of Foreign Languages*
Yunus Emre GÜNER, *Zonguldak Bülent Ecevit University, School of Foreign Languages, Department of Foreign Languages*

Danışma Kurulu / *Advisory Board*

Hakan AKBULUT, *Ankara University, Faculty of Medicine, Department of Medical Oncology*
Volkan HANCI, *Dokuz Eylül University, Faculty of Medicine, Department of Anesthesiology and Reanimation*
K. Gonca AKBULUT, *Gazi University, Faculty of Medicine, Department of Physiology*
Tahir Kansu BOZKURT, *Greatwestwern Hospital, United Kingdom*
Ahmet GÜL, *Istanbul University, Istanbul Faculty of Medicine, Department of Rheumatology*
Derya KARADENİZ, *Istanbul University, Cerrahpaşa Faculty of Medicine, Department of Neurology*
Güler ÖZTÜRK, *Medeniyet University, Faculty of Medicine, Department of Physiology*
Atakan SEZER, *Trakya University, Faculty of Medicine, Department of General Surgery*
Bülent Sabri CIĞALI, *Trakya University, Faculty of Medicine, Department of Anatomy*
Dikmen DÖKMECİ, *Trakya University, Faculty of Medicine, Department of Pharmacology*
Gülnur KIZILAY, *Trakya University, Faculty of Medicine, Department of Histology and Embryology*
Nurettin AYDOĞDU, *İnönü University, Faculty of Medicine, Department of Physiology*

Medical Journal of Western Black Sea

Batı Karadeniz Tıp Dergisi

Zonguldak Bülent Ecevit Üniversitesi Yayın Organıdır.
Official Journal of Zonguldak Bulent Ecevit University.
Yılda üç kez yayınlanır / Published three times per year

ISSN: 2822-4302
eISSN: 2587-0602

Cilt / Volume 10 Sayı / Number 1 Nisan / April 2026

Yayıncı / *Publisher*
Zonguldak Bülent Ecevit Üniversitesi Yayın Organıdır
Official Journal of Zonguldak Bülent Ecevit University

Yayıncı Adresi / *Publisher Address*
Zonguldak Bülent Ecevit Üniversitesi
İncivez, Üniversite Cd., 67100 Zonguldak Merkez/Zonguldak

Yayın Türü / *Publication Type*
Uluslararası hakemli dergi / *International peer-reviewed journal*

Yayın Sıklığı ve Dili / *Publication Frequency and Language*
Yılda üç sayı (Nisan, Ağustos, Aralık), Türkçe ve İngilizce
Triannual (April, August, December), Turkish and English

Basım tarihi / *Printing date: 30.04.2026*
Asitsiz kağıda basılmıştır / *Printed on acid-free paper*

Kapak Görseli / *Cover Picture:* *Karaca Ocak Ö*, Sayfa/*Page* 9

Yayıncılık Hizmetleri / *Publishing Services*

Buluş Tasarım ve Matbaacılık Hizmetleri San. Tic.
Mebusevleri Mah., Turgut Reis Cad., No: 11/1, Çankaya, Ankara
Tel: 0312 222 44 06 www.bulustasarim.com.tr

Baskı / *Printed at*

Vadi Grafik Tasarım ve Reklamcılık Ltd. Şti.
İvedik Organize San. 1420. Cadde No: 58/1, Ostim-Y. Mahalle, Ankara,
Tel: +90 (312) 395 85 71 72 Faks: 0 (312) 395 85 72 E-mail: info@stepdijital.com

Bu dergide kullanılan kağıt ISO 9706: 1994 standardına ("Requirements for Permanence") uygundur.
The paper used to print this journal conforms to ISO 9706: 1994 standard (Requirements for Permanence).

Dizinlenme / *Abstracting and Indexing*

Batı Karadeniz Tıp Dergisi, TR Dizin ve Türkiye Atıf Dizini tarafından indekslenmektedir.
Medical Journal of Western Black Sea is indexed in TR Index and Turkiye Citation Index.

Tüm yayın politikalarına ve yazar rehberine https://dergipark.org.tr/tr/pub/mjwbs adresinden ulaşabilirsiniz.
You can reach all publication policies and author guidelines from https://dergipark.org.tr/en/pub/mjwbs.

22
PAP

ÇEVRE BİLGİSİ / *ENVIRONMENTAL INFORMATION*

Bu dergide kullanılan kağıdın üreticisi olan şirket ISO 14001 çevre yönetim sertifikasına sahiptir. Üretici şirket tüm odun elyafını sürdürülebilir şekilde temin etmektedir. Şirketin ormanları ve plantasyonları sertifikalıdır. Üretimde kullanılan su arıtılarak dönüşümlü kullanılmaktadır. Bu derginin basımında ağır metaller ve film kullanılmamaktadır. Alüminyum basım kalıplarının banyo edilmesinde kullanılan sıvılar arıtılmaktadır. Kalıplar geri dönüştürülmektedir. Basımda kullanılan mürekkepler zehirli ağır metaller içermemektedir.

Bu dergi geri dönüştürülebilir, imha etmek istediğinizde lütfen geri dönüşüm kutularına atınız.

The company that manufactures the paper used in this journal has an ISO 14001 environmental management certificate. The company obtains all wood fiber in a sustainable manner. The forests and plantations of the company are certified. The water used in production is purified and used after recovery. Heavy metals or film are not used for the publication of this journal. The fluids used for developing the aluminum printing templates are purified. The templates are recycled. The inks used for printing do not contain toxic heavy metals.

This journal can be recycled. Please dispose of it in recycling containers.

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi
Cilt / Volume 10 Sayı / Number 1 Nisan / April 2026

EDİTÖRDEN

Değerli Okuyucularımız,

Batı Karadeniz Tıp Dergisi olarak 2026 yılının ilk sayısını sizlerle buluşturmaktan memnuniyet duyuyoruz. Kuruluşumuzdan bu yana bilimsel üretimi destekleyen, etik yayıncılığı esas alan ve klinik uygulamalara katkı sunan nitelikli çalışmaları yayımlama anlayışımızı kararlılıkla sürdürmekteyiz. Her yeni sayımızda, farklı tıp disiplinlerinden özgün araştırmaları bir araya getirerek bilimsel bilgi birikimine katkı sunmaya devam ediyoruz.

Geride bıraktığımız süreçte dergimizin akademik görünürlüğünü artırmaya yönelik önemli adımlar atılmıştır. TÜBİTAK ULAKBİM TR Dizin ve Türkiye Atıf Dizini başta olmak üzere mevcut indekslerdeki yerimizi güçlendirirken, uluslararası prestijli veri tabanlarında yer alma hedefimiz doğrultusunda gerekli başvurular tamamlanmış ve değerlendirme süreçleri başlatılmıştır. Bu bağlamda, dergimizin uluslararası bilim dünyasında daha güçlü bir konuma ulaşması adına hedeflerimize emin adımlarla yürümekteyiz.

Nisan 2026 sayımızda yer alan çalışmalar; bilimsel niteliği yüksek, metodolojik açıdan güçlü ve günlük klinik pratiğe katkı sağlayacak içeriklerden oluşmaktadır. Dergimizin yalnızca bir yayın platformu değil, aynı zamanda akademik etkileşimi güçlendiren, bilgi paylaşımını teşvik eden ve bilimsel iş birliğini destekleyen bir platform olma misyonunu sürdürdüğüne inanıyoruz.

Bu vesileyle, dergimize değerli çalışmalarıyla katkı sunan tüm yazarlarımıza, titiz değerlendirmeleriyle yayın kalitemizi yükselten hakemlerimize, Danışma Kurulumuza, Alan Editörlerimize, Editör Yardımcılarımıza, Türkçe ve İngilizce Dil Redaksiyon Kurullarımıza, Yayın Kurulumuza ve teknik ekibimize içten teşekkürlerimi sunarım. Ayrıca kurumsal destekleri için Rektörümüz Prof. Dr. İsmail Hakkı ÖZÖLÇER'e şükranlarımı ifade ederim.

Batı Karadeniz Tıp Dergisi'nin 2026 yılı Nisan sayısında yer alan çalışmaların mesleki gelişiminize katkı sağlamasını, yeni araştırmalara ilham vermesini ve bilimsel üretime değer katmasını temenni ederim.

Saygılarımla,

Doç. Dr. Emrah KESKİN
Baş Editör
Nisan 2026

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi
Cilt / Volume 10 Sayı / Number 1 Nisan / April 2026

EDITORIAL

Dear Readers,

We are pleased to present the first issue of 2026 of the Medical Journal of Western Black Sea. Since our establishment, we have remained committed to publishing high-quality studies that support scientific productivity, uphold ethical publishing principles, and contribute to clinical practice. With each new issue, we continue to bring together original research from various medical disciplines and contribute to the advancement of scientific knowledge.

During the recent period, significant steps have been taken to further enhance the academic visibility of our journal. While strengthening our presence in the databases in which we are currently indexed, including TÜBİTAK ULAKBİM TR Dizin and the Turkish Citation Index, the necessary applications for prestigious international indexing databases have been completed and evaluation processes have been initiated. In this context, we continue to move forward with determination toward our goals of achieving a stronger position within the global scientific community.

The studies included in our April 2026 issue consist of scientifically valuable works with strong methodology and the potential to contribute directly to daily clinical practice. We believe that our journal continues to fulfill its mission not only as a publication platform, but also as a medium that promotes academic interaction, encourages knowledge sharing, and supports scientific collaboration.

On this occasion, I would like to sincerely thank all our authors for their valuable contributions, our reviewers for their meticulous evaluations that enhance the quality of our publication, our Advisory Board, Section Editors, Associate Editors, Turkish and English Language Editing Boards, Editorial Board members, and our technical team for their dedicated efforts. I would also like to express my gratitude to our Rector, Prof. Dr. İsmail Hakkı ÖZÖLÇER, for his institutional support.

I hope that the articles presented in the April 2026 issue of the Medical Journal of Western Black Sea will contribute to your professional development, inspire new research, and add value to scientific production.

Sincerely,

Emrah Keskin, MD, Assoc. Prof.
Editor-in-Chief
April 2026

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi
Cilt / Volume 10 Sayı / Number 1 Nisan / April 2026

İÇİNDEKİLER

Özgün Araştırmalar / Original Researchs

1 **Investigation of factors associated with the level of temporomandibular dysfunction in university students**
Üniversite öğrencilerinde temporomandibular disfonksiyon düzeyi ile ilişkili faktörlerin incelenmesi
Merve Bulguroğlu, Halil İbrahim Bulguroğlu, Gökçe Kılıç, Selcan Suiçmez, Serenay Zorlu, Cansu Gevrek Aslan

9 **In-silico engineering of an antimicrobial growth factor fusion protein to prevent post-laser infection in hemorrhoidal wounds**
Hemoroidal yaralarda lazer sonrası enfeksiyonu önlemek için antimikrobiyal büyüme faktörü füzyon proteinlerinin bilgisayar simülasyonlu tasarımı
Özlem Karaca Ocak, Zoha Naeem

25 **The correlation between serum YKL-40, proteinuria and inflammation in patients with systemic lupus erythematosus**
Sistemik lupus eritematozuslu hastalarda serum YKL-40, proteinüri ve inflamasyon düzeyi arasındaki ilişki
İlyas Öztürk, Fatma Betül Güzel, Serhat Kura, Enes Süleyman Gün, Burak Okyar, Sinan Kazan, Muhammed Seyithanoğlu, Gözde Yıldırım Çetin, Orçun Altunören, Özkan Güngör

31 **Predictors of early readmission in patients awaiting interval cholecystectomy after conservative treatment of acute cholecystitis: A retrospective cohort study**
Akut kolesistitte konservatif tedavi sonrası interval kolesistektomi bekleyen hastalarda erken yeniden yatışın prediktörleri: Retrospektif kohort çalışması
Kazım Duman

39 **Investigation of the effects of odontoid bone anatomical location and structure on cervical lordosis angle in cervical CT examinations: A clinical study**
Servikal BT incelemelerinde odontoid kemik anatomik yerleşiminin ve yapısının servikal lordoz açısı üzerine olan etkilerinin araştırılması: Bir klinik araştırma
Alemiddin Özdemir, Ahmet Melih Erdoğan, Özge Sevimoğlu, Buket Oğuz, Buse Kaymakcı, Selma Çalışkan, Bülent Bakar

49 **Clinical and radiological response assessment after whole brain radiotherapy for patients with non-small cell lung cancer with brain metastases**
Küçük hücreli dışı akciğer kanserli beyin metastazlı hastalarda tüm beyin radyoterapisi sonrası klinik ve radyolojik yanıt değerlendirme
Funda Çukurçayır, Mert Saynak, Canan Çelebi, Alaattin Özen, Görkem Türkkan, Kamuran İbiş, Mustafa Akın, Zafer Koçak

60 **Evaluation of viral respiratory pathogens detected in a tertiary university hospital between 2021 and 2025**
Üçüncü basamak bir üniversite hastanesinde 2021-2025 tarihleri arasında saptanan viral solunum yolu etkenlerinin değerlendirilmesi
Canberk Çınar, Sümeyye Özkaya, İlknur Bıyık, Yeliz Tanrıverdi Çaycı, Demet Gür Vural, Kemal Bilgin, Asuman Birinci

65 **Comparative analysis of pH measurements: Urine dipstick and pH meter**
pH ölçümlerinin karşılaştırmalı analizi: İdrar dipstik ve pH metre
Murat Can, Berrak Güven, Osman Emre Ünlü

70 **Applications and research trends of large language models in healthcare: A bibliometric analysis**
Sağlık hizmetlerinde büyük dil modellerinin uygulamaları ve araştırma eğilimleri: Bibliyometrik bir analiz
Hakan Öztürk, Elvan Hayat

81 **Clinical characteristics and hematological inflammatory markers in pediatric peripheral facial paralysis: A retrospective study**
Pediatrik periferik fasiyal paralizide klinik özellikler ve hematolojik inflamatuvar belirteçler: Retrospektif bir çalışma
Meral Karadağ, Melek Yaman Ortaköylü

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi
Cilt / Volume 10 Sayı / Number 1 Nisan / April 2026

Evaluation of practical application of Tell-Hashomer, Livneh and Yalçınkaya-Özen criteria for diagnosis of Familial Mediterranean fever in children
89 Çocuklarda Ailevi Akdeniz Ateşi tanısında Tel-Hashomer, Livneh ve Yalçınkaya-Özen kriterlerinin pratikteki uygulamasının değerlendirilmesi
Tuğba Kazlı, Müferet Ergüven

The effect of a surgical nursing course structured with role-playing and flipped learning methods on students' motivation and attitudes toward clinical practice
97 Rol oynama ve ters yüz öğrenme yöntemi ile yapılandırılmış cerrahi hemşireliği dersinin öğrencilerin motivasyonuna ve klinik uygulamaya yönelik tutumuna etkisi
Sibel Altıntaş, Sevim Çelik, Elif Karahan, Özge Uçar, Meryem Yücel

Postoperative frequency of normocalcemic elevated parathormone in patients surgically treated for primary hyperparathyroidism: A single-center experience
106 Primer hiperparatiroidizm sebebiyle cerrahi tedavi uygulanan hastalarda ameliyat sonrası normokalsemik yüksek parathormon sıklığı: Tek merkez deneyimi
Gülşah Elbüken, Ogün İrem Bilen, Beyza Olcay Öztürk, Hünkar Ağgül, Sibel Özkan-Gürdal, Burcu Tokuç, Sayid Shafi Zuhur

Adneksiyal kitlesi olan kadınlarda siberkondri düzeyini etkileyen faktörler: Tek merkezli kesitsel çalışma
116 Factors affecting the level of cyberchondria in women with adnexal masses: A single-center cross-sectional study
Hande Nur Öncü, Gökçen Ege, Neslihan Öztürk, Candost Hanedan, Vakkas Korkmaz

Sensory perception profiles in children with epilepsy: A prospective controlled study
122 Epilepsili çocuklarda duyusal algı profilleri: Prospektif kontrollü bir çalışma
Seda Ayaz Taş, Merve Başol Göksülük, Özge Karanlık Özcan, Seda Yakıt Yeşilyurt, Sedef Kestellioğlu, Sezen Tezcan, Ayşegül Danış

Effect of intraoperative magnesium sulfate infusion on early postoperative opioid consumption in laparoscopic hysterectomy: A randomized, double-blind, controlled trial
133 Laparoskopik histerektomide intraoperatif magnezyum sülfat infüzyonunun erken postoperatif opioid tüketimine etkisi: Randomize, çift kör, kontrollü çalışma
İlke Tamdoğan, Dilek Yeniay, Fatma Alkan Bayburt, Tuna Albayrak, Azime Bulut

Effects of preemptive analgesia and preoperative anxiety on emergence delirium after pediatric adenotonsillectomy
142 Pediatrik adenotonsillektomi sonrası uyanma deliryumu üzerine preemptif analjezi ve preoperatif anksiyetenin etkileri
Gamze Küçükosman, Bengü G. Köksal, Keziban Bollucuoğlu, Rahşan D. Okyay, Dilay İ. Alp, Özcan Pişkin, Hilal Ayoğlu

Relationship between demodex infestation and clinical parameters in polycystic ovary syndrome: A prospective observational study
151 Polikistik over sendromunda demodex enfestasyonu ile klinik parametreler arasındaki ilişki: Prospektif gözlemsel çalışma
Mehmet Bülbül, Mustafa Kaplanoğlu, Tuncay Çelik, İpek Çakılkaya, Emel Dündar

Clinical and radiological outcome comparison of unilateral vs. bilateral percutaneous vertebroplasty for osteoporotic vertebral compression fractures
157 Osteoporotik omurga kompresyon kırıklarında tek taraflı ve çift taraflı perkütan vertebroplasti uygulamalarının klinik ve radyolojik sonuçlarının karşılaştırılması
Levent Aydın, Buse Sarıgül, Tufan Agah Kartum, Gonca Gül Öndüç, Saime Ayça Şahin, Ali Fatih Ramazanoğlu

Olgu Sunumu / Case Report

Hemichorea-hemiballismus developing due to non-ketotic hyperglycemia: A rare case report
164 Non-ketotik hiperglisemiye bağlı gelişen hemikore-hemiballismus: Nadir bir olgu sunumu
Merve İrem Atıcı, Baycan Kuş, Taner Şahin, İrem Kömürcü, Merva Tuna

Derleme / Review

Artificial intelligence in pain: A comprehensive review
168 Ağrıda yapay zekâ: Kapsamlı bir derleme
İlhan Celil Özbek, Erkan Özduran, Volkan Hancı

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):1-8
DOI: 10.29058/mjwbs.1698765

Investigation of factors associated with the level of temporomandibular dysfunction in university students

Üniversite öğrencilerinde temporomandibular disfonksiyon düzeyi ile ilişkili faktörlerin incelenmesi

Merve Bulguroğlu[1], Halil İbrahim Bulguroğlu[1], Gökçe Kılıç[2], Selcan Suiçmez[1], Serenay Zorlu[1], Cansu Gevrek Aslan[1]

[1]Department of Physiotherapy and Rehabilitation, Ankara Medipol University, Faculty of Health Sciences, Ankara, Türkiye
[2]Dr. Dt. Gökçe Kılıç Private Dental Clinic, Ankara, Türkiye

Cite this article as: *Bulguroğlu M et al. Investigation of factors associated with the level of temporomandibular dysfunction in university students. Med J West Black Sea. 2026;10(1):1-8.*

ABSTRACT

Aim: Temporomandibular disorders (TMD) and factors associated with them are frequently observed in university students. The aim of this study was to investigate the relationship between different severities of TMD and deep neck flexor muscle endurance, craniovertebral angle, upper neck mobility, and stress levels in university students.

Material and Methods: Our cross-sectional study included 228 students from the Department of Physiotherapy and Rehabilitation at Ankara Medipol University. TMD severity was assessed using the Fonseca Anamnestic Index (FAI), deep neck flexor muscle endurance using the Craniocervical flexion test (CCFT), head-forward posture using the craniovertebral angle (CVA), upper neck mobility using the cervical flexion-rotation test (CFRT), and stress levels using the Perceived Stress Scale (PSS).

Results: Moderate negative correlations were found between FAI and CVA (r=-0.495, p<0.001), CFRT-right (r=-0.486, p<0.001), and CFRT-left (r=-0.443, p<0.001). Weak negative correlations were observed between FAI and both the deep neck flexor activation score (r=-0.352, p<0.001) and performance index score (r=-0.323, p<0.001). In addition, a weak positive correlation was found between FAI and PSS (r=0.327, p<0.001). Significant differences were observed in all measurements across the four TMD severity groups (p<0.05).

Conclusion: TMD in university students is associated with craniovertebral angle, deep neck flexor muscle endurance, upper neck mobility, and stress levels. As TMD severity increases, all these parameters are negatively affected. This study highlights the importance of evaluating TMD and related symptoms in university students.

Keywords: Posture, stress, students, temporomandibular joint disorders

ÖZ

Amaç: Üniversite öğrencilerinde temporomandibular bozukluklar (TMB) ve bu bozukluklarla ilişkili faktörler sıklıkla gözlemlenmektedir. Bu çalışmanın amacı, üniversite öğrencilerinde farklı temporomandibular bozukluk şiddetleri ile derin boyun fleksör kas dayanıklılığı, kraniovertebral açı, üst boyun mobilitesi ve stres düzeyleri arasındaki ilişkiyi incelemektir.

Gereç ve Yöntemler: Kesitsel bir çalışma olarak planlanan araştırmamıza, Ankara Medipol Üniversitesi Fizyoterapi ve Rehabilitasyon bölümünden çalışmaya 228 öğrenci dahil edildi. Bireylerin TMB şiddeti Fonseca Anamnestik İndeksi (FAİ) ile, derin boyun fleksör kas enduransı Kranioservikal fleksiyon testi (KSFT) ile, baş önde postür kraniovertebral açı (KVA) ile, üst boyun mobilitesi servikal fleksiyon-rotasyon testi (SFRT) ile, stres düzeyleri ise Algılanan Stres Ölçeği (ASÖ) kullanılarak değerlendirildi.

Bulgular: FAİ ile KVA (r=-0.495, p<0.001), SFRT-sağ (r=-0.486, p<0.001) ve SFRT-sol (r=-0.443, p<0.001) arasında orta düzeyde negatif anlamlı bir korelasyon bulundu. Derin boyun fleksörleri-aktivasyon skoru (r=-0.352, p<0.001) ve derin boyun fleksörleri-performans indeks skoru (r=-0.323, p<0.001) ile FAİ arasında zayıf düzeyde negatif anlamlı bir korelasyon belirlendi. Ayrıca, FAİ ile ASÖ arasında zayıf düzeyde pozitif anlamlı bir korelasyon saptandı (r=0.327, p<0.001). TMB şiddeti farklı olan dört grup arasında tüm ölçüm sonuçlarında anlamlı fark bulundu (p<0.05).

Sonuç: Üniversite öğrencilerinde temporomandibular bozukluklar, kraniovertebral açı, derin boyun fleksör kas dayanıklılığı, üst servikal mobilite ve stres düzeyleri ile ilişkilidir. Bozukluk şiddeti arttıkça tüm bu parametreler olumsuz yönde etkilenmektedir. Bu çalışma, üniversite öğrencilerinde TMB ve ilişkili semptomların değerlendirilmesinin önemini vurgulamaktadır.

Anahtar Kelimeler: Öğrenciler, postür, stres, temporomandibular eklem bozuklukları

Corresponding Author: Halil İbrahim Bulguroğlu ✉ fztibrahim@hotmail.com **Received:** 13.05.2025 **Revision:** 01.09.2025 **Accepted:** 26.02.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Temporomandibular disorder (TMD) was prevalent among university students, with more than half demonstrating mild to severe symptoms.
- TMD severity showed moderate negative correlations with craniovertebral angle and upper cervical mobility.
- Deep neck flexor muscle endurance decreased as TMD severity increased.
- Perceived stress levels were positively associated with TMD severity.
- Cervical posture and mobility exhibited large effect sizes across TMD severity groups, indicating clinical relevance.

INTRODUCTION

The term "temporomandibular disorders" (TMD) refers to conditions affecting the muscles and joints responsible for the movement of the temporomandibular joint and its structures. The etiology of TMD is multifactorial. Individuals with TMD may be asymptomatic or may present with symptoms such as pain, limited movement, and joint sounds (1). Previous studies have revealed that 85% to 95% of individuals experience at least one TMD symptom during their lifetime, with peak incidence occurring between 20 and 40 years of age (2). Furthermore, epidemiological studies have shown that TMD prevalence among students ranges from 50% to 77%, while in Turkish student populations this rate varies between 47% and 81% (3). These findings underscore the importance of investigating TMD-related factors in young adult populations, particularly in academic settings.

Cervical spine problems and postural disorders may also be associated with TMD symptoms due to altered head posture influencing mandibular position and masticatory muscle activity. The craniocervical and craniomandibular systems interact anatomically, as the mandibular muscles connect the hyoid, cranial, and cervical vertebrae, suggesting a direct relationship between the temporomandibular joint (TMJ) and the cervical spine (4).

Considering the biomechanical link, extension movement is seen in the craniocervical junction during standard mouth opening, and limitations in the upper cervical spine may reduce mouth opening capacity. Reduced mouth opening capacity or pain at maximum mouth opening may cause different craniocervical postures (5). Studies have shown that individuals with TMD have lower upper cervical mobility values than those without TMD (6,7). Cranio-cervical flexion (CCF) is provided by the deep cervical flexor (DCF) muscles at the atlanto-occipital joint and upper cervical segments C1-2. These are the main postural muscles in the anterior neck that support cervical lordosis. Studies have shown that deep cervical flexor strength and endurance are reduced, superficial muscle activity is increased, and mandibular functions are negatively affected in individuals with TMD (8).

There has been much research on the link between forward head posture (FHP) and TMD, including a case–control study of individuals with myogenic TMD (2), a study investigating cranio-cervico-mandibular function and postural stability in patients with TMD (8), and a cross-sectional study assessing masticatory muscle pressure pain thresholds in patients with TMD (9). FHP is characterized by a head deviating forward in the sagittal plane due to increased extension in the upper cervical spine and increased flexion in the lower cervical spine. When assessed by craniovertebral angle (CVA), individuals with TMD have significantly higher FHP than those without TMD (8).

In most individuals with TMD, psychological factors are highly influential in the onset and progression of the dysfunction (10). Individuals with TMD report exposure to stressful life situations before developing their symptoms. Evidence suggests that stress leads to increased muscle activity, which causes pain in areas close to the TMJ. Such pain and dysfunction may negatively affect mastication, speech, and cervical function, and can also impair academic and professional performance, ultimately reducing social participation and quality of life (10,11).

University students, due to long-term poor posture related to excessive computer and smartphone use, are considered a significant risk group (12). The progression of TMD symptoms in this population can negatively affect both academic performance and quality of life. Therefore, early identification of individuals with TMD and determination of risk factors are important for preventing progression and protecting public health.

The aim of this study was to investigate the relationship between different severities of temporomandibular disorders (TMD) and deep neck flexor muscle endurance, craniovertebral angle, upper neck mobility, and stress levels in university students.

MATERIAL and METHODS

Study Design and Participants

Using the Pearson correlation coefficient r=0.30 with 80% power (alpha=0.05, two-way) in the G*power program, the power analysis for the study's sample size showed that 200 participants were needed to perform the correlation analysis. Our study, which was designed to be cross-sectional, involved 234 students aged 18 to 22 who were enrolled in Ankara Medipol University's Department of Physiotherapy and Rehabilitation from 2024–2025. Six students who did not want to make the evaluations left the study, and the study was completed with 228 participants. Individuals with neurological or chronic systemic diseases, individuals with psychiatric disorders, individuals with a history of trauma in the last six months, individuals with chronic pain syndromes such as trigeminal neuralgia, individuals with rheumatic or orthopedic problems that may affect TMJ, individuals who have undergone musculoskeletal surgery, individuals with diagnosed TMD, and individuals who have received any dental treatment in the last six months were not included in the study. Before starting the study, Ankara Medipol University Non-Interventional Clinical Research Ethics Committee approval was obtained (Date: 25/12/2024 Decision No: 170). Before the study, each participant received information about it and signed the "Informed Consent Form." Our study followed the principles of the Helsinki Declaration.

Measuring Methods

The individuals' demographic information was obtained. The questionnaires assessed TMD and stress levels, while the upper neck mobility, craniovertebral angle, and deep neck flexor muscle endurance were assessed using objective measurement methods. All measurements were made outside the exam period.

Ten questions from the Fonseca Anamnestic Index (FAI) were utilized in this investigation to categorize the presence and severity of TMD (13). Joints, head and neck pain, pain when chewing, parafunctional behaviors, reduced joint mobility, occlusal disorders, and mental stress are all covered in the questionnaire. Each question has three possible answers: "Yes: 10 points", "no: 0 points", and "Sometimes: 5 points". The questionnaire's results are classified into four categories based on the participants' answers: mild TMD (20–40), moderate TMD (45–65), severe TMD (70–100), and no TMD (0–15). According to Kaynak et al. (14), whose Turkish validity and reliability assessment was carried out, the questionnaire's Cronbach's alpha value was 0.80. Our study had a Cronbach's alpha of 0.79.

The endurance of the DNF muscles was assessed using the Cranio Cervical Flexion Test (CCFT) (15). An uninflated Pressure Biofeedback Unit (Chattanooga Stabilizer, USA) was positioned between the earlobe and the chin prominence, with the subject in the chin position and the neck in a neutral posture. It was then inflated to an initial pressure of 20 mmHg. Individuals were first instructed to place their tongues on the roof of their mouths, with their lips close together and their teeth touching, thus minimizing muscle activity in the jaw area. A practice session was then conducted to reduce the possibility of not understanding and performing the test as recommended (16). While the examiner gave visual feedback, the subject was asked to perform a slow and smooth craniocervical flexion without lifting the head and to nod as if to indicate "yes." In incremental increments of 2 mmHg, the DNFs' activation and performance index scores were measured between 20 and 30 mmHg. The pressure level that could be maintained for ten repeats of ten seconds each was used to calculate the activation score. The performance index score was computed based on how many repetitions the person could keep the position for 10 seconds at the pressure level. For example, if the individual could reach the second pressure level (24 mmHg) and correctly hold the craniocervical flexion for 10 seconds five times, the performance index is $4 \times 5 = 20$. The highest activation score is 10 mmHg, and the highest performance index score is 100.

FHP was evaluated using the CVA measurement (17). Markers were placed on the ear tragus, and the spinous process of the seventh cervical vertebra (C7) was used to employ lateral photography for measuring. The participant was 200 cm away from the digital camera positioned at acromion level to record CVA. The tripod held the camera in place. This technique shows high reliability (ICC = 0.94). Participants were asked to maintain a comfortable posture by sharing their weight symmetrically on both lower extremities with their feet shoulder-width apart. Before taking the photograph, each participant was asked to flex and extend their neck three times. Participants were then instructed to look straight ahead in their natural posture. An angle was measured between a line that extended from the ear tragus to the C7 vertebra and a horizontal line that went through the C7 spinous process. As a result of the measurement, the photographs were analyzed with the Image-J program (18). The smaller the CVA, the more forward the head posture. The average value of the three measurements was recorded as CVA.

Upper neck mobility was assessed with the Cervical flexion-rotation test (CFRT) (19). The participant's comfort was a priority throughout the assessment. The cervical range of motion device was securely attached to the head of the participant while lying on his back, and the examiner brought him to the limit of cervical flexion. In this flexion position, the head and neck were passively rotated within pain and physiological limits. Measurements were made without encoun-

tering resistance or pain. The interval was recorded. The average of three range of motion measurements on both sides was calculated, with a 30-second rest between each repetition, and this represented the final values for the left and right ranges of motion.

The students' stress levels were assessed using the 14-item Perceived Stress Scale (PSS) designed by Cohen et al. (20). The scale includes positive and negative questions about stress. It is a self-report scale designed to assess the perception of the degree to which a given situation is considered stressful. Each item is scored between 0 and 4; a total PSS score of 0 to 56 is obtained by summing all items. The higher the score, the greater the perceived stress. According to validity and reliability research by Eskin et al. (21), the Turkish version had a Cronbach's alpha value with subscales ranging from 0.66 to 0.88. The Cronbach's alpha value in this investigation was 0.75.

Statistical Analysis

The statistical analyses of the study were performed using the 'Statistical Package for Social Sciences ' (SPSS) version 21.0 (SPSS Inc. Chicago, IL, USA). To assess the normal distribution of variables, visual methods (histograms, probability plots), skewness and kurtosis values, and analytical tests were considered. Given the sample size, the Kolmogorov–Smirnov test was primarily used, while the Shapiro–Wilk test was reported only for comparative purposes. The results were consistent across all approaches. Categorical variables were expressed as percentage (%) and number of people (n), while mean and standard deviation (SD) values for continuous variables were also defined. A one-way ANOVA test was used to compare the evaluation parameters between groups. Effect sizes for the one-way ANOVA were calculated using partial eta squared (ηp^2) and interpreted as small (0.01), medium (0.06), and large (0.14), according to conventional thresholds. Pearson correlation analysis was used to determine the relationship between evaluation parameters. The statistical significance level was consistently set at $p < 0.05$ in all tests.

RESULTS

The demographic characteristics of the 228 university students who participated in the study are shown in Table 1. The university students' mean CVA was 44.93±6.60. In CCFT, mean DNF- activation scores were 5.82±2.03, while mean DNF- performance index scores were 64.54±26.94. The mean angles of the CFRT measurements were 41.83±11.97 on the right side and 40.07±10.85 on the left side. The mean scores of PSS were 29.38±8.71 (Table 2).

Table 3 shows the severity levels of Temporomandibular disorders in the participants. In our study, moderate negative correlations were observed between FAI and CVA (r=-0.495) (p<0.05), CFRT-R (r=-0.486) (p<0.05), and CFRT-L (r=-0.443) (p<0.05). Weak negative correlations were found between FAI and DNF-Activation score (r=-0.352) (p<0.05) and DNF-performance index score (r=-0.323) (p<0.05). A weak correlation was observed between FAI and PSS (r=0.327) (p<0.05, Table 4).

Table 1: Demographic characteristics of participants

		Participants (n=228)	
		$\bar{X}\pm s$	
Age (years)		19.32±2.41	
BMI (kg/m²)		22.21±3.42	
		n	%
Gender	Female	149	65.35
	Male	79	34.65

x̄±s: mean±SD, **m:** meter, **kg:** kilogram, **BMI:** body mass index, **n:** sample size.

Table 2: Participants' Fonseca anamnestic index, Craniovertebral angle, Cranio-cervical flexion test, Cervical flexion-rotation test and stress measurement results

	Participants (n=228)
	$\bar{X}\pm s$
FAI score (0-100)	25.74±22.04
CVA (°)	44.93±6.60
CCFT	
DNF- Activation score	5.82±2.03
DNF- Performance index score	64.54±26.94
CFRT	
Right (°)	41.83±11.97
Left (°)	40.07±10.85
PSS score	29.38±8.71

x̄±s: mean±SD, **FAI:** fonseca anamnestic index, **CVA:** craniovertebral angle, **CCFT:** cranio-cervical flexion test, **DNF:** deep neck flexor, **CFRT:** cervical flexion-rotation test, **(o):** angle, **PSS:** perceived stress scale, **n:** sample size.

Table 3: Temporamandibular disorders severity levels of participants

	Participants (n=228)	
	n	%
FAI Score (0-100)		
No TMD (0–15)	108	47.37
Mild TMD (20–40)	76	33.33
Moderate TMD (45–65)	32	14.03
Severe TMD (70–100)	12	5.27

FAI: fonseca anamnestic index, **TMD:** temporamandibular disorders, **n:** sample size.

In Table 5, significant differences were observed among the four groups according to TMD severity levels. Effect size analysis (ηp^2) indicated large effects for craniovertebral angle and cervical flexion–rotation test measurements, and moderate effects for deep neck flexor endurance and perceived stress.

DISCUSSION

This study showed that TMD is prevalent among university students. Its severity correlates with craniovertebral angle, DNF endurance, upper neck mobility, and stress levels.

Our study determined that the average FAI score of university students was 25.74±22.04, and only 47.37% did not have TMD. Studies in the literature show that the frequency of TMD is high in university students (11,22,23). Studies like ours have indicated that students' TMD severity levels are mostly mild. In addition, our study's 5.27% severe TMD level shows the importance of differential diagnosis in this population. University years are when various health problems occur, and unfortunately, people usually do not care too much. Since TMD etiology is very complex, it is confused with many health problems and, unfortunately, is mainly ignored by students' self-medication (24). Therefore, differential diagnoses, especially in this population, can reduce the progression of TMD and prevent different problems.

We found the mean CVA to be 44.93±6.60. In addition, a moderate negative correlation was found between FAI and CVA ($r=-0.495$, $p<0.05$), indicating that CVA decreased as TMD severity increased. Studies similar to ours in the literature show that CVA, used in evaluating head forward posture, is associated with TMD (9,25). One study stated that forward head posture is typical in individuals with TMD, but this forward head posture is not associated with TMD symptoms in the absence of pain (9). In another study, it was stated that the position of the head posture affects TMD (25). Our study observed a relationship between CVA, which determines the forward head posture, and TMD. The muscles in the craniomandibular and craniocervical regions should work in harmony, but there may be imbalances of strength and tension between these muscles in forward head postures. This situation may even affect the position of the mandible (26). Considering the ergonomic problems and hormonal irregularities that are particularly common among university students (27), the likelihood of developing forward head posture may increase. The moderate strength of the correlation, together with the observed group differences, suggests that cervical posture may be clinically relevant in individuals with higher TMD severity.

The correlation between the increase in TMD severity level and CVA supports our results. Although studies indicate that DNF and upper cervical mobility are affected in individuals

Table 4: Relationship between participants' Fonseca anamnestic index scores and Craniovertebral angle, Cranio-cervical flexion test, Cervical flexion-rotation test and stress level measurement scores

	Participants (n=228)
	Fonseca Anamnestic Index Score (0-100)
CVA	r= -0.495 p< 0.001*
DNF-Activation score	r= -0.352 p< 0.001*
DNF-Performance index score	r= -0.323 p< 0.001*
CFRT-R	r= -0.486 p< 0.001*
CFRT-L	r= -0.443 p< 0.001*
PSS score	r= 0.327 p< 0.001*

Pearson correlation analysis, **r:** correlation coefficient, **CVA:** craniovertebral angle, **DNF:** deep neck flexors, **CFRT:** cervical flexion-rotation test, **R:** right, **L:** left, **PSS:** perceived stress scale, **n:** sample size.

Table 5: Comparison of participants' Craniovertebral angle, Cranio-cervical flexion test, Cervical flexion-rotation test and stress level measurements according to temporomandibular severity levels

	TMD Severity Levels					
	No TMD	**Mild TMD**	**Moderate TMD**	**Severe TMD**	**p**	**ηp^2**
CVA	49.29±4.65	43.39±4.44	39.28±5.79	30.95±2.28	0.001*	0.347
DNF-activation score	6.33±1.92	5.61±1.03	5.12±1.82	4.53±1.18	0.039*	0.102
DNF-performance score	74.64±21.26	57.94±28.64	53.40±29.36	46.00±35.55	0.003*	0.130
CFRT-R	45.99±7.99	41.98±13.74	33.25±10.16	26.10±3.49	0.001*	0.228
CFRT-L	45.18±7.52	37.42±11.75	33.42±9.31	26.99±4.82	0.001*	0.260
PSS	27.40±7.87	28.85±8.30	34.64±6.67	37.15±8.06	0.014*	0.129

A one-way ANOVA, **TMD:** temporomandibular disorders, **CVA:** craniovertebral angle, **DNF:** deep neck flexors, **CFRT:** cervical flexion-rotation test, **R:** right, **L:** left, **PSS:** perceived

with TMD (9,10), no study has evaluated their relationships with different TMD levels. As a result of our study, the mean DNF-Activation score was 5.82±2.03, while the DNF-Performance index score was 64.54±26.94. In addition, in the CFRT angle evaluating upper cervical mobility, the right side average was 41.83±11.97, and the left side average was 40.07±10.85. Weak negative correlations were found between FAI and DNF-Activation ($r=-0.352$, $p<0.05$) and DNF-Performance Index ($r=-0.323$, $p<0.05$), while moderate negative correlations were observed between FAI and CFRT angles ($r=-0.486$ right; $r=-0.443$ left; $p<0.05$). In line with our study, it has been shown in the literature that individuals with TMD have lower CCFT scores, which provide DNF measurement, and lower CFRT angles than healthy populations (4-7). Considering the anatomical position and functional duties of the longus capitis and longus colli muscles, which are responsible for DNF, and their proximity to the C1-2 region, where upper cervical region mobility is provided, it is typical for both DNF and upper cervical mobility to be affected by postural changes, especially head posture (28). In particular, due to imbalances caused by postural changes, the superficial muscles are more activated, which can negatively affect the activities of deep muscles (26). Our study found that head-forward posture was more common in university students with TMD. It is usual for the deep muscles to weaken with this posture (26). In addition, the primary function of the DNF muscles is to provide neck stabilization (29). The observed weak correlations with DNF measures and moderate correlations with upper cervical mobility suggest that postural imbalances may progressively affect both stabilization and mobility in this region.

Studies have shown a relationship between TMD and stress levels in university students (30,31). In our study, in line with the literature, a weak positive correlation was observed between FAI and PSS scores ($r=0.327$, $p<0.05$), indicating that stress levels increased with the progression of TMD severity. The academic stress experienced by university students and the stress levels created in their bodies by their irregular lives may create a vicious cycle for TMD (11). In addition, emotional stress in students may have caused muscle hyperactivity and may have prepared the ground for the formation of TMD. Studies have indicated that psychological factors are effective in the onset and progression of dysfunction in most individuals with TMD (31,32). Recent systematic reviews have further highlighted that stress, anxiety, and depression are closely associated with TMD and headache disorders across different age groups (32) and that depression is significantly linked with TMD in adults (33). Moreover, clinical studies suggest that TMD management approaches such as dry needling and face yoga may also improve psychosocial outcomes, including depression and sleep quality (34). Our findings, consistent with these reports, suggest that psychological stress plays a contributory but relatively weak role in the progression of TMD among university students, potentially influencing their academic performance and psychosocial well-being.

This study has some limitations. First, its cross-sectional design precludes causal inferences, and future longitudinal research is needed to clarify temporal relationships. Second, lifestyle factors such as smoking, bruxism, and sleep quality (34), as well as gender-related differences, were not assessed in detail, which may have influenced the outcomes. Third, the single-center sample consisting only of physiotherapy and rehabilitation students may limit the generalizability of the findings. Despite these limitations, this study provides novel insights by examining TMD severity in relation to cervical posture, muscle endurance, mobility, and stress in a relatively large university population.

Conclusion

Our study emphasizes the importance of identifying TMD and its associated factors in university students. If this disorder is not addressed, it can lead to increased problems later in life. Therefore, when symptoms that may be related are observed, it is crucial to evaluate TMD and take appropriate actions. Future studies could explore the effects of different types of exercise on TMD and associated factors in university students.

Acknowledgments

The authors thank all the participants who volunteered to participate the study.

Author Contributions

Study conception and design: **Merve Bulguroğlu, Halil İbrahim Bulguroğlu;** data collection: **Merve Bulguroğlu, Selcan Suiçmez, Serenay Zorlu, Cansu Gevrek Aslan, Gökçe Kılıç;** analysis and interpretation of results: **Cansu Gevrek Aslan, Halil İbrahim Bulguroğlu;** draft manuscript preparation: **Merve Bulguroğlu, Halil İbrahim Bulguroğlu, Gökçe Kılıç.** The authors reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Financial Support

The authors declare that the study received no funding.

Ethical Approval

This study was carried out with the approval of the Ankara Medipol University Non-Interventional Clinical Research Ethics Committee, dated 25/12/2024 and numbered 170.

REFERENCES

1. Kapos FP, Exposto FG, Oyarzo JF, Durham J. Temporomandibular disorders: a review of current concepts in aetiology, diagnosis and management. Oral Surg. 2020;13(4):321-334. https://doi.org/10.1111/ors.12473

2. Akarca EM, Karabulut Açıkgöz T, Senyilmaz D, Canay S. Prevalence and Severity of Temporomandibular Disorders among the Turkish Dental Students Using Fonseca's Anamnestic Index. Clin Dent Res. 2025;49(2):110-118.
3. Altın FN, Aslan Telci E, Özden F. Evaluation of factors associated with the temporomandibular joint dysfunction in university students. Bull Fac Phys Ther. 2024;29:83. https://doi.org/10.1186/s43161-024-00249-9
4. Yüzbaşıoğlu Ü, Kaynak BA, and Taş S. Assessment of Cervical Joint Position Sense and Head Posture in Individuals with Myogenic Temporomandibular Dysfunctions and Identifying Related Factors: A Case-Control Study. J Oral Rehabil. 2025;52:160-168. https://doi.org/10.1111/joor.13885
5. Lam AC, Liddle LJ, and MacLellan CL. The Effect of Upper Cervical Mobilization/Manipulation on Temporomandibular Joint Pain, Maximal Mouth Opening, and Pressure Pain Thresholds: A Systematic Review and Meta-Analysis. Arch Rehab Res Clin. 2022;5:100242. https://doi.org/10.1016/j.arrct.2022.100242
6. Grondin F, Hall T, Laurentjoye M, Ella B. Upper cervical range of motion is impaired in patients with temporomandibular disorders. Cranio. 2015;33(2):91-99. https://doi.org/10.1179/0886963414Z.00000000053
7. Ferreira MP, Waisberg CB, Conti PCR, Bevilaqua-Grossi D. Mobility of the upper cervical spine and muscle performance of the deep flexors in women with temporomandibular disorders. J Oral Rehabil. 2019;46(12):1177-1184. https://doi.org/10.1111/joor.12858
8. Miçooğulları M, Yüksel İ, and Angın S. Effect of pain on cranio-cervico-mandibular function and postural stability in people with temporomandibular joint disorders. Korean J Pain. 2024;37:164-177. https://doi.org/10.3344/kjp.23301
9. Yao Y, Cai B, Fan S, Yang HX, Zhang YX, Xu LL. The association between forward head posture and masticatory muscle pressure pain thresholds in patients with temporomandibular joint dissorders: a cross-sectional observational study. Clin Oral Investig. 2023;27(1):353-360. https://doi.org/10.1007/s00784-022-04739-9
10. Dos Santos EA, Peinado BRR, Frazão DR, Né YGS, Fagundes NCF, Magno MB, Maia LC, Lima RR, de Souza-Rodrigues RD. Association between temporomandibular disorders and anxiety: A systematic review. Front Psychiatry.2022;13:990430. https://doi.org/10.3389/fpsyt.2022.990430
11. Augusto VG, Perina KCB, Penha DSG, Dos Santos DCA, Oliveira VAS. Temporomandibular Dysfunction, Stress and Common Mental Disorder in University Students. Acta Ortop Bras. 2016;24:330-333. https://doi.org/10.1590/1413-785220162406162873
12. Eitivipart AC, Viriyarojanakul S, and Redhead L. Musculoskeletal disorder and pain associated with smartphone use: A systematic review of biomechanical evidence. Hong Kong Physiother J. 2018;38:77-90. https://doi.org/10.1142/S1013702518300010
13. Fonseca DMD, Bonfante G, Valle ALD, Freitas SFTD. Diagnóstico pela anamnese da disfunção craniomandibular. RGO. 1994; 42:23-28.
14. Kaynak BA, Taş S, and Salkın Y. The accuracy and reliability of the Turkish version of the Fonseca anamnestic index in temporomandibular disorders. Cranio. 2023;41:78-83. https://doi.org/10.1080/08869634.2020.1812808
15. Jull GA, O'Leary SP, and Falla DL. Clinical assessment of the deep cervical flexor muscles: the craniocervical flexion test. J Manipulative Physiol Ther. 2008;3:525-533. https://doi.org/10.1016/j.jmpt.2008.08.003
16. Kelly M, Cardy N, Melvin E, Reddin C, Ward C, Wilson F. The craniocervical flexion test: an investigation of performance in young asymptomatic subjects. Man Ther. 2013;18(1):83-86. https://doi.org/10.1016/j.math.2012.04.008
17. Salahzadeh Z, Maroufi N, Ahmadi A, et al. Assessment of forward head posture in females: observational and photogrammetry methods. J Back Musculoskelet Rehabil. 2014;27(2):131-139. https://doi.org/10.3233/BMR-130426
18. Jabbar KM and Gandomi F. The comparison of two corrective exercise approaches for hyperkyphosis and forward head posture: A quasi-experimental study. J Back Musculoskelet Rehabil. 2021;34:677-687. https://doi.org/10.3233/BMR-200160
19. Takasaki H, Hall T, Oshiro S, Kaneko S, Ikemoto Y, Jull G. Normal kinematics of the upper cervical spine during the Flexion-Rotation Test - In vivo measurements using magnetic resonance imaging. Man Ther. 2011;16(2):167-171. https://doi.org/10.1016/j.math.2010.10.002
20. Cohen S, Kamarck T, and Mermelstein R. A global measure of perceived stress. J Health Soc Behav. 1983;24:385-96. https://doi.org/10.2307/2136404
21. Eskin M. The adaptation of the perceived stress scale into Turkish: a reliability and validity analysis. Journal of New Symposium. 2013;51:132-140.
22. Karthik R, Hafila MIF, Saravanan C, Vivek N, Priyadarsini P, Ashwath B. Assessing Prevalence of Temporomandibular Disorders among University Students: A Questionnaire Study. J Int Soc Prev Community Dent. 2017;7(Suppl 1):S24-S29. https://doi.org/10.4103/jispcd.JISPCD_146_17
23. Özdinç S, Ata H, Selçuk H, Can HB, Sermenli N, Turan FN. Temporomandibular joint disorder determined by Fonseca anamnestic index and associated factors in 18- to 27-year-old university students. Cranio. 2020;38:327-332. https://doi.org/10.1080/08869634.2018.1513442
24. Farias SJS, Ramos AG, Kogawa EM, Lia EN, de Medeiros RA. Prevalence of self-medication among university students diagnosed with temporomandibular disorders. J Oral Facial Pain Headache. 2024;38:58-63.
25. Kawasaki T, Ohji S, Aizawa J, Sakai T, Hirohata K, Koseki T, Kuruma H, Okawa A, Jinno T. Characteristics of head and neck alignment and function of deep cervical flexor muscles in patients with nonspecific neck pain. J Bodyw Mov Ther. 2024;39:565-571. https://doi.org/10.1016/j.jbmt.2024.03.063
26. Çobanoğlu G, Demirkan MY, Ecemiş ZB, Atalay Güzel N. Forward Head Posture and Its Effect on Muscle Activation. Gazi Sağlık Bil. 2024;9:85-93. https://doi.org/10.52881/gsbdergi.1376080
27. Ikenna UC, Nwobodo LN, Ezeukwu AO, Ilo IJ, Ede SS, Okemuo AJ, Okoh CF. Relationship between the development of musculoskeletal disorders, physical activity level, and academic stress among undergraduates students of University of Nigeria. J Educ Health Promot. 2022;11:399. https://doi.org/10.4103/jehp.jehp_416_22
28. Kennedy E, Albert M, and Nicholson H. Do longus capitis and colli really stabilise the cervical spine? A study of their fascicular anatomy and peak force capabilities. Musculoskelet Sci Pract. 2017;32:104-113. https://doi.org/10.1016/j.msksp.2017.10.005

29. Sherwin R, Henricksen J, Kapila J, Adams L, Likness A, Hooper TL, Sizer P. The effect of tongue-to-palate on deep neck flexor activity. J Electromyogr Kinesiol. 2024;79:102938. https://doi.org/10.1016/j.jelekin.2024.102938

30. Saczuk K, Lapinska B, Wawrzynkiewicz A, Witkowska A, Arbildo-Vega HI, Domarecka M, Lukomska-Szymanska M. Temporomandibular Disorders, Bruxism, Perceived Stress, and Coping Strategies among Medical University Students in Times of Social Isolation during Outbreak of COVID-19 Pandemic. Healthcare (Basel). 2022;10:740. https://doi.org/10.3390/healthcare10040740

31. Paz-Cortés MM, Martín-Vacas A, Mendoza PA, Rozalén M, Alvaredo MÁ. Analysis of the Effect of Exams on Perceived Stress and Temporomandibular Joint Status in Physiotherapy Students: A Pilot Study. Medicina (Kaunas). 2024;60:952. https://doi.org/10.3390/medicina60060952

32. Minervini G, Franco R, Marrapodi MM, et al. The Association between COVID-19 Related Anxiety, Stress, Depression, Temporomandibular Disorders, and Headaches from Childhood to Adulthood: A Systematic Review. Brain Sci. 2023;13(3):481. https://doi.org/10.3390/brainsci13030481

33. Uehara LM, Tardelli JDC, Botelho AL, Valente MLDC, Dos Reis AC. Association between depression and temporomandibular dysfunction in adults - a systematic review. Cranio. 2025;43(3):461-467. https://doi.org/10.1080/08869634.2022.2161985

34. Ozmen EE, Unuvar BS. The effects of dry needling and face yoga on pain, depression, function, and sleep quality in patients with temporomandibular dysfunction. Explore (NY). 2024;20(5):102980. https://doi.org/10.1016/j.explore.2024.01.006

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):9-24
DOI: 10.29058/mjwbs.1822432

In-silico engineering of an antimicrobial growth factor fusion protein to prevent post-laser infection in hemorrhoidal wounds

Hemoroidal yaralarda lazer sonrası enfeksiyonu önlemek için antimikrobiyal büyüme faktörü füzyon proteinlerinin bilgisayar simülasyonlu tasarımı

Özlem Karaca Ocak[1], **Zoha Naeem**[2]

[1]Department of General Surgery, Üsküdar University Faculty of Medicine, İstanbul, Türkiye
[2]Department of Bioinformatics, University of Central Punjab, Faculty of Biotechnology, Lahore, Pakistan

Cite this article as: *Karaca Ocak Ö and Naeem Z. In-silico engineering of an antimicrobial growth factor fusion protein to prevent post-laser infection in hemorrhoidal wounds. Med J West Black Sea. 2026;10(1):9-24.*

ABSTRACT

Aim: Post laser wound infection is a recurrent complication after hemorrhoidal surgery mainly because polymicrobial flora is colonized in the anorectal environment and the mucosal repair is impaired. This study aims to overcome the drawbacks of current therapeutic options and to provide a novel generation of prophylactic biomolecule with antimicrobial activities, monitored to meet the specific conditions of the microenvironment of anorectal wound repair.

Material and Methods: To overcome this two-fold problem of infection and slow wound healing, a current study developed an in-silico bioengineering intervention to design and characterize a new antimicrobial-growth factor fusion protein fusing selected antimicrobial peptide (AMPs) with the fibroblast growth factor 10 (FGF10) domain.

Results: The Antimicrobial Peptide Database was searched with four cationic AMPs, which included AP00608, AP02570, AP03795, and AP00952, and screened using AllerTOP v2.0 and ToxiPred, which confirmed their non-allergenicity and non-toxicity. The obtained engineered fusion construct possessed desirable physicochemical properties (32.84 kDa; pI 10.19; instability index 38.52; GRAVY -0.33; solubility 0.48) which implies high stability and hydrophilicity. Secondary structure prediction showed 27.34% alpha-helices and 16.26% beta-strands, whereas AlphaFold probability and ERRAT (94.28% proved structural accuracy). ClusPro 2.0 showed good binding affinities to FGFR2b (1NUN) and FGFR2 (1EV2) with energies of -1161.9 kcal mol -1 and -1144.4 kcal mol -1 respectively. Following 500 Desmond molecular dynamics simulations determined the thermodynamic stability, compact folding, and the maintenance of hydrogen bonds within the complexes.

Conclusion: All these findings suggest that the proposed fusion construct has a potential to serve as a dual-acting therapeutic - preventing microbial colonization and stimulating epithelial repair in post-laser hemorrhoidal wounds. This computational research therefore provides the basis of experimental validation and translational development of multifunctional wound-healing biologics in the future.

Keywords: Antimicrobial peptides, fibroblast growth factor 10, molecular docking, desmond simulation, in-silico fusion protein, wound healing, post-laser hemorrhoid infection

ÖZ

Amaç: Lazer sonrası yara enfeksiyonu, hemoroid cerrahisi sonrası tekrarlayan bir komplikasyondur, çünkü anorektal ortamda polimikrobiyal flora kolonize olur ve mukozal onarım bozulur. Bu çalışma, mevcut tedavi seçeneklerinin dezavantajlarını ortadan kaldırmayı ve anorektal yara onarımının mikroçevresinin özel koşullarını karşılamak üzere izlenen, antimikrobiyal aktiviteye sahip yeni nesil bir profilaktik biyomolekül sağlamayı amaçlamaktadır.

Gereç ve Yöntemler: Enfeksiyon ve yavaş yara iyileşmesi gibi bu iki sorunu aşmak için, mevcut bir çalışmada, seçilen antimikrobiyal peptidleri (AMP'ler) fibroblast büyüme faktörü 10 (FGF10) alanıyla birleştiren yeni bir antimikrobiyal-büyüme faktörü füzyon proteini tasarlamak ve karakterize etmek için bir in-silico biyomühendislik müdahalesi geliştirilmiştir.

Bulgular: Antimikrobiyal Peptid Veritabanı, AP00608, AP02570, AP03795 ve AP00952 dahil olmak üzere dört katyonik AMP ile taranmış ve AllerTOP v2.0 ve ToxiPred kullanılarak taranarak alerjenik ve toksik olmadıkları doğrulanmıştır. Elde edilen mühendislik füzyon yapısı, istenen fizikokimyasal özelliklere (32,84 kDa; pI 10,19; kararsızlık indeksi 38,52; GRAVY -0,33; çözünürlük 0,48) ve bu da yüksek stabilite ve hidrofiliklik anlamına gelmektedir. İkincil yapı tahmini %27,34 alfa sarmal ve %16,26 beta sarmal gösterirken, AlphaFold olasılığı ve ERRAT (%94,28 yapısal doğruluk kanıtladı). ClusPro 2.0, FGFR2b (1NUN) ve FGFR2 (1EV2) ile sırasıyla -1161,9 kcal mol -1 ve -1144,4 kcal mol -1 enerjilerle iyi bağlanma afiniteleri gösterdi. 500 ikonik Desmond moleküler dinamik simülasyonu, termodinamik stabiliteyi, kompakt katlanmayı ve kompleksler içindeki hidrojen bağlarının korunmasını belirledi.

Sonuç: Tüm bu bulgular, önerilen füzyon yapısının, lazer sonrası hemoroidal yaralarda mikrobiyal kolonizasyonu önleyen ve epitel onarımını uyaran çift etkili bir tedavi yöntemi olarak kullanılma potansiyeline sahip olduğunu göstermektedir. Bu hesaplamalı araştırma, bu nedenle gelecekte çok işlevli yara iyileştirici biyolojik ürünlerin deneysel olarak doğrulanması ve translasyonel geliştirilmesinin temelini oluşturmaktadır.

Anahtar Kelimeler: Antimikrobiyal peptitler, fibroblast büyüme faktörü 10, moleküler kenetlenme, desmond simülasyonu, in-silico füzyon proteini, yara iyileşmesi, lazer sonrası hemoroid enfeksiyonu

Corresponding Author: Özlem Karaca Ocak ✉ karacaocako@gmail.com **Received:** 14.11.2025 **Revision:** 19.02.2026 **Accepted:** 05.03.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- A novel antimicrobial–growth factor fusion protein was engineered to prevent infection in post-laser hemorrhoidal wounds.
- Candidate antimicrobial peptides were screened for allergenicity and toxicity using AllerTOP and ToxiPred.
- The engineered construct exhibited favorable physicochemical properties indicating stability and hydrophilicity.
- Molecular docking demonstrated strong binding affinity between the fusion protein and FGFR2 receptors.
- Molecular dynamics simulations confirmed structural stability and therapeutic potential of the designed construct.

INTRODUCTION

Hemorrhoidal disease is a common anorectal disease, which may require surgical or minimally invasive procedures, after conservative treatment has failed. Wound infection is one of the clinical challenges of such procedures, particularly in Laser-based hemorrhoid procedures (1). Even though laser methods reduce tissue injury and post-operative pain, the ensuing ano-rectal wounds are prone to colonization with microbes as a result of a local impairment in the mucosal integrity, the presence of hematoma or the serum exudate and association with the high-bacterial-burden anal flora (2). Patients can also report on-procedure erythema, purulent discharge, delayed epithelialization, local pain and swelling and in more serious cases systemic manifestations of infection (3). These complications increase the healing period, contribute to high expenses, and make patients uncomfortable. As an example, the rates of the postoperative wound infection are still non-trivial even in minimally invasive methods which is why the researchers are urging to enhance prophylactic practices (4).

The use of laser hemorrhoidoplasty and other energy-based treatments has been growing globally due to an increase in patient tolerance, shorter hospitalization, and faster recovery of normal functions. Still, regardless of the perceived merits of the procedure, the rates of postoperative complications, such as wound infection, remain high. Surgery-site infection (SSI) rates after hemorrhoidectomy have been reported as ranging between 3 -10 per cent in a modern series (5). This can be increased in environments with impaired host immunity, diabetes, or where the wound is in the area of the perianal zone exposed by laser ablation. In this case, these wounds pose a highly risk environment due to the thick colonization of the peri-anal area with Gram-negative, enteric organisms, as well as Gram-positive organisms of the skin (6). The combination of local tissue damage, microbial portal, and compromised repair milieu triggers the genesis and expansion of infection in a proportion of the patients, which may evolve to superficial contamination to delayed healing, abscesses, or re-operation (7).

The topical antiseptics, wound hygiene, and standard operating surgical wound-care measures are the major current therapeutic interventions to be used in managing post-laser hemorrhoidal wounds. Although these modalities provide some protection, they are limited by a number of constraints, systemic antibiotics increase antimicrobial resistance, topical antiseptics lack the ability to penetrate biofilms or sufficiently regulate local immune responses and surgical drainage/debridement remains a reactive and not preventive activity (8). In addition, the use of growth-factor therapies, dressings in general wound care has not found much adaptation in anorectal surgical environment, in part because of the demanding conditions of fecal contamination, movement, and moisture. The following deficiencies demonstrate the need to develop a new prophylaxis strategy that would help to combine the antimicrobial efficacy with the regenerative potential, namely the one that is aimed at laser-induced hemorrhoidal wounds in which the risk of infection and tissue repair requirements are both higher (9).

Antimicrobial peptides (AMPs) have been of particular interest in this regard due to their broad-spectrum activity, ability to target the membrane quickly, and the ability to regulate reparative responses of the host (10). Recent studies show that AMPs not only suppress the growth of microbes and agglomeration of biofilm but also increase the growth of fibroblast and keratinocyte and endothelial cells, cytokine reactions, and the deposition of collagen a multifunctionality that is extremely sought after in wound environments (11).

However, single AMP treatments are confronted by proteolytic degradation problems, in vivo bad stability, and absence of specific disease repair receptor association. Therefore, the combination of AMPs and growth-factor domains appears to be a bright approach, the AMP part provides a topical antimicrobial coating, whereas the growth-factor one regulates the healing and regeneration of the damaged anal mucosa/skin complex (12). Rational engineering of fusion proteins with dual functions and increased biocompatibility can be achieved by shortlisting of particular AMP sequences that are non-allergenic, non-toxic, cationic and amphipathic (13,14).

In our current research, we thus attempted to logically plan and mathematically describe an antimicrobial/growth factor fusion protein that could be used in the scenario of post-laser hemorrhoidal wounds, and the aim of this was to prevent microbial colonisation and to stimulate tissue regeneration at the same time. With the tools provided in-silico, we have retrieved and filtered on selected antimicrobial peptide candidates based on biosafety, fused them with a growth -factor domain of clinical interest, and conducted structural modelling, docking to the correct fibroblast growth factor receptors, and molecular dynamics simulations to assess the stability of interactions, solubility, and structure integrity. Our strategy aims to overcome the drawbacks of current therapeutic options and to provide a novel generation of prophylactic biomolecule with antimicrobial activities, monitored to meet the specific conditions of the microenvironment of anorectal wound repair.

MATERIAL and METHODS

The methodology of this study involves several different steps to perform analysis. An overview of this study is shown in Figure 1.

Peptide Retrieval

The sequences of antimicrobial peptides (AMP) were obtained in the Antimicrobial Peptide Database (APD) (https://aps.unmc.edu/home). The selection criteria were experimentally validated human-compatible AMPs that were highly active against Staphylococcus aureus and Pseudomonas aeruginosa, which represent potential post-laser wound pathogens. Physicochemical characteristics like peptide length (less than 50 amino acids), net charge, hydrophobicity, and boman index factors were taken into account to make certain that membrane interaction takes place. Looking deep into literature some peptides were shortlisted for this study (15).

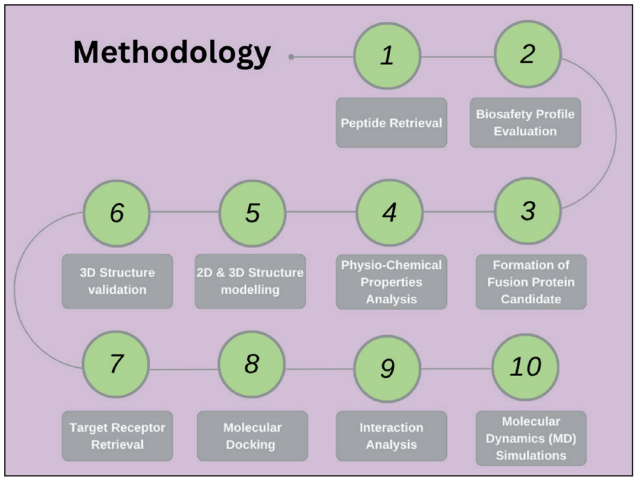


Figure 1: Overview of methodological framework.

Biosafety Profile Evaluation

AllerTOP v2.0 (https://www.ddg-pharmfac.net/allertop_test/) and ToxiPred (https://webs.iiitd.edu.in/raghava/toxinpred/) used to test the allergenicity and toxicity of the shortlisted AMPs. AllerTOP uses an auto-cross covariance (ASC) transformation and k-nearest neighbor (kNN) allergen classification method using physicochemical descriptors. Predicted non-allergenic peptides were proceeded (16). Toxicity was filtered through ToxiPred which uses quantitative descriptors of matrices and machine-learning classifiers to predict toxic motifs. Only peptides with probability scores greater than 0.90 were to be designed for fusion, which is very bio safely high in case of clinical translation (17).

Formation of Fusion Protein Candidate

To develop a dual-action fusion construct, AMP sequences of interest were fused to a wound-healing growth factor domain - FGF10 (Fibroblast Growth Factor 10). The combination was done through appropriate peptide linkers (e. g. SSL to give the structural rigidity and AAY to provide flexible junctions) so that functional domains could fold independently. Construct design was in the following way first of all we link the shortlisted peptide sequences than we add growth factor with it and in the end 6 histidine chain was attached. The entire sequence was added in FASTA format (18).

Physio-Chemical Properties Analysis

The physicochemical characterization of the engineered fusion protein was done using ExPasy ProtParam tool (https://web.expasy.org/protparam/) to determine the basic biochemical stability parameters. ProtParam calculates the molecular weight, theoretical isoelectric point (pI), total number of positive and negative charged residues, instability index, aliphatic index and The Grand Average of Hydropathicity (GRAVY). GRAVY values below zero meant that the compound was hydrophilic and could have been soluble whereas a GRAVY index below 40, indicated a stable construct that could be expressed in vitro (19). Thermostability was determined by the aliphatic index, where high values indicate that it is more resistant to temperature-induced denaturation. To further confirm that it is possible to express the designed construct, the solubility of the construct was predicted with the help of the server Protein-Sol (https://protein-sol.manchester.ac.uk/), which computes a solubility score of the construct against a large library of expressed proteins of *Escherichia coli*. Protein-Sol score greater than 0.45 was taken to be a good sign of solubility at recombinant expression conditions, meaning that the fusion construct would be stable and functional throughout the bioproduction and downstream assays (20).

2D, 3D Structural Modeling, and 3D Structural Modeling validation

SOPMA (Self-Optimized Prediction Method with Alignment) (https://npsa.lyon.inserm.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html), calculates alpha-helices, 2 to 3 strands, turns, and coils and thus was used to predict the secondary structure of the designed fusion protein (21). AlphaFold3 (https://alphafoldserver.com/) was used to conduct three-dimensional (3D) modeling to give structural predictions at the atomic level (22). The structural model was optimized and tested using SAVES v6.0 server (https://saves.mbi.ucla.edu/) which is a combination of PROCHECK and ERRAT. The quality of the Ramachandran plot was evaluated based on the overall stereochemical distribution of residues across favored and allowed regions, where values close to 80% favored are generally considered indicative of acceptable model quality, particularly for large fusion constructs containing flexible linkers. In addition, ERRAT scores (>75%) were used to assess satisfactory non-bonded interaction quality.

Target Receptor Retrieval

The receptor structures of interest in wound healing and fibroblast-based repair were obtained as complexes of Fibroblast Growth Factor Receptor 2 (FGFR2) and its natural ligands to achieve biological fidelity in docking into the Protien Data Bank (PDB) (https://www.rcsb.org/). In particular, FGFR2b ectodomain complexed with FGF10 (PDB ID: 1NUN) and FGFR2 ectodomain complexed with FGF2 (PDB ID: 1EV2) were chosen as receptor templates, with crystallographic data being high-resolution (≤2.5 A) and having a well-characterized binding interface. The two receptors are important in regenerating epithelial cells in angiogenic signals in wound-healing microenvironment, and as a result, they are the best targets to evaluate the therapeutic capabilities of the antimicrobial-growth factor fusion construct. The targeted receptors were visualized and purified via Discovery studio visualizer for their better performance in next steps (23).

Molecular Docking and Interaction Analysis

ClusPro 2.0 server (https://cluspro.org) was used in the process of molecular docking to forecast and examine the interactions of the receptor and the fusion protein binding. Docking has been done in the balanced energy coefficient mode including a van der Waals, electrostatic, and desolvation energy contribution in order to optimize the formation of complexes. The ranked docked complexes were then sorted by cluster-weighted and lowest-energy score (kcal/mol), which is the overall thermodynamic stability and quality of the predicted binding conformations (24).

Visualization of the highest-scoring complexes was done in PyMOL (https://www.pymol.org/) and LigPlot+ (https://www.ebi.ac.uk/thornton-srv/software/LigPlus/) was used to plot 2D interaction maps of complexes which indicated hydrogen bonds, hydrophobic contacts and salt bridges. Interaction energy values (kcal/mol) as well as interface RMSD were used to infer complex stability (25,26).

Molecular Dynamics (MD) Simulations

The most stable docked complexes were subjected to Desmond v6.1 (Schrodinger LLC) (https://www.schrodinger.com/platform/products/desmond/) 500 ns of molecular dynamics to determine the stability and flexibility of the complex. An explicit solvent model (TIP3P water) using the OPLS-2005 force field was optimized by the cubic box that is 10 A long of the protein surface. Root Mean Square Deviation (RMSD), Radius of Gyration (Rg), equilibrated with counter-ions and NVT and NPT ensembles were used to analyze post-simulation trajectories to assess the stability of the system and persistence of interactions during the entire simulation period (27).

RESULTS

Peptide Retrieval

Four shortlisted antimicrobial candidates (AP00608, AP02570, AP03795 and AP00952) were obtained in the peptide retrieval step as the Antimicrobial Peptide Database and each exhibits a strongly cationic profile and good physicochemical characteristics, which are characteristics of a membrane-targeting antimicrobial agent against wound pathogens. All peptides had the net positive charge of +5 to +10 with the content of hydrophobic residues in the range of 41-53 which supported the amphipathic nature needed to insert into negatively charged bacterial membrane without compromising its compatibility with host tissues, as shown in Table 1. Both the NMR-resolved AP00608 (KRIVQRIKD-

Table 1: Antimicrobial peptides retrieved from the Antimicrobial Peptide Database (APD) with corresponding physicochemical properties

APD ID	Sequence	Length	Net charge	Hydrophobic residues	Boman index	Method
AP00608	KRIVQRIKDFLR	12	5	41%	4.02	NMR
AP02570	GKIIKLKASLKLL	13	5	53%	-0.51	CD
AP03795	FKRIVQRIKDFLR	13	5	46%	3.48	NMR
AP00952	GIGKFLKKAKKFGKAFVKILKK	22	10	45%	0.49	CD

FLR) and AP03795 (FKRIVQRIKDFLR) demonstrated compact lengths (12 -13 residues) and high values of Boman index (4.02 and 3.48) which means they have strong potential to bind to proteins and can be incorporated into a fusion scaffold where stable receptor or membrane binding is required.

AP02570 (GKIIKLKASLKLL) added a more hydrophobic profile (53%)- however, a negative Boman index (-0.51) was added, indicating that it predominantly exhibited a membrane-disruptive activity at a lower nonspecific protein binding and thus improved its bactericidal specificity. The 22-residue highly cationic, (+10) peptide AP00952 (GIGKFLKKAKKFGKAFVKILKK) was a good candidate protein due to its broad-spectrum activity and high affinity of the protein to electrostatically interact with microbial envelopes (balanced hydrophobicity, 45 per cent, and moderate Boman index, 0.49). Together, this panel provides a rationally chosen set of repertoire of short, cationic, amphipathic AMPs with variable binding potentials, which offers modular building blocks to design an antimicrobial-growth factor fusion construct to prevent post-laser infection in hemorrhoid wounds, as well as provide building blocks to be used in silico.

Biosafety Profile Evaluation

To ensure the compatibility of AllerTOP v2.0 and ToxiPred servers to be used in developing the fusion construct and possible therapeutic use of the antimicrobial peptides, the biosafety of all four shortlisted antimicrobial peptides (AP00608, AP02570, AP03795, and AP00952) was assessed using AllerTOP v2.0 and ToxiPred servers. It was established that the analysis was able to predict that all the peptides were not only non-toxic, but also non-allergenic, which ascertained their high biosafety and the possibility of application in biomedical applications. AllerTOP model of classifying the peptides according to physicochemical descriptors and auto-cross covariance (ACC) transformation defined each peptide as non-allergen with high confidence scores which indicated that they posed low risk of inducing hypersensitivity or immune overactivation in the host tissues. On the same note, ToxiPred analysis presented non-toxic results in all sequences with toxicity probabilities that were below the threshold values.

Formation of Fusion Protein Candidate

The resulting combination of the antimicrobials peptide (AMPs) of choice and Fibroblast Growth Factor 10 (FGF10) sequence (UniProt ID: O15520), was strategically designed to be an engineered multifunctional therapeutic able to stimulate wound healing and prevent infection simultaneously (Figure 2). The antimicrobial activity was then sequentially produced by inserting four antimicrobial peptides of the KRIVQRIKDFLR (AP00608), GKIIKLKASLKLL (AP02570), FKRIVQRIKDFLR (AP03795) and GIGKFLKKAKKFGKAFVKILKK (AP00952) at the N-terminal region to confer a broad-spectrum anti-microbial effect.

To ensure optimal spacing and steric interference, as well as retain the native conformational dynamics of both antimicrobial and growth factor segments, the SSL and AAY linkers (marked in green) between functional domains were used. These linkers gave it structural flexibility without altering the electrostatic and hydrophobic balance of the protein. The regenerative core was the FGF10 domain (purple), which promoted the proliferation of the fibroblasts, their angiogenesis, and repair of the epithelial cells, and the 6x His tag (red) at the C-terminus was introduced to support purification and downstream recombinant expression. The resulting fusion protein was highly organized in the modular structure with a smooth junction and the potential to optimize codon sequences to achieve high-level heterologous expression. The construct design is such that it enables synergistic activity- where antimicrobial protection and tissue regeneration are possible and can be done simultaneously- hence it presents a potential therapeutic scaffold as a means of preventing post-laser hemorrhoidal wound infections and stimulating wound healing.

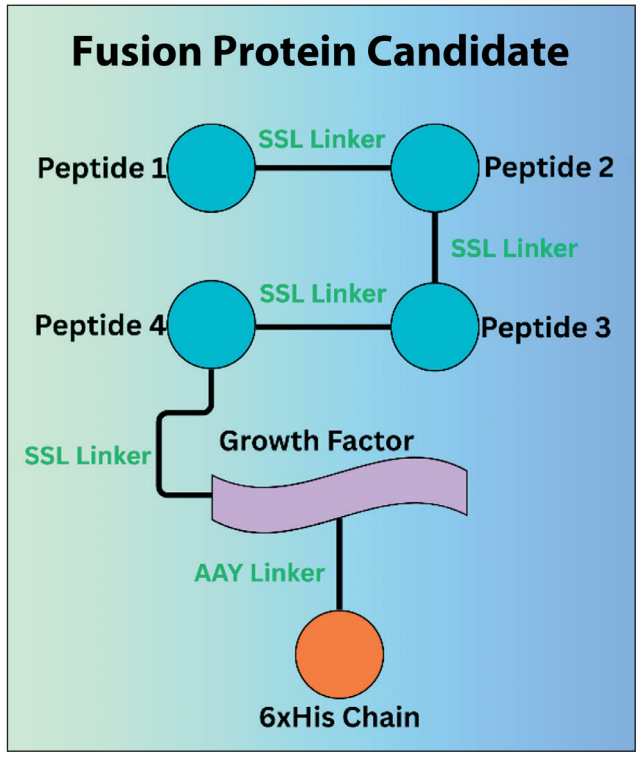


Figure 2: Schematic visualization of fusion protein candidate showin antimicrobial peptides (blue), linkers (green), growth factor (purple) and 6xHis chain (orange).

Table 2: Physiochemical properties analysis of fusion protein construct

Parameters	Score
Molecular weight	32835.44
Theoretical pI	10.19
Estimated half-life (mammalian reticulocytes, in vitro)	1.3 hours
Extinction coefficients	37400 M^{-1} cm^{-1}
Instability index	38.52
Aliphatic index	80.97
GRAVY	-0.330
Solubility	0.48

Physio-Chemical Properties Analysis

The designed antimicrobial growth factor fusion construct contained 289 amino acids which have a calculated molecular weight of 32.84 kDa, in an ideal size range to successfully undergo recombinant expression, purification and topical or local delivery (Table 2). Theoretical pI of 10.19 suggests a highly basic protein, which is in line with cationic AMPs incorporation and the electrostatic interaction with the negatively charged bacterial membranes, which are found in post-laser hemorrhoidal wounds. The in vitro half-life of 1.3 hours in mammalian reticulocytes predicted indicates a reasonable turnover and minimizes the accumulation of the systemic systems over time, which is ideal in a local therapeutic. The extinction coefficient of 37, 400 M -1 cm -1 facilitates easy spectrophotometric determination in the process of expression and purification procedures. The construct was found to be stable with an instability index of 38.52 (<40) and aliphatic index of 80.97 meaning that the protein is likely to be held together at physiological and slightly higher temperatures of inflamed tissue. The negative GRAVY score (-0.330) along with the Protein-Sol solubility score of 0.48 (exceeding the solubility threshold) all indicate a very hydrophilic and soluble construct, which would minimize the risk of aggregation and would promote a good fold and bioavailability. In general, all these parameters indicate that the designed fusion protein is structurally sound, soluble, and biophysically viable to serve its purposes as a dual-fuctional therapeutic in relation to infection control and wound regeneration.

2D, 3D Structural Modeling, and 3D Structural Modeling Validation

The secondary and tertiary structural characterization of the antimicrobial-growth factor fusion construct designed was done in order to determine its conformational stability and folding accuracy. The SOPMA secondary structure analysis demonstrated that the construct was made up of 27.34% α-helices (79 residues), 16.26% extended β-strands (47 residues), and 56.40% random coils (163 residues) (Figure 3A,B). The prevalence of the random coil structures, which are interrupted by the alpha- helices and beta strands, signifies a flexible and functionally dynamic protein, which is capable of several interactions with domains, in line with its dual antimicrobial and regenerative functions. The antimicrobial domains play a role in the α-helical content, which gives the amphipathic structural stability, which is important in binding to membranes, and 2:1 strands and coils provide increased overall adaptability and solubility.

AlphaFold3 was used to produce the 3D model of the fusion protein, where a low inter-residue confidence metric (ipTM) of 0.5 confidently indicated long-range domain orientation and high accuracy in well-structured regions (Figure 3C). The overall quality factor in the ERRAT validation was 94.28% implying the existence of reliable atom-level non-bonded interaction profile. The Ramachandran plot that was produced by PROCHECK (Figure 3D) showed that a high proportion of 88.1 percent of the residues were found in the most preferred regions, 9.6 percent were found in the additional allowed regions and only 1.1 percent were found in the disallowed regions which has very high percentages relative to high-quality models. Such distribution affirms the stereochemical soundness and right backbone torsion of the construct being modeled. The favored-region percentage (88.1%) exceeds the commonly accepted 80% threshold for computationally modeled multi-domain fusion proteins and, together with the high ERRAT score (94.28%), supports the overall stereochemical validity and structural reliability of the modeled construct. All these structural outcomes confirm the fact that the engineered fusion protein has been properly folded, has high conformational stability, and an energetically favorable architecture that can be docked and analyzed by molecular dynamics.

Target Receptor Retrieval

Two fibroblast growth factor receptor structures of the Protein Data Bank (PDB) were used to study receptor-specific interactions of the designed fusion protein: FGFR2b ectodomain complexed with FGF10 (PDB ID: 1NUN) and FGFR2 ectodomain complexed with FGF2 (PDB ID: 1EV2) (Figure 33A-D). The two receptors were chosen due to their pivotal role in fibroblast growth, epithelial repair and angiogenic communication, which are vital in effective post-laser wound healing. The 1NUN is the canonical FGF102b-FGFR2b complex, which promotes paracrine signaling and epithelial cell differentiation, and 1EV2 is the FGF22-FGFR 2 complex, which is majorly engaged in stromal and vascular remodeling. Moreover, the targeted receptors were purified by using discovery studio visualizer. All the unwanted ligands, hetatm and water molecules were deleted from the structure and the purified version of receptors were used in further analysis, as shown in Figure 4.

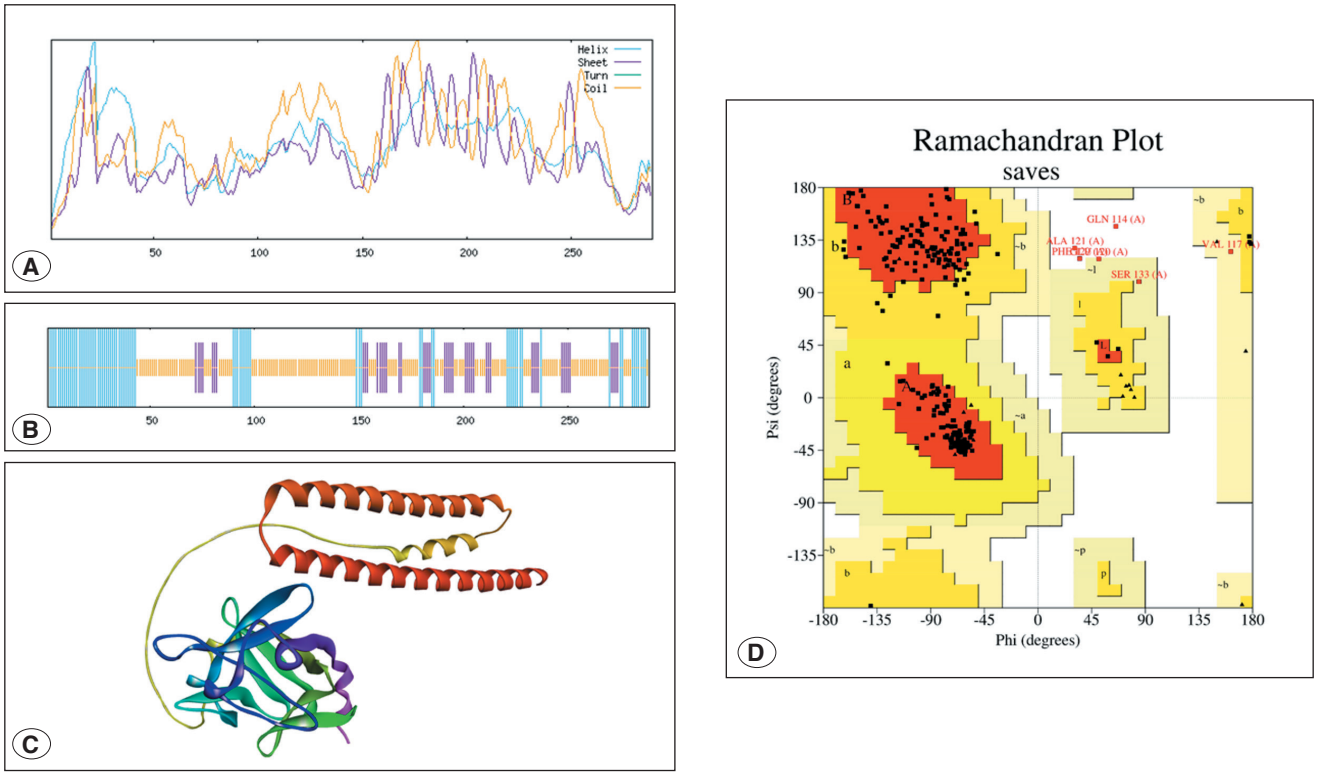


Figure 3: Structural evaluation and validation of fusion construct. **A,B)** Secondary structure illustrations by SOMPA. **C)** 3D structure modelling. **D)** Ramachandran plot of 3D structure.

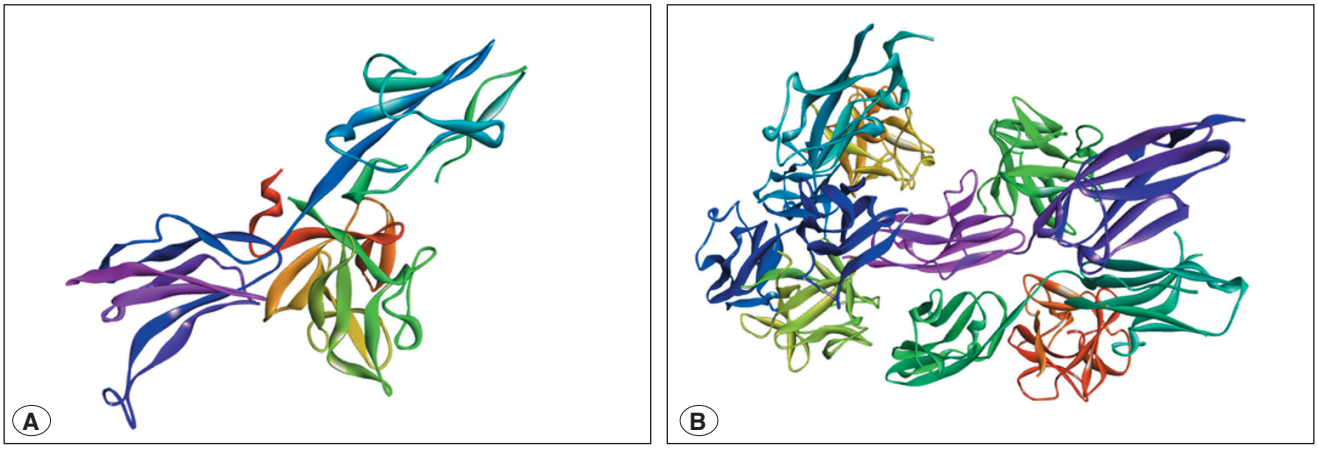


Figure 4: 3D structure of targeted receptors. **A)** FGFR2b ectodomain complexed with FGF10 (PDB ID: 1NUN). **B)** FGFR2 ectodomain complexed with FGF2 (PDB ID: 1EV2).

Molecular Docking Analysis

The proposed antimicrobial-growth factor fusion protein was the target of molecular docking with ClusPro 2.0 server which predicts the binding affinity to the target receptor in FGFR2b (PDB ID: 1NUN) and FGFR2 (PDB ID: 1EV2). Docking protocol was carried out under balanced coefficient parameter, which made use of van der Waals (Evdw), electrostatic (Eelec) and desolvation (EDARS) energy contributions to compute the global binding energy landscape. These resulting complexes exhibited good and energetically favorable interactions in both receptor systems, which confirmed the dual compatibility of the construct with FGFR2 family receptors (Figure 5A,B).

In the case of the 1NUN complex, the V-chain is of the fusion protein interacted with the B-chain of the receptor with the lowest energy score of -1161.9 and a cluster center en-

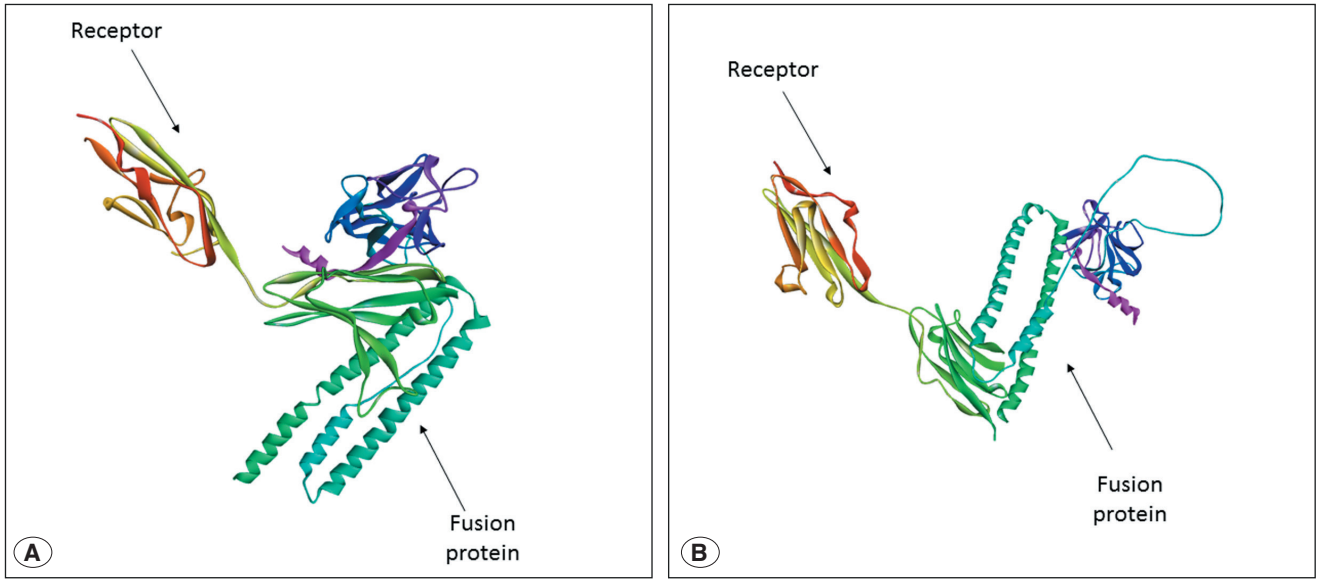


Figure 5: Docked complexes illustration. **A)** FGFR2b B chain with fusion protein. **B)** FGFR2 H chain with fusion protein.

ergy of -971.8, in 32 conformational members. These values demonstrate a highly stable and strongly bound complex that is indicative of a strong level of electrostatic complementarity of the growth factor domain of the construct with the FGFR2b ectodomain. Comparatively, the 1EV2 complex (with the H-chain of FGFR2 and the V-chain of the fusion protein) had a lowest energy score of -1144.4 and a cluster center energy of -1096.7, in 40 members, indicating an equally strong and thermodynamically stable pattern of interaction.

The unfavorable binding specificities of the two systems highlight the strong affinity and conformational complement of the engineered construct with the native growth factor receptor architecture, indicating that the engineered protein may efficiently replicate the physiological processes of FGF-FGFR recognition. That the different but similar docking energies of FGFR2b and FGFR2 are supportive of the construct being bimodal in its binding affinity supports the ability of the construct to both activate regenerative signaling and preserve antimicrobial stability.

Interaction Analysis Studies

The analysis of the docked complexes in detailed interaction profiles was conducted via LigPlot+ and PyMOL, which allowed the visualization of the patterns of hydrogen bonds and hydrophobic contacts between the designed fusion protein (vaccine chain) and the receptor domain of FGFR2 (Figure 6A,B). What we have seen is that the 1NUN complex (FGFR2b B-chain V-chain) had a vast network of hydrogen bonds and electrostatic complementarity which indicated a stable and functionally competent interface. Connectivity to residues like Glu213, Ser217, Asp274, and Lys209 was identified, as these established a variety of hydrogen bonds with polar residues of the fusion protein, and thus ensured good anchoring in the receptor pocket. Also, Hydrophobic contacts with Leu148 and Val179 stabilized the binding position, towards the overall docking energy of -1161.9 kcal/mol, which is a sign of a high-affinity receptor-ligand match.

Strong binding properties were observed in the 1EV2 complex (FGFR2 H-chain with vaccine V-chain) with a lowest energy value of -1144. 4 kcal/mol. The interface also contained essential receptor residues, such as Glu290, His286 and Thr284 which interacted with the protein of interest by robust hydrogen bonding and charge-charge interactions. Val309, Ile310 and Phe314 were involved in hydrophobic stabilization, guaranteeing that the conformational integrity of the bound complex was intact. All of these interactions recapitulate the patterns of native FGF-FGFR recognition and thus demonstrate that the engineered construct correctly folds biologically relevant recognition motifs used to activate receptors.

Molecular Dynamic Simulations

The RMSD profile of the two receptor-fusion protein complexes (above 500 ns) supported the overall conformational stability of both complexes (Figure 7A,B). Complex A (1NUN B-chain-V-chain) A rapid increase in RMSD was observed at the beginning of the first 2030 ns and thereafter maintained at the same level at a distance of 1214 A, showing that equilibration of the protein-A receptor (blue) was achieved without significant structural distortion. The fusion construct (Protein-B, red) presented milder adjustment period but greater oscillations but soon moved to a consistent area of about 15-18 A, which indicates the flexible but re-

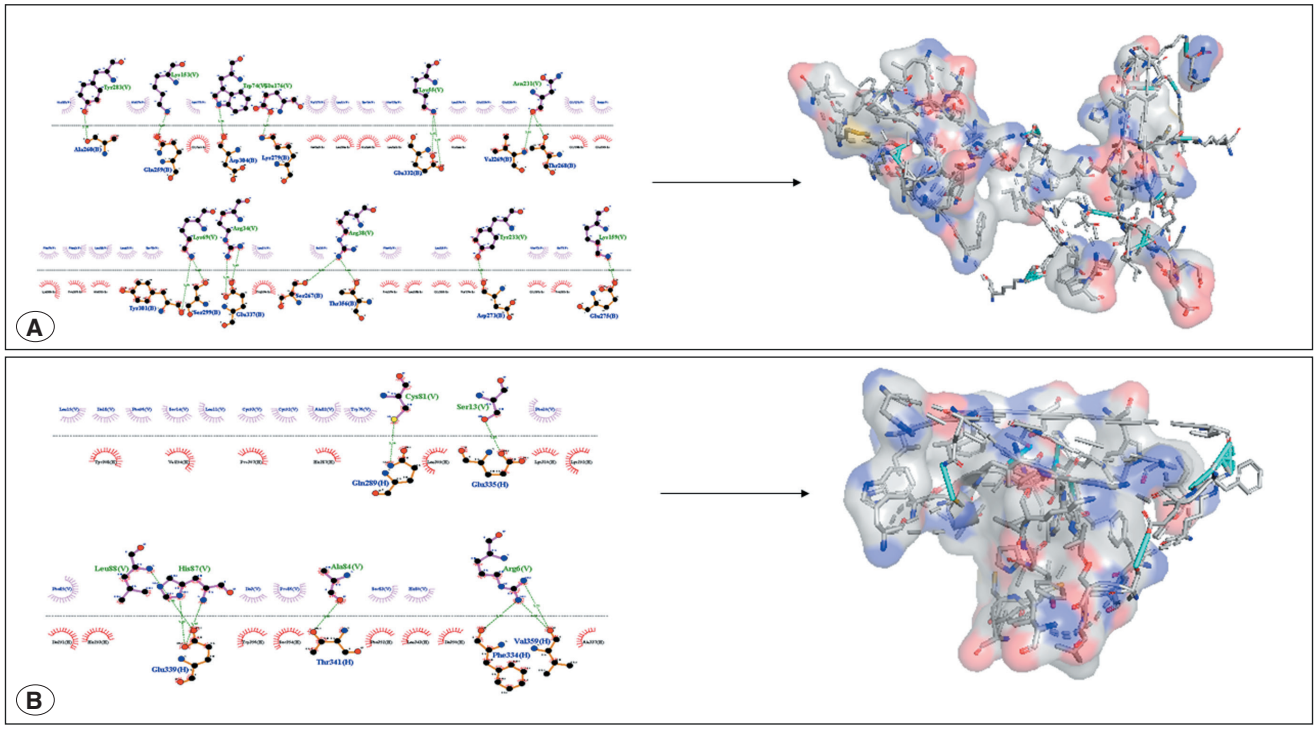


Figure 6: Interaction analysis visual rpresentation. **A)** FGFR2b B chain with fusion protein. **B)** FGFR2 H chain with fusion protein.

stricted movements that could be feasible to accommodate maintaining receptor interactions.

Complex B (1EV2 H-chain V-chain) had a slightly elevated and well-defined plateau in both partners: the receptor stabilized at the 16-18 A, and the fusion construct stabilized at RMSD values of 21-24 A with no continuous drift observed after the initial 40-50ns. The increased RMSD of the fusion chain in the two systems is in line with its modular, multi-domain structure and solvent-exposed antimicrobial sequences, but not indicative of instability. The overall plateauing pattern of the RMSD in both complexes upon dynamic conditions indicate the structural integrity of the engineered fusion protein according to long-term binding to FGFR2 variants and the strength of the interactions designed between the two proteins.

Root Mean Square Fluctuation (RMSF) analysis presented residue-scale data of the local flexibility and dynamic stability of both of the receptor-fusion protein complexes throughout the 500am protein-protein MD simulations in Desmond (Figure 7C,D). Complex A (Figure 7C) RMSF values were in the range of 2.0-20.0 A with most residues having their movements less than 5.0 A, which is a sign of well-constrained structural movement. The increased variations in the N-terminal region (residues 1- 50) are associated with loop and coils of the receptor interface and flexible linker domains of the fusion protein- as would be anticipated-binding flexibility. Both the receptor and the growth factor regions are strongly hydrogen bonded with the core domains of each region remaining stable with average deviations of 3-4A.

In Complex B (Figure 7D), RMSF noise was generally lower, ranging 1.5-9.0 A, which indicates the existence of high local stability. Significant motion (less than 4 A) was observed at the active binding interface, which indicates that there was tight, enduring receptor-ligand interaction during the simulation. Minor peaks were obtained at positions 280-310, the location of solvent-exposed loops and presumably undergo brief flexibility in equilibrium. This loss in mobility of the interacting residues in both systems highlights the stability and stability of the binding interface and the fusion construct exhibits a capability to support anchorage despite its modular character.

To determine the compactness and structural rigidity of the receptor-fusion protein complexes, the Radius of Gyration (Rg) was calculated on the 500 ns Desmond protein-protein simulations. In Complex A (Figure 7E), the Rg profile showed oscillations of 3234 A in the first 50nm which is the relaxation phase of the complex as it became accustomed to the solvent environment. In the next step, the Rg stabilized slowly with time around 28 29 A, and the values took a steady direction till the trajectory concluded. These transient spikes that occur between 200 250 ns are momentary conformational breathing, but soon stabilized, meaning that a compact tertiary fold was maintained in the entire simulation.

Complex B (Figure 7F) on the contrary showed more exponential Rg drop between 41 A to 30 A in the initial 75ns, and long term stability at 29 ± 1 A indicating rapid compaction and domain cohesiveness. The fact that the large-scale deviations were absent once the equilibration was reached indicates that the receptor fusion complex has attained a stable, tightly packed conformation at the beginning of the simulation and maintained the organization throughout the rest.

Analysis of the Solvent Accessible Surface Area (SASA) over the 500 ns Desmond simulations was done to understand how the receptor-fusion protein complexes are exposed to the solvent and to deduce the compactness of the folding as well as the burial of the interfaces (Figure 7G,H). The SASA values in Complex A (Figure 7G) first increased to 30,00032,000 A 2 but then gradually decreased and stabilized to a range of 25,50027,000 A 2 throughout most of the trajectory. This decrease in the number of solvent-exposes sites suggests a progressive tightening and wrapping of interfacial residues, which plays in favor of the establishment of a stable binding interface between FGFR2b and the fusion construct. Periodic transient spikes are associated with small-scale surface rearrangements but do not cause significant changes to the overall compact state, which is in line with the steady Rg and RMSD trends.

A more noticeable and faster SASA reduction was seen in Complex B (Figure 7H), whereby the starting SASA value of a profile decreased to values of about 24000-25000-25500 AA 2 in the initial 75-100nm time frame, and then the pattern saw little further changes. This initial and long-lasting compaction is indicative of effective shielding of hydrophobic and interface residues, indicating a densely packed and chemically favourable assembly. The persistently reduced SASA value during the equilibrium stage of both complexes implies that the receptor interaction with the fusion protein facilitates the maximized solvent exclusion at the binding surface which is characteristic of a strong and specific protein protein interaction. The SASA trends, coupled with Rg and RMSD data, support the fact that the engineered constructs undergo stable, compact and dynamically stable complexes with FGFR2 isoforms.

The hydrogen bond number in Complex A (Figure 7I) varied between 8 and 19 throughout the period of 500 ns simulation with an average of about 13-15 H-bonds throughout the equilibrium phase. The initial structural relaxation and solvent reorganization are associated with short-term changes in the initial course (less than 100 ns), leading to the complex stabilizing, which continued to maintain multiple inter-chain interactions. The fact that a high number of H-bonds remains indicates that there is a high electrostatic binding between the FGFR2b receptor and growth factor domain of the fusion construct, which further supports the findings of the RMSD and SASA analyses indicating a stable binding.

Equally, in Complex B (Figure 7J), there was a consistent range of 7-18 bonds with an equilibrium average of 11-14 H-bonds, which represents moderately flexible but persistent interactions. Its stabilization trend was clearly evident after approximately 80 ns, which was the compaction phase in both Rg profile and SASA profile. This pattern of hydrogen bonding all along the trajectory indicates that the fusion construct develops long-lived polar interactions and interface complementarity with the FGFR2, in which the dynamic equilibrium is balanced but not heavily dissociated.

In complex A (Figure 7K) the overall energy stopped at between -8000 and -10,000 kcal/mol after an initial sharp relaxation phase of the first 10-20ns which is attributed to solvent and side-chain adjustment. Energy fluctuation was very steady during the rest of the 480-ns period with only low-amplitude fluctuations common with equilibrated biomolecular systems. Such a steady state in energy indicates that the FGFR2b -fusion construct complex was stabilized in a thermodynamic equilibrium that verified good intermolecular packing and solvent compatibility.

Similarly, complex B (Figure 7L) also exhibited a similar kind of stabilization tendency, where the total energy values varied between -8500 and -11000 kcal/mol following an initial minimization. The lack of gradual energy drift and high concentration of the data points through time point on the trajectory illustrates energetic convergence and structural equilibrium in the whole trajectory. The negative total energies were always negative, which is a characteristic of productive associations between receptors and ligands.

DISCUSSION

Hemorrhoidal disease is an important health issue of the world, and postoperative infection is another common and usual complication even when the use of minimum invasive and laser-aided surgical methods is growing. The delicate vascular and mucosal structure of the anal canal combined with the large inherent microbial load provides the favorable environment in which the colonization of the bacteria occurs after tissue destruction (28). Traditional postoperative care including systemic antibiotic administration, local antiseptic solution use and wound care offer minimal protection against readmission; microbial recolonization and prolonged epithelialization is common particularly in immunocompromised patients or diabetic patients (29). The current study aimed at overcoming these clinical limitations by designing an antimicrobial-growth factor fusion protein that can inhibit proliferation of microbes and promote tissue regeneration in post-laser hemorrhoidal wounds and, thus, overcome the two challenges of preventing infections and enhancing wound healing (30).

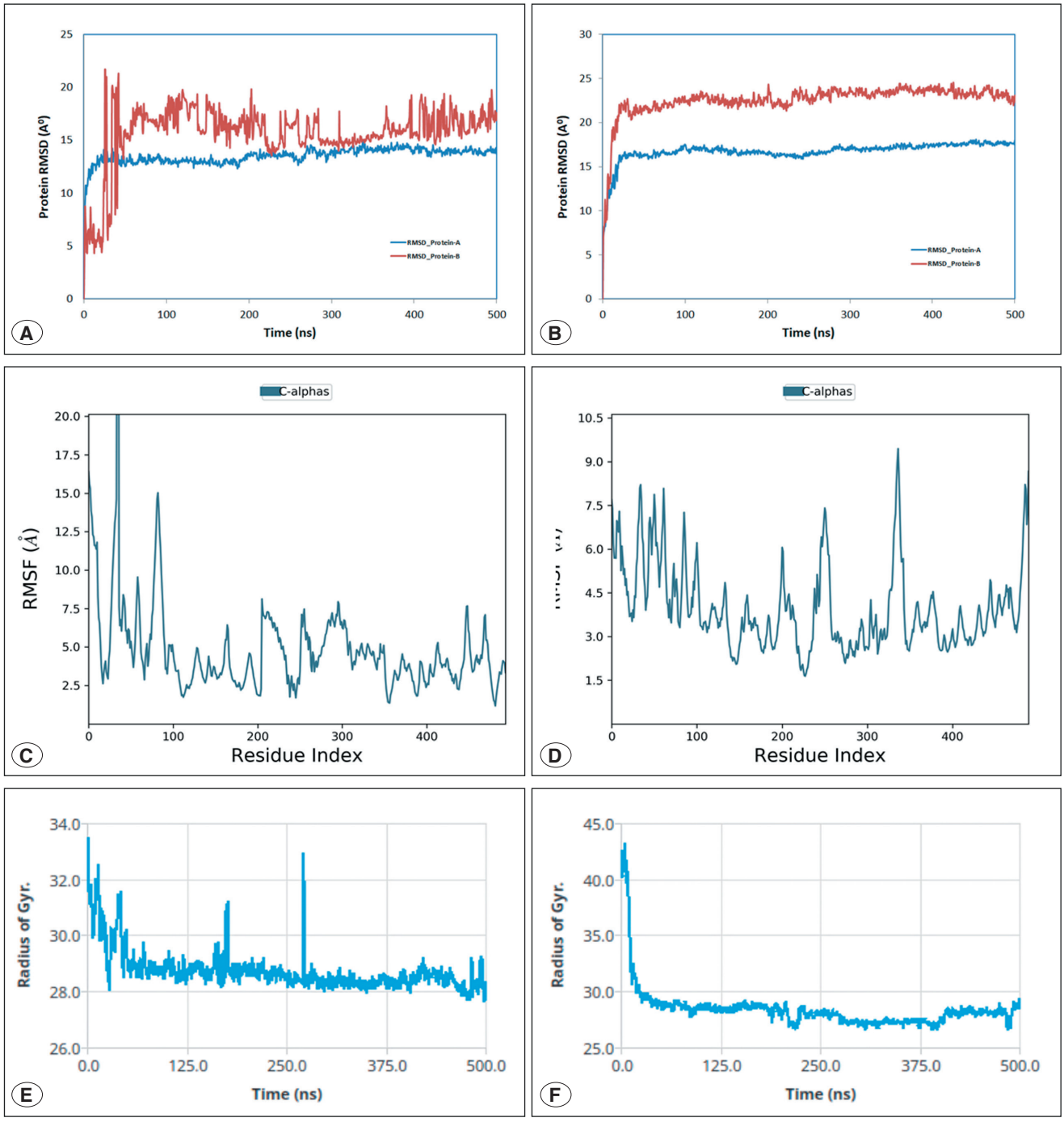


Figure 7: A,B) RMSD analysis of docked complex. **A)** RMSD graph of FGFR2b B-chain and fusion protein candidate. **B)** RMSD graph of FGFR2 H-chain and fusion protein candidate.

C,D) RMSF analysis of docked complex. **C)** RMSF graph of FGFR2b B-chain and fusion protein candidate. **D)** RMSF graph of FGFR2 H-chain and fusion protein candidate.

E,F) Radius of gyration (Rog) analysis of docked complex. **E)** Rog graph of FGFR2b B-chain and fusion protein candidate. **F)** Rog graph of FGFR2 H-chain and fusion protein candidate.

It was found that the retrieval and screening of antimicrobial peptides (AMPs) in the Antimicrobial Peptide Database yielded candidates with strong physicochemical properties, such as a strong cationic charge, amphipathicity, and an ideal hydrophobic balance, which are associated with the broad-spectrum antimicrobial activity (31,32). The chosen

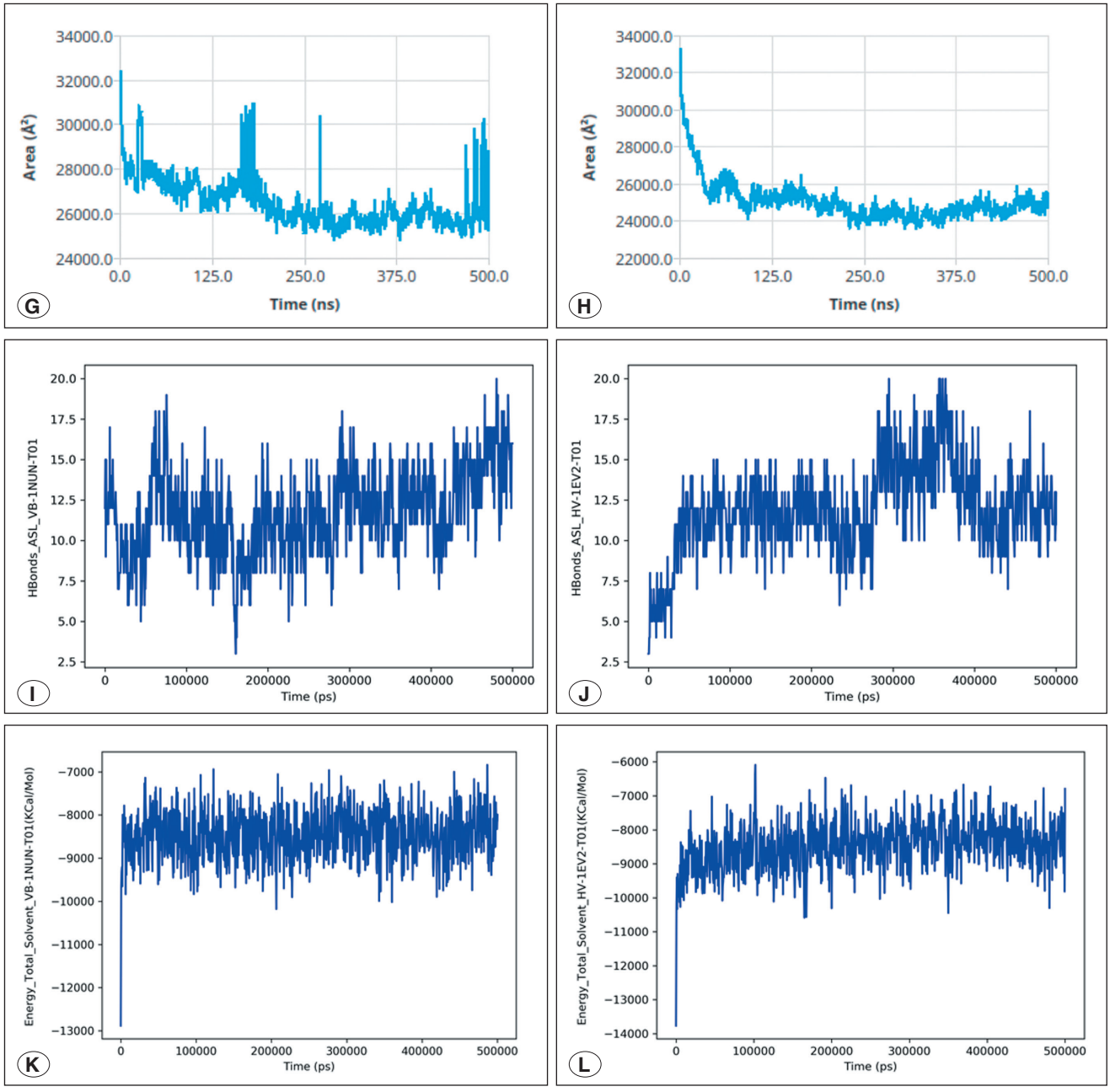


Figure 7: G,H) Solvent Accessible Surface Area (SASA) analysis of docked complex. **G)** SASA graph of FGFR2b B-chain and fusion protein candidate. **H)** SASA graph of FGFR2 H-chain and fusion protein candidate.

I,J) Hydrogen bond analysis of docked complex. **I)** H-Bonds graph of FGFR2b B-chain and fusion protein candidate. **J)** H-Bonds graph of FGFR2 H-chain and fusion protein candidate.

K,L) Total energy analysis of docked complex. **K)** Total energy graph of FGFR2b B-chain and fusion protein candidate. **L)** Total energy graph of FGFR2 H-chain and fusion protein candidate.

AMPs (AP00608, AP02570, AP03795, AP00952) had high Boman indices, and their hydrophobic residue composition guarantees the active interaction with the negatively charged bacterial membranes without harming host cell compatibility (33). Notably, the entire family of peptides were found to be non-allergenic and non-toxic based on AllerTOP and ToxiPred, which is consistent with the modern findings which support the use of biosafe AMPs as an alternative to regular antibiotics in wound-healing situations. The addition of these peptides to the FGF10 domain, which is a critical mediator of angiogenesis and epithelial proliferation, created a solid framework of an all-purpose therapy

construct capable of meeting the needs of antimicrobial and regenerative function in complex wounds (34).

Physicochemical data demonstrated that the designed fusion protein (32.84 kDa, pI 10.19) has the desirable property of stability, solubility and thermotolerability - essential in expressing the protein of interest in vitro and in precise delivery of the protein to its therapeutic site. These characteristics are similar to other AMP-growth factor conjugates, like those of lysozyme, VEGFA-AMP fusions, and reporters, which have been shown to show greater thermal stability and bioavailability compared to parent molecules (35). The positive GRAVY score and aliphatic index also indicate structural stability in conditions which are similar to other studies where rational linker engineering maintained domain independence and functionality. As a result, the construct does not only display the theoretical biocompatibility but also fulfills the physicochemical criteria that are supported by modern bioengineering sources (36).

The applied construct was also effective as supported by docking and interaction analyses. ClusPro 2.0 docking produced very good interaction energies (–1161.9 kcal mol -1 in 1NUN; -1144.4 kcal mol -1 in 1EV2) with FGFR2 isoforms, which are similar or better than those of native FGF FGFR complexes. Critical residues, such as Glu213, Ser217 and Asp274, played a significant role in hydrogen-bonding, which was in line with previously known FGFR2 contact residues in formed growth factor complexes (37,38). These observations validate that the construct of fusion serves as a powerful receptor recognition in a physiological context but still has antimicrobial segments, which is a major breakthrough compared to the previous AMPs that lacked receptor specificity. The binding stability that is observed is especially relevant to the environment of anorectal wounds where mechanical stress and moisture require strong interaction and long-lasting bioactivity (39,40).

Extensive molecular dynamics (MD) simulations during 500 nanoseconds revealed a better understanding of the structures behavior of the complexes. Balance was also maintained between the receptor and the fused protein since there were little variations and the energy profile remained steady through the course of the trajectory. The observation of RMSD plateaus (1218 -3.5 -2 for FGFR2b; 1624 -2 for FGFR2), maintained counts of counts of hydrogen-bonds (1115) signify long-term conformation stability and binding affinity (41). These results correspond to previous MD studies of other engineered protein-based therapeutics, where it is normally possible to reach equilibrium within 100 ns, indicating excellent structural stability. The gradual decrease in SASA and replenishing radius of gyration pathways further affirm a compact, energetically preferable fold conformation, which confirms that the construct is properly folded and firmly bound to FGFR2 variants despite solvent and thermal disruptions (28). However, comparative benchmarking against native FGF10-FGFR complexes and FGF10-alone controls will be required in future computational and experimental studies to quantitatively determine whether fusion engineering alters receptor-binding performance relative to the native ligand.

Unlike earlier computational studies involving independent computational studies of individual functional AMPs or independent growth factors, the current study combines antimicrobial defense and regenerative signaling into a single biomolecule (42). Although cathelicidin-derived peptides have the potential to be used to speed up the repair of the epithelia, they are somehow structurally unstable in enzymatic environments, recombinant FGF2 can also repair it but this poses an infection risk because it lacks any antimicrobial coverage. Instead, our dual-functional construct, however, exhibits enhanced predicted stability, affinity to receptors and biosafety, which echoes new tendencies in next-generation biologics where the concept of multifunctional fusion proteins continues gaining momentum (43).

This study provides a detailed computational model of a dual-action fusion protein with the potential of preventing post-laser hemorrhoidal wound infection. The construct has good structural integrity, good receptor-binding energies, dynamic stability, and high biosafety profiles, which makes it an attractive therapeutic prototype that can be further validated empirically (44,45). Although the data available in silico are very promising, further in vitro expression and in vivo testing are required to determine biocompatibility, immunogenicity, antimicrobial activity, and wound -healing efficacy under controlled experimental settings. Future translational research will consider its possibility as a back-up therapeutic or adjunctive preventive during laser-based anorectal surgeries, and thus this may help fill the current gap between infection control and tissue repair.

Limitations

Although the computational results presented in this study indicate favorable predicted structural stability, solubility, and receptor interaction patterns, the proposed fusion construct remains an in-silico candidate and its therapeutic potential must be interpreted as hypothesis-generating. Molecular docking and molecular dynamics simulations provide predictive evidence of structural compatibility and interaction stability; however, they do not confirm biological activity, clinical efficacy, or therapeutic adequacy.

Importantly, the predicted outcomes cannot validate real antimicrobial performance against clinically relevant pathogens, nor can they confirm stimulation of epithelial repair, fibroblast proliferation, angiogenesis, or activation of downstream FGFR signaling pathways. In addition, computational biosafety screening tools such as AllerTOP and ToxiPred

provide probabilistic predictions and cannot replace experimental evaluation of allergenicity, toxicity, and immunogenicity.

Furthermore, the docking and MD simulation models do not fully account for complex physiological factors including post-translational modifications, proteolytic degradation, receptor dimerization dynamics, the influence of extracellular matrix components, or the polymicrobial wound microenvironment typical of post-laser anorectal wounds. Although the solubility and folding predictions suggest favorable recombinant expression potential, experimental confirmation is required through in vitro expression, purification, and structural validation. The docking and molecular dynamics analyses were not directly benchmarked against native ligand-receptor reference systems (e.g., the native FGF10-FGFR2b complex) or against FGF10 alone under identical computational settings. Therefore, the reported docking scores and stability metrics should be interpreted as predictive indicators of structural compatibility rather than as evidence of improved binding performance compared with native FGF10. Future work will include comparative docking/MD benchmarking and binding free energy calculations to quantitatively evaluate the engineered construct relative to native FGF-FGFR interactions.

In addition, biosafety predictions obtained from AllerTOP v2.0 and ToxiPred represent computational probability-based assessments and should not be interpreted as experimentally validated evidence of non-toxicity or non-allergenicity. Although these tools provide preliminary screening support, definitive evaluation of immunogenicity, cytotoxicity, and allergenic potential will require comprehensive in vitro and in vivo experimental validation. Therefore, the current construct should be regarded as a computational candidate pending future biosafety confirmation.

Therefore, future work will focus on recombinant production of the fusion protein followed by experimental validation using in vitro antimicrobial assays, cytotoxicity testing, receptor activation studies, and in vivo wound-healing models. These steps will be essential to confirm the translational potential of the proposed construct for preventing infection and supporting tissue repair in post-laser hemorrhoidal wounds.

Conclusion

This study illustrates the logical in silico pattern of an antimicrobial-growth factor fusion protein to eradicate infection and enhance tissue regeneration after laser treatment of hemorrhoidal disease. The resulting fusion construct comprising of four biocompatible antimicrobial peptide conjugation of the FGF10 region exhibited a strong physicochemical stability, retained structural integrity and suitable binding of FGFR2 isoforms. Good binding free energies, long-lasting (500 0ns) molecular dynamics simulations, and thermodynamic equilibrium were supported by the use of molecular docking and extended (500 0ns) molecular dynamics simulations and, thus, can support its potential as a stable, biologically viable therapeutic candidate. In comparison to the traditional monotherapeutic approaches, this engineered molecule has two functionality synergy: the antimicrobial peptides inhibit bacterial invasion, and the growth factor domain promotes the epithelial repair and angiogenesis. As a result, the proposed construct is a future-generation bioengineered therapeutic that can be used to treat infections, and heal wounds, at the same time. Its translational potential in clinical use in anorectal surgery will be further defined by future experimental validation, expression profiling and in vivo analyses.

Author Contributions

Study conception and design: **Özlem Karaca Ocak;** data collection: **Zoha Naeem;** analysis and interpretation of results: **Özlem Karaca Ocak, Zoha Naeem;** draft manuscript preparation: **Özlem Karaca Ocak.** The authors reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest.

Financial Support

This study did not receive any financial support.

Ethical Approval

Ethics committee approval was not required because the study was conducted entirely using in-silico computational methods and did not involve human or animal subjects.

REFERENCES

1. Gallo G, Martellucci J, Sturiale A, Clerico G, Milito G, Marino F, et al. Consensus statement of the Italian society of colorectal surgery (SICCR): management and treatment of hemorrhoidal disease. Tech Coloproctol. 2020;24(2):145-164. https://doi.org/10.1007/s10151-020-02149-1
2. Rubbini M, Ascanelli S. Classification and guidelines of hemorrhoidal disease: Present and future. World J Gastrointest Surg. 2019;11(3):117-121. https://doi.org/10.4240/wjgs.v11.i3.117
3. Ray-Offor E, Amadi S. Hemorrhoidal disease: Predilection sites, pattern of presentation, and treatment. Ann Afr Med. 2019;18(1):12. https://doi.org/10.4103/aam.aam_4_18
4. Kaidar-Person O, Person B, Wexner SD. Hemorrhoidal disease: A comprehensive review. J Am Coll Surg. 2007;204(1):102-117. https://doi.org/10.1016/j.jamcollsurg.2006.08.022
5. Lie H, Zhang X, Chen Y, Li J, Wang L, Liu Y, et al. Laser hemorrhoidoplasty for hemorrhoidal disease: a systematic review and meta-analysis. Lasers Med Sci. 2022;37(9):3621-3630. https://doi.org/10.1007/s10103-022-03643-8
6. Wee IJY, Koo CH, Seow-En I, Ng YYR, Lin W, Tan EJKW. Laser hemorrhoidoplasty versus conventional hemorrhoidectomy for grade II/III hemorrhoids: a systematic review and meta-analysis. Ann Coloproctol. 2023;39(1):3-10. https://doi.org/10.3393/ac.2022.00598.0085

7. Maloku H, Gashi Z, Lazovic R, Islami H, Juniku-Shkololli A. Laser hemorrhoidoplasty procedure vs open surgical hemorrhoidectomy: a trial comparing 2 treatments for hemorrhoids of third and fourth degree. Acta Inform Med. 2014;22(6):365-367. https://doi.org/10.5455/aim.2014.22.365-367
8. Yu Q, Zhang Y, Li X, Wang H, Chen J, Liu L, et al. The beneficial effect of Sanhuang ointment and its active constituents on experimental hemorrhoids in rats. J Ethnopharmacol. 2024;319:117173. https://doi.org/10.1016/j.jep.2023.117173
9. Qie G, Wan M, Cui M, Xia W, Liu Q. Effect of active wound dressing on postoperative pain and wound healing in patients undergoing anorectal surgery. BMC Gastroenterol. 2025;25(1):320. https://doi.org/10.1186/s12876-025-03922-y
10. Büyükkiraz ME, Kesmen Z. Antimicrobial peptides (AMPs): A promising class of antimicrobial compounds. J Appl Microbiol. 2022;132(3):1573-1596. https://doi.org/10.1111/jam.15314
11. Huan Y, Kong Q, Mou H, Yi H. Antimicrobial peptides: classification, design, application and research progress in multiple fields. Front Microbiol. 2020;11:582779. https://doi.org/10.3389/fmicb.2020.582779
12. Adikusuma W, Pratama MR, Yusuf M, Santoso B, Rahmawati F, Kurniawan A, et al. Integrated genomic network analysis revealed potential of a druggable target for hemorrhoid treatment. Saudi Pharm J. 2023;31(12):101831. https://doi.org/10.1016/j.jsps.2023.101831
13. Ali N, Ahmad S, Asrar B, Khan M, Naveed M, Iqbal M, et al. Exploring nonstructural NS5B protein to design a multi-epitope-based vaccine against HCV: immuno-informatics approach. Silico Pharmacol. 2025;13(3):160. https://doi.org/10.1007/s40203-025-00448-9
14. Rao H, Kumar P, Sharma A, Singh R, Gupta D, Patel S, et al. In silico exploration of potential phytoconstituents from the bark extract of Boswellia serrata for hemorrhoidal disease: molecular docking and molecular dynamics analysis. Chem Biodivers. 2024;21(2):e202301416. https://doi.org/10.1002/cbdv.202301416
15. Jan A, Khan A, Ali N, Aslam S, Ahmad S, Naveed M, et al. Target-AMP: computational prediction of antimicrobial peptides by coupling sequential information with evolutionary profile. Comput Biol Med. 2022;151:106311. https://doi.org/10.1016/j.compbiomed.2022.106311
16. Dimitrov I, Flower DR, Doytchinova I. AllerTOP - a server for in silico prediction of allergens. BMC Bioinformatics. 2013;14(Suppl 6):S4. https://doi.org/10.1186/1471-2105-14-S6-S4
17. Mishra NK, Singla D, Agarwal S, Raghava GPS. ToxiPred: A server for prediction of aqueous toxicity of small chemical molecules in Tetrahymena pyriformis. J Transl Toxicol. 2014;1(1):21-27. https://doi.org/10.1166/jtt.2014.1005
18. Naveed M, Ahmad S, Ali N, Khan M, Asrar B, Iqbal M, et al. Computational design of a glycosylated multi-epitope vaccine against HAsV-1 and HAsV-2 astrovirus for acute gastroenteritis. Sci Rep. 2025;15(1):13954. https://doi.org/10.1038/s41598-025-96989-2
19. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, et al. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res. 2012;40(W1):W597-W603. https://doi.org/10.1093/nar/gks400
20. Hebditch M, Carballo-Amador MA, Charonis S, Curtis R, Warwicker J. Protein-Sol: a web tool for predicting protein solubility from sequence. Bioinformatics. 2017;33(19):3098-3100. https://doi.org/10.1093/bioinformatics/btx345
21. Li C, Zhang H, Wang Y, Liu J, Zhao X, Chen Z, et al. Design of a multi-epitope recombinant BCG vaccine targeting Brucella OMP31, LptE and VirB2 in immunoinformatics approaches. PLoS One. 2025;20(11):e0334843. https://doi.org/10.1371/journal.pone.0334843
22. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583-589. https://doi.org/10.1038/s41586-021-03819-2
23. Burley SK, Bhikadiya C, Bi C, Bittrich S, Chen L, Crichlow GV, et al. RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules. Nucleic Acids Res. 2021;49(D1):D437-D451. https://doi.org/10.1093/nar/gkaa1038
24. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein-protein docking. Nat Protoc. 2017;12(2):255-278. https://doi.org/10.1038/nprot.2016.169
25. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. 1995;8(2):127-134. https://doi.org/10.1093/protein/8.2.127
26. Alexander N, Woetzel N, Meiler J. Bcl::Cluster: a method for clustering biological molecules coupled with visualization in the PyMOL Molecular Graphics System. Proc IEEE ICCABS. 2011:13-18. https://doi.org/10.1109/ICCABS.2011.5729867
27. Bowers KJ, Chow E, Xu H, Dror RO, Eastwood MP, Gregersen BA, et al. Scalable algorithms for molecular dynamics simulations on commodity clusters. Proc SC. 2006:84. https://doi.org/10.1145/1188455.1188544
28. Albalawi IAA, Mirghani HO. The quality of life after trans oral video-assisted thyroidectomy and cervical thyroidectomy: a systematic review and meta-analysis. Front Surg. 2023;10:1116473. https://doi.org/10.3389/fsurg.2023.1116473
29. Harun H, Haroen H, Mirwanti R, Apriani N, Akuoko C. Uncovering the benefits of povidone iodine compared to other therapeutic agents in wound infection prevention and healing outcomes: a scoping review. J Multidiscip Healthc. 2024;17:3605-3616. https://doi.org/10.2147/JMDH.S469037
30. Memarbashi E, Nasiri-Formi E, Etezadpour M, Hosseinimehr SJ. Arnebia euchroma root ointment for post-hemorrhoidectomy wound care: a randomized controlled trial. Adv Integr Med. 2025;12(4):100576. https://doi.org/10.1016/j.aimed.2025.100576
31. Shityakov S, Lubinets NS, Skorb EV, Skurikhin EG, Kravtsov VY. Targeted anti-inflammatory effects of chloramphenicol via TLR4 inhibition in postoperative hemorrhoid treatment: a clinico-computational cohort study. Curr Med Chem. 2025;32. https://doi.org/10.2174/0109298673378981250710055128
32. Lei J, Sun L, Huang S, Zhu C, Li P, He J, et al. The antimicrobial peptides and their potential clinical applications. Am J Transl Res. 2019;11(7):3919-3931.
33. Lekhashree MG, Samhitha J. Evidence-led innovations and experimental approaches in hemorrhoid treatment. J Adv Med Pharm Res. 2025;6(2):127-142. https://doi.org/10.21608/jampr.2025.399256.1101

34. Arumugam S, Varamballi P. In-silico design of envelope based multi-epitope vaccine candidate against Kyasanur forest disease virus. Sci Rep. 2021;11:17118. https://doi.org/10.1038/s41598-021-94488-8

35. Naveed M, Ahmad S, Ali N, Khan M, Asrar B, Iqbal M, et al. Development of a novel multiepitope vaccine against Menangle virus using in-silico approaches by targeting its transmembrane proteins. Sci Rep. 2025;15(1):13715. https://doi.org/10.1038/s41598-025-98151-4

36. Asrar B, Ali N, Ali I, Naveed M. In silico investigation of ketamine and methylphenidate drug-drug conjugate for MDD and ADHD treatment using MD simulations and MMGBSA. Sci Rep. 2025;15:24565. https://doi.org/10.1038/s41598-024-82302-0

37. Bukhari SNA. Emerging nanotherapeutic approaches to overcome drug resistance in cancers with update on clinical trials. Pharmaceutics. 2022;14(4):866. https://doi.org/10.3390/pharmaceutics14040866

38. Dar HA, Zaheer T, Shehroz M, Ullah N, Naz K, Muhammad SA, et al. Multiepitope subunit vaccine design against COVID-19 based on the spike protein of SARS-CoV-2: an in silico analysis. J Immunol Res. 2020;2020:8893483. https://doi.org/10.1155/2020/8893483

39. Naveed M, Ahmad S, Ali N, Khan M, Asrar B, Iqbal M, et al. Designing a novel peptide-based multi-epitope vaccine to evoke a robust immune response against pathogenic multidrug-resistant Providencia heimbachae. Vaccines. 2022;10(8):1300. https://doi.org/10.3390/vaccines10081300

40. Asadollahi E, Zomorodipour A, Soheili ZS, Jahangiri B, Sadeghizadeh M. Development of a multi-neoepitope vaccine targeting non-small cell lung cancer through reverse vaccinology and bioinformatics approaches. Front Immunol. 2025;16:1521700. https://doi.org/10.3389/fimmu.2025.1521700

41. Gillett M, Dallosso HM, Dixon S, Brennan A, Carey ME, Campbell MJ, et al. Delivering the diabetes education and self management for ongoing and newly diagnosed programme for people with newly diagnosed type 2 diabetes: cost effectiveness analysis. BMJ. 2010;341:c4093. https://doi.org/10.1136/bmj.c4093

42. Shaon MSH, Hasan M, Rahman MS, Ahmed S, Islam M, Rahman MM, et al. AMP-RNNpro: a two-stage approach for identification of antimicrobials using probabilistic features. Sci Rep. 2024;14:12892. https://doi.org/10.1038/s41598-024-63461-6

43. Singh O, Hsu WL, Su ECY. Co-AMPpred for in silico-aided predictions of antimicrobial peptides by integrating composition-based features. BMC Bioinformatics. 2021;22:389. https://doi.org/10.1186/s12859-021-04305-2

44. Das NC, Gorai S, Gupta PSS, Panda SK, Rana MK, Mukherjee S. Immune targeting of filarial glutaredoxin through a multi-epitope peptide-based vaccine: a reverse vaccinology approach. Int Immunopharmacol. 2024;133:112120. https://doi.org/10.1016/j.intimp.2024.112120

45. Dobhal S, Sharma A, Singh P, Kumar V, Tiwari S, Singh R, et al. In silico identification of MHC displayed tumor associated peptides in ovarian cancer for multi-epitope vaccine construct. Endocr Metab Immune Disord Drug Targets. 2024;24(12):1401-1413. https://doi.org/10.2174/0118715303169428231205173914

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):25-30
DOI: 10.29058/mjwbs.1793565

The correlation between serum YKL-40, proteinuria and inflammation in patients with systemic lupus erythematosus

Sistemik lupus eritematozuslu hastalarda serum YKL-40, proteinüri ve inflamasyon düzeyi arasındaki ilişki

İlyas Öztürk[1], Fatma Betül Güzel[1], Serhat Kura[2], Enes Süleyman Gün[2], Burak Okyar[3], Sinan Kazan[4], Muhammed Seyithanoğlu[5], Gözde Yıldırım Çetin[3], Orçun Altunören[1], Özkan Güngör[1]

[1]Department of Internal Medicine, Division of Nephrology, Kahramanmaraş Sütçü İmam University Faculty of Medicine, Kahramanmaraş, Türkiye

[2]Department of Internal Medicine, Kahramanmaraş Sütçü İmam University Faculty of Medicine, Kahramanmaraş, Türkiye

[3]Department of Internal Medicine, Division of Rheumatology, Kahramanmaraş Sütçü İmam University Faculty of Medicine, Kahramanmaraş, Türkiye

[4]Department of Internal Medicine, Division of Nephrology, Afyonkarahisar Health Science University, Faculty of Medicine, Afyonkarahisar, Türkiye

[5]Department of Biochemistry, Kahramanmaras Sutcu Imam University, Faculty of Medicine, Kahramanmaraş, Türkiye

Cite this article as: *Öztürk İ et al. The correlation between serum YKL-40, proteinuria and inflammation in patients with systemic lupus erythematosus. Med J West Black Sea. 2026;10(1):25-30.*

This study was presented as an oral presentation at the 22nd Nephrology Winter School held in Antalya-Türkiye on 7-10 March 2024.

ABSTRACT

Aim: YKL-40 is a protein associated with acute and chronic inflammation and cancer. Neutrophils, macrophages, and cancer cells normally produce it. It plays a role in cell proliferation, differentiation, angiogenesis, inflammation, and extracellular matrix remodeling. It is also a marker of endothelial dysfunction and atherosclerosis. Serum YKL-40 has been reported to be associated with proteinuria in diabetic and renal transplant patients. In this study, we aimed to investigate, for the first time in the literature, the relationship between the amount of proteinuria and inflammatory markers and YKL-40 in patients with systemic lupus erythematosus (SLE).

Material and Methods: Fifty-one patients with SLE who presented to the adult nephrology outpatient clinic between June 2022 and January 2023 were included in the study. Blood and urine test results were obtained from the patients' routine examinations. ELISA was used to determine YKL-40.

Results: The mean age of the patients was 38±14 years and 94% were female. The median creatinine was 0.72 (0.36-2.62) mg/dl, the mean estimated glomerular filtration rate was 99±31 ml/min, the mean C-reactive protein (CRP) was 8±7 mg/L, the median erythrocyte sedimentation rate (ESR) was 14.5 (4-72) mm/hour, the median amount of proteinuria was 190 (69-4959) mg/g and the mean YKL-40 was 12.9±9.2 pg/ml. In correlation analysis, YKL-40 was positively correlated with glucose, CRP, and ESR [respectively (p=0.01, r=0.26), (p=0.02, r=0.06), (p=0.01, r=0.11)], and negatively correlated with calcium, albumin, in-

ÖZ

Amaç: YKL-40; akut ve kronik inflamasyon yanında kanser ile de ilişkili bir proteindir. Hücre proliferasyonu, diferansiasyonu, anjiogenez, inflamasyon ve ekstrasellüler matriks remodellinginde görevli bu proteinler normalde nötrofiller, makrofajlar ve kanser hücreleri tarafından üretilir. Ayrıca endotel disfonksiyonunun ve aterosklerozun da bir göstergesidir. Diyabetik ve böbrek nakilli hastalarda serum YKL-40 düzeyinin proteinüri ile ilişkili olduğu bildirilmiştir. Biz de bu çalışmada, literatürde ilk defa, Sistemik Lupus Eritematozus (SLE) hastalarında YKL-40 düzeyi ile proteinüri miktarı ve inflamasyon belirteçleri arasındaki ilişkiyi incelemeyi amaçladık.

Gereç ve Yöntemler: Çalışmaya Haziran 2022–Ocak 2023 tarihleri arasında Nefroloji Polikliniğe başvuran 51 SLE hastası dahil edildi. Kan ve idrar tetkik sonuçları hastaların rutin tetkiklerinden elde edildi. YKL-40 düzeyi ELISA yöntemiyle ölçüldü.

Bulgular: Hastaların yaş ortalaması 38±14 yıl olup %94'ü kadındı. Ortanca kreatin 0.72 (0.36-2.62) mg/dl, ortalama tahmini glomerüler filtrasyon hızı 99±31 ml/dk, ortalama C-reaktif protein (CRP) 8±7 mg/L, ortanca eritrosit sedimentasyon hızı (ESH) 14.5 (4-72) mm/sa, ortanca proteinüri miktarı 190 (69-4959) mg/g ve ortalama YKL-40 düzeyi 12.9±9.2 pg/ml'ydi. Korelasyon analizinde, YKL-40; glukoz, CRP, ESH ile pozitif korele [sırasıyla (p=0.01, r=0.26), (p=0.02, r=0.06), (p=0.01, r=0.11)], kalsiyum, albümin, intakt parathormon ve kompleman-3 düzeyi ile negatif korele [sırasıyla (p<0.001, r=-0.41), (p=0.01, r=-0.37),

Corresponding Author: İlyas Öztürk ✉ drilyasozturk@gmail.com **Received:** 30.09.2025 **Revision:** 29.10.2025 **Accepted:** 12.12.2025

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

tact parathyroid hormone, and complement-3 [respectively (p<0.001, r=-0.41), (p=0.01, r=-0.37), (p=0.01, r=-0.41), (p<0.001, r=-0.52)]. No correlation was found between YKL-40 and the amount of proteinuria (p=0.68, r=0.18).

Conclusion: Serum YKL-40 was found to be associated with inflammatory markers such as CRP and ESR in SLE patients, but not with the amount of proteinuria.

Keywords: Inflammation, proteinuria, systemic lupus erythematosus, Ykl-40

(p=0.01, r=-0.41), (p<0.001, r=-0.52)] bulundu. YKL-40 düzeyi ile proteinüri miktarı arasında ilişki tespit edilmedi (p=0.68, r=0.18).

Sonuç: SLE hastalarında YKL-40 düzeyi ile CRP ve ESH gibi inflamasyon markerları arasında ilişki bulunmakla birlikte, proteinüri miktarıyla bir ilişki bulunamamıştır.

Anahtar Kelimeler: Inflamasyon, proteinüri, sistemik lupus eritematozus, Ykl-40

Highlights

- YKL-40 is a protein associated with cell proliferation, differentiation, angiogenesis, and inflammation.
- YKL-40 is produced by neutrophils, macrophages, and cancer cells and is an indicator of endothelial dysfunction and atherosclerosis.
- Serum YKL-40 levels have been reported to be associated with proteinuria in diabetic and kidney transplant patients.
- Our study found a correlation between YKL-40 levels and inflammation markers in Systemic Lupus Erythematosus patients, but no correlation with the amount of proteinuria.

INTRODUCTION

YKL-40, also known as chitinase-3-like protein-1, has a glycoprotein structure and the 40 in its name comes from its molecular weight of 40 kDa. It is known to be associated with acute and chronic inflammation and even cancer (1). Neutrophils, macrophages, and cancer cells normally produce it. It plays a role in cell proliferation and differentiation, angiogenesis, inflammation, and extracellular matrix remodeling (2). It is also a marker of endothelial dysfunction and atherosclerosis (1).

YKL-40 is associated with the amount of albuminuria in patients with type 1 and type 2 diabetes mellitus (DM), and it has been reported that high serum YKL-40 may predict all-cause mortality in these patients (3, 4). It has also been reported to be associated with proteinuria in renal transplant patients (5).

YKL-40 is known to be high in patients with nephrotic syndrome. Moreover, an association between YKL-40 and arterial stiffness has been previously reported in the literature (6). It has also been reported that YKL-40 is higher in patients with end-stage renal disease (ESRD) on dialysis and not on dialysis compared to the healthy population, and that YKL-40 increases with decreasing estimated glomerular filtration rate (eGFR) in ESRD patients (7).

In addition, in recent years, the fact that serum YKL-40 levels may increase in autoinflammatory diseases and may even be related to disease activity has become a popular research topic (8). For all these reasons, high YKL-40 may be a practical and useful follow-up parameter, as it can give an idea of prognosis and disease activation in diseases with renal involvement.

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease characterized by inflammation and immune-mediated damage in multiple organ systems, including the mucocutaneous, musculoskeletal, hematological, and kidney systems. Approximately 40% of SLE patients develop lupus nephritis, and an estimated 10% of individuals with lupus nephritis develop end-stage renal disease within 10 years (9). Therefore, disease monitoring and early detection of renal involvement are crucial for patient surveillance. In this study, we aimed to investigate, for the first time in the literature, the relationship between YKL-40 and the amount of proteinuria and inflammatory markers in SLE patients.

MATERIAL and METHODS

This cross-sectional study was conducted at Kahramanmaras Sutcu Imam University Faculty of Medicine Nephrology Outpatient Clinic. The study was approved by the Ethics Committee of Kahramanmaras Sutcu Imam University Faculty of Medicine on 14.02.2022 with the number 2022/17. Patients were informed before the study and informed consent forms were signed. The study was conducted in accordance with the tenets of the Declaration of Helsinki.

Patients over 18 years of age who presented to the Adult Nephrology Outpatient Clinic between June 2022 and January 2023 and were diagnosed with SLE were included in the study. The 2012 SLICC criteria were used to diagnose SLE (10). Patients with additional diseases known to affect serum YKL-40, including DM, renal transplantation, coronary artery disease, malignancy, additional inflammatory disease, and local or systemic infection were not included in the study.

Patients' age, gender, and additional disease information were obtained from the hospital's automated system files; blood and urine test results were obtained from patients' test results during routine examinations. Serum YKL-40 was assessed by ELISA (Fine Test, Wuhan, China). The kit was able to detect serum YKL-40 in the range of 62.5-4000 pg/ml.

Statistical Analysis: SPSS 20.0 was used to analyze the study data. Continuous variables determined by measuring patients were expressed as mean±SD and categorical data determined by counting were expressed as numbers of patients or percentages. Kolmogorov-Smirnov and/or Shapiro-Wilk tests were used to analyze the normal distribution of continuous data. Pearson or Spearman correlation analysis was used for correlation analysis, depending on the distribution characteristics of the data. A p-value of less than 0.05 was considered statistically significant.

RESULTS

Fifty-one patients with SLE were included in the study. The mean age of the patients was 38±14 years and 94% were female. The median creatinine was 0.72 (0.36-2.62) mg/dl, the mean eGFR was 99±31 ml/min, the mean C-reactive protein (CRP) was 8±7 mg/L, the median erythrocyte sedimentation rate (ESR) was 14.5 (4-72) mm/hour, the median amount of proteinuria was 190 (69-4959) mg/g and the mean YKL-40 was 12.9±9.2 pg/ml. The general characteristics and laboratory results of the patients are detailed in Table 1.

A review of patient files revealed that 21 patients underwent kidney biopsies due to suspected lupus nephritis. The presence of lupus nephritis was confirmed by kidney biopsy in 20 patients. Only one patient's biopsy was not definitively evaluated due to insufficient sample size. In three patients with clinical suspicion of lupus nephritis, kidney biopsies were not performed due to reasons such as contraindications to kidney biopsy, patient refusal to undergo the procedure, or low platelet count at presentation.

In the correlation analysis, YKL-40 was positively correlated with glucose (p=0.01, r=0.26), CRP (p=0.02, r=0.06), and ESR (p=0.01, r=0.11); and negatively correlated with calcium (p<0.001, r=-0.41), albumin (p=0.01, r=-0.37), intact parathyroid hormone (iPTH) (p=0.01, r=-0.41) and complement 3 (C3) (p<0.001, r=-0.52). No correlation was found between YKL-40 and the amount of proteinuria (p=0.68, r=0.18). Detailed information on the correlation analysis is shown in Table 2.

Table 1: General characteristics and laboratory measurements of the patients

Parameters	Mean±SD, Median or n(%) (n=51)
Age (y)	38±14
Gender (Female) (%)	94.1
Glucose (mg/dL)	86.5 (72-252)
BUN (mg/dL)	12 (5-98)
Creatinine (mg/dL)	0.72 (0.36-2.62)
eGFR (ml/dk)	99±31
Sodium (mmol/L)	140 (130-145)
Potassium (mmol/L)	4.3±0.4
Calcium (mg/dL)	9.05 (7.40-9.80)
Albumin (mg/dL)	41.5 (18-47)
Ferritin (mg/L)	44 (6-598)
iPTH (pg/mL)	45 (8-154)
Proteinuria (mg/g)	190 (69-4959)
C3 (mg/L)	0.79±0.22
C4 (mg/L)	0.15±0.06
ANA (%)	86.3
Anti-dsDNA (%)	7.8
CRP (mg/L)	8±7
ESR (mm/hour)	14.5 (4-72)
YKL-40 (pg/mL)	12.9±9.2

Data are presented as mean±SD , median (min–max), or n (%), as appropriate.
BUN: blood urea nitrogen, **eGFR:** estimated glomerular filtration rate, **iPTH:** intact parathyroid hormone, **C3:** complement 3, **C4:** complement 4, **CRP:** C-reactive protein, **ESR:** erythrocyte sedimentation rate

Table 2: Correlation analysis for YKL-40

Parameters	p	R
Age (y)	0.83[a]	-0.05
Glucose (mg/dL)	0.01[b]	0.26
BUN (mg/dL)	0.30[b]	0.08
Creatinine (mg/dL)	0.46[b]	-0.06
Sodium (mmol/L)	0.09[b]	-0.13
Potassium (mmol/L)	0.73[a]	0.04
Calcium (mg/dL)	<0.001[b]	-0.41
Albumin (mg/dL)	0.01[b]	-0.37
Ferritin (mg/L)	0.06[b]	0.03
iPTH (pg/mL)	0.01[b]	-0.41
Proteinuria (mg/g)	0,68[b]	0.18
C3 (mg/L)	<0.001[a]	-0.52
C4 (mg/L)	0.21[a]	-0.15
CRP (mg/L)	0.02[a]	0.06
ESR (mm/hour)	0.01[b]	0.11

[a]:Pearson [b]:Spearman correlation analysis
BUN: blood urea nitrogen, **iPTH:** intact parathyroid hormone, **C3:** complement 3, **C4:** complement 4, **CRP:** C-reactive protein, **ESR:** erythrocyte sedimentation rate

DISCUSSION

Our study is the first to investigate the relationship between serum YKL-40 and the extent of proteinuria and inflammatory markers in SLE patients. Previous studies in the literature have shown that YKL-40 has a high diagnostic potential, especially in young female patients, and can provide insight into disease activation. However, these studies did not evaluate the amount of proteinuria. In this respect, our study is the first in the literature. As a result of our study, we found that YKL-40 was associated with inflammatory markers such as CRP and ESR in SLE patients, but not with the amount of proteinuria.

It has been noted that the increase in circulating YKL-40 is more pronounced in patients with coronary artery disease and that this increase in patients with coronary artery disease may be associated with major cardiovascular events (11, 12). El-Ghorab et al. reported that serum YKL-40 level is a highly sensitive predictor of atherosclerosis in CKD patients, with a sensitivity of 74% in CKD (13).

An attempt has been made to explain the relationship between proteinuria and CVD risk by the fact that systemic endothelial dysfunction also causes vascular damage in the kidneys and increases proteinuria (14). Similarly, Tatar et al. reported that there is a correlation between YKL-40 and proteinuria in kidney transplant recipients and that this correlation may play a role in the pathogenesis of cardiovascular damage in this patient group (5). However, the relationship between YKL-40 and proteinuria is unclear. Proteinuria alone does not always indicate renal damage but is associated with increased systemic vascular permeability (15). The increase in vascular permeability in proteinuric patients may be due to endothelial dysfunction. The relationship between endothelial dysfunction markers and micro and macroalbuminuria has been attempted to be explained in this way (16). The main idea of our study was that YKL-40 is an indicator of endothelial dysfunction and atherosclerosis and therefore can give an idea about proteinuria.

This relationship has been studied in different patient groups before in the literature. In a study conducted by Kocyigit et al. on patients with nephrotic syndrome, it was found that there is a relationship between serum YKL-40 and endothelial dysfunction and arterial stiffness and that the amount of proteinuria is significantly higher in patients with nephrotic syndrome with high serum YKL-40, therefore serum YKL-40 may be an indicator of proteinuria in patients with nephrotic syndrome (6). A study by Persson et al. found that high serum YKL-40 in patients with type 2 DM and increased proteinuria could predict all-cause mortality (4). Similarly, other studies in the literature have reported an association between the degree of albuminuria and serum and/or urinary YKL-40 in diabetic patients (17-19). In our study, unlike previous studies, we found no similar association between YKL-40 and proteinuria in SLE patients. The fact that proteinuria was assessed using spot urine and the small sample size may be a contributing factor. Because our study was conducted retrospectively, upon reviewing patient files, we observed that 24-hour urine data were missing. This was due to the difficulty of collecting 24-hour urine samples and the fact that routine outpatient clinic follow-ups were performed using the first morning urine sample, which is known to provide the closest proteinuria results, rather than the 24-hour urine sample. Although similar results are known to be obtained, we may not have been able to detect a possible association because the 24-hour urine sample was not used. We also believe that the small sample size may be another factor, and we note this as a limitation of our study.

A study by Keskin et al. found that YKL-40 is an inflammatory marker that increases with the stage of ESRD, is associated with markers of endothelial dysfunction, and is also correlated with CRP and uric acid (7). A study by Lorenz et al. found that high YKL-40 in ESRD patients may be a marker for risk prediction of cardiovascular mortality (20). Similarly, El-Ghorab et al. reported that YKL-40 was a highly sensitive indicator of CRP, ESR, and increased carotid intima-media thickness in CKD patients, but they could not detect any relationship with ferritin and albumin levels (13). In our study, a correlation was found between YKL-40 and CRP, ESR and hypoalbuminemia, but no correlation was found with ferritin level.

There are few studies in the literature on the relationship between YKL-40 and SLE patients. A study by Ali et al. found that YKL-40 has a high diagnostic potential in SLE, especially in young women. This study reported that YKL-40 levels were associated with disease duration, anemia, thrombocytopenia, positive anti-double stranded DNA (anti-dsDNA), and antiphospholipid antibody (APL) levels. The study was conducted in juvenile SLE patients, and YKL-40 levels were reported to increase with increasing age. The relationship between SLE diagnostic criteria and YKL-40 levels was analyzed, and no association was reported with nephrological involvement (21). The study was conducted in a younger patient population and examined general renal findings such as nephrological involvement. Specific findings such as proteinuria were not examined. Furthermore, the relationship with SLE clinical findings was analyzed, but no analysis was conducted regarding markers such as ESR or CRP. In these respects, it differs from our study. A study by Gamal El-Agamy et al. reported that serum YKL-40 was higher in SLE patients than in healthy controls and that YKL-40 correlated with disease activity, CRP, ESR, and serum creatinine (22). Similarly, in our study, we observed a correlation between YKL-40 and CRP, ESR, and hypoalbuminemia which are traditional indicators of SLE activation.

The relationship between YKL-40 and various autoimmune diseases has been investigated in the literature. It has been previously reported in the literature that it can provide insight into the diagnosis, prognosis and disease activity of diseases such as Rheumatoid Arthritis, Psoriasis, Behçet's disease and Inflammatory bowel disease, in addition to SLE (8). A study by Chen et al. found that serum YKL-40 was a cytokine associated with disease activity, as were TNF-α and IL-6, in patients with primary Sjögren's syndrome. It was emphasized that these three cytokines are also related to CRP and ESR. In addition, serum YKL-40 is associated with lung and kidney involvement, neutrophil count, leukopenia, and anemia in these patients (23).

The first limitation of our study is the number of patients. It would be appropriate to support the issue with studies conducted in larger groups of patients. Second, proteinuria was assessed with spot urine. Another limitation is that serum YKL-40 was measured with a single assay. It is more likely that more reliable results will be obtained if serial measurements are used.

Conclusion

Our study is the first to investigate the relationship between serum YKL-40, proteinuria, and inflammatory markers in SLE patients. We found that serum YKL-40 was associated with inflammatory markers such as CRP and ESR in SLE patients, but not with the amount of proteinuria. The small number of patients may be a factor influencing this and further studies are needed.

Author Contributions

Study conception and design: **Ilyas Ozturk, Ozkan Gungor;** data collection: **Fatma Betul Guzel, Serhat Kura, Enes Suleyman Gun, Burak Okyar, Muhammed Seyithanoglu;** analysis and interpretation of results: **Ilyas Ozturk, Sinan Kazan, Gozde Yildirim Cetin, Orcun Altunoren, Ozkan Gungor;** draft manuscript preparation: **Ilyas Ozturk, Ozkan Gungor.** The author(s) reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors have no conflict of interest to declare.

Ethical Approval

The study was approved by the Ethics Committee of Kahramanmaras Sutcu Imam University Faculty of Medicine on 14.02.2022 with the number 2022/17. Patients were informed before the study and informed consent forms were signed.

REFERENCES

1. Rathcke CN, Vestergaard H. YKL-40, a new inflammatory marker with relation to insulin resistance and with a role in endothelial dysfunction and atherosclerosis. Inflamm. res. 55, 221–227 (2006). https://doi.org/10.1007/s00011-006-0076-y.
2. Okyay GU, Er RE, Tekbudak MY, Paşaoğlu Ö, Inal S, Öneç K, et al. Novel Inflammatory Marker in Dialysis Patients: YKL-40. Therapeutic Apheresis and Dialysis, 2013; 17(2), 193-201. https://doi.org/10.1111/j.1744-9987.2012.01141.x.
3. Rathcke CN, Persson F, Tarnow L, Rossing P, Vestergaard H. YKL-40, a marker of inflammation and endothelial dysfunction, is elevated in patients with type 1 diabetes and increases with levels of albuminuria. Diabetes Care, 2009; 32: 323–28. https://doi.org/10.2337/dc08-1144.
4. Persson F, Rathcke CN, Gall MA, Parving H-H, Vestergaard H, Rossing P. High YKL-40 levels predict mortality in patients with type 2 diabetes. Diabetes Res Clin Pract, 2012; 96:84–89. https://doi.org/10.1016/j.diabres.2011.12.008.
5. Tatar E, Gungor O, Celtik A, Sisman AR, Yaprak M, Asci G, et al. Correlation between serum YKL-40 (Chitinase-3-like protein 1) level and proteinuria in renal transplant recipients. Ann Transplant, 2013: 18: 95-100. https://doi.org/10.12659/AOT.883844.
6. Kocyigit I, Gungor O, Dogan E, Karadavut S, Karakukcu C, Eroglu E, et al, Oymak O. (2014). The serum YKL-40 level is associated with vascular injury and predicts proteinuria in nephrotic syndrome patients. J Atheroscler Thromb, 2015; 22: 257-264. https://doi.org/10.5551/jat.26385.
7. Keskin GSY, Helvacı Ö, Yayla Ç, Paşaoğlu ÖT, Keskin Ç, Arınsoy ST, et al. Relationship between plasma YKL-40 levels and endothelial dysfunction in chronic kidney disease. Turkish journal of medical sciences, 2019; 49(1), 139-146. https://doi.org/10.3906/sag-1804-169.
8. Tizaoui K, Yang JW, Lee KH, Kim JH, Kim M, Yoon S, et al. The role of YKL-40 in the pathogenesis of autoimmune diseases: a comprehensive review. Int J Biol Sci, 2022; 18(9), 3731-3746. https://doi.org/10.7150/ijbs.67587.
9. Siegel CH, Sammaritano LR. Systemic Lupus Erythematosus: A Review. JAMA. 2024;331(17):1480–1491. https://doi.org/10.1001/jama.2024.2315.
10. Petri M, Orbai AM, Alarcón GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation andvalidation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum. 2012; 64: 2677–2686. https://doi.org/10.1002/art.34473.
11. Kucur M, Isman FK, Karadag B, Vural VA, Tavsanoglu S. Serum YKL-40 levels in patients with coronary artery disease. Coron Artery Dis 2007;18(5):391–6. https://doi.org/10.1097/MCA.0b013e328241d991.
12. Rathcke CN, Vestergaard H. YKL-40—an emerging biomarker in cardiovascular disease and diabetes. Cardiovasc Diabetol 2009;8:61. https://doi.org/10.1186/1475-2840-8-61.
13. El-Ghorab OAAN, Elshafey MHR, Negm MSA, Elkhateeb MAHT. Evaluation of YKL-40 as a Biomarker of Inflammation and Atherosclerosis in Chronic Kidney Disease Patients. Journal of Advances in Medicine and Medical Research. 2022;34(21),10-26.
14. Currie G, Delles C.Proteinuria and its relation to cardiovascular disease. Int J Nephrol Renovas Dis. 2014; 7: 13-24. https://doi.org/10.2147/IJNRD.S40522.
15. Joles JA, Stroes ESG, Rabelink TJ. Endothelial function in proteinuric renal disease. Kid Int 1999; 56: 71.pp57-61. https://doi.org/10.1046/j.1523-1755.1999.07115.x.
16. Schalkwijk CG, Stehouwer CDA. Vascular complications in diabetes mellitus: the role of endothelial dysfunction. Clin Sci (Lond). 2005; 109(2):143-159. https://doi.org/10.1042/CS20050025.

17. Brix JM, Hollerl F, Koppensteiner R, Schernthaner G, Schernthaner GH. YKL-40 in type 2 diabetic patients with different levels of albuminuria. Eur J Clin Invest 2011;41(6):589–96. https://doi.org/10.1111/j.1365-2362.2010.02446.x.

18. Rondbjerg AK, Omerovic E, Vestergaard H. YKL-40 levels are independently associated with albuminuria in type 2 diabetes. Cardiovasc Diabetol 2011;10:54. https://doi.org/10.1186/1475-2840-10-54.

19. Lee JH, Kim SS, Kim IJ, Song SH, Kim YK, In Kim J, et al. Clinical implication of plasma and urine YKL-40, as a proinflammatory biomarker, on early stage of nephropathy in type 2 diabetic patients. J Diabetes Complications 2012; 26: 308-312. https://doi.org/10.1016/j.jdiacomp.2012.04.012.

20. Lorenz G, Schmalenberg M, Kemmner S, Haller B, Steubl D, Pham D, et al. Mortality prediction in stable hemodialysis patients is refined by YKL-40, a 40-kDa glycoprotein associated with inflammation. Kidney International, 2018; 93(1), 221-230. https://doi.org/10.1016/j.kint.2017.07.010.

21. Ali AA, Yousef RN, Elsheikh MS, Salamah AR, Wu LL, Alnaggar AR, et al. YKL-40 in serum: a promising biomarker of juvenile SLE and strongly correlated with disease duration. Irish Journal of Medical Science (1971 -), 2024; 193(3), 1403-1409. https://doi.org/10.1007/s11845-023-03545-w.

22. Gamal El-Agamy A, Abdel Noor R, M Mabrouk M, Ahmed El-Tatawy R, Salah Hegab D. Assessment of Serum Level of Chitinase-3-Like Protein-1 (Ykl-40) in Systemic Lupus Erythematosus Patients and Its Relation to Disease Activity. Journal of Advances in Medicine and Medical Research, 2023; 35(1), 28-37. https://doi.org/10.9734/jammr/2023/v35i14931.

23. Chen C, Liang Y, Zhang Z, Zhang Z, Yang Z. Relationships between increased circulating YKL-40, IL-6 and TNF-α levels and phenotypes and disease activity of primary Sjögren's syndrome. International Immunopharmacology, 2020; 88, 106878. https://doi.org/10.1016/j.intimp.2020.106878.

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):31-38
DOI: 10.29058/mjwbs.1836406

Predictors of early readmission in patients awaiting interval cholecystectomy after conservative treatment of acute cholecystitis: A retrospective cohort study

Akut kolesistitte konservatif tedavi sonrası interval kolesistektomi bekleyen hastalarda erken yeniden yatışın prediktörleri: Retrospektif kohort çalışması

Kazım Duman

Department of General Surgery, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, İstanbul, Türkiye

Cite this article as: *Duman K. Predictors of early readmission in patients awaiting interval cholecystectomy after conservative treatment of acute cholecystitis: a retrospective cohort study. Med J West Black Sea. 2026;10(1):31-38.*

ABSTRACT

Aim: Although interval cholecystectomy is a common strategy in the treatment of acute cholecystitis, a significant proportion of patients are readmitted due to complications during the waiting period. This study aims to identify factors affecting early readmission in patients scheduled for interval cholecystectomy after conservative treatment.

Material and Methods: In this retrospective cohort study, patients hospitalized with acute cholecystitis between January 2023 and December 2024 who received conservative treatment and were scheduled for interval cholecystectomy were evaluated. Patients were classified as Early Presentation Group (EPG) and Normal Presentation Group (NPG) based on whether they were readmitted within six weeks after discharge. Demographic, clinical, and laboratory data were compared; risk factors were evaluated by logistic regression analysis.

Results: A total of 110 patients were included in the study (NPG: 78, EPG: 32). In multivariate analysis, male gender (OR: 5.108; 95% CI: 1.758-14.852; p=0.003) and diabetes mellitus (OR: 8.521; 95% CI: 2.933-24.756; p<0.001) were identified as independent predictors of early readmission. Age, inflammatory markers, and gallbladder stone size showed no significant association.

Conclusion: Male gender and diabetes mellitus are independent risk factors for early readmission in patients awaiting interval cholecystectomy. Early surgical intervention should be considered instead of prolonged conservative treatment in patients with these risk factors.

Keywords: Acute cholecystitis, cholecystectomy, patient readmission, diabetes mellitus, risk factors

ÖZ

Amaç: Akut kolesistit tedavisinde interval kolesistektomi yaygın bir strateji olmasına rağmen, bekleme döneminde hastaların önemli bir kısmı komplikasyonlar nedeniyle yeniden hastaneye yatmaktadır. Bu çalışma, konservatif tedavi sonrası interval kolesistektomi planlanan hastalarda erken yeniden yatışı etkileyen faktörleri belirlemeyi amaçlamaktadır.

Gereç ve Yöntemler: Bu retrospektif kohort çalışmasında, Ocak 2023-Aralık 2024 tarihleri arasında akut kolesistit tanısıyla yatırılarak konservatif tedavi uygulanan ve interval kolesistektomi planlanan hastalar değerlendirildi. Hastalar, taburculuk sonrası altı hafta içinde yeniden başvurup başvurmamalarına göre Erken Başvuru Grubu (EBG) ve Normal Başvuru Grubu (NBG) olarak sınıflandırıldı. Demografik, klinik ve laboratuvar verileri karşılaştırıldı; risk faktörleri lojistik regresyon analizi ile değerlendirildi.

Bulgular: Çalışmaya 110 hasta dahil edildi (NBG= 78, EBG= 32). Çok değişkenli analizde erkek cinsiyet (OR: 5,108; %95 GA: 1,758-14,852; p=0,003) ve diyabetes mellitus (OR: 8,521; %95 GA: 2,933-24,756; p<0,001) erken yeniden yatış için bağımsız prediktörler olarak saptandı. Yaş, inflamatuar belirteçler ve safra kesesi taşı boyutu anlamlı ilişki göstermedi (p>0.05).

Sonuç: Erkek cinsiyet ve diyabetes mellitus, interval kolesistektomi bekleyen hastalarda erken yeniden yatış için bağımsız risk faktörleridir. Bu risk faktörlerine sahip hastalarda uzun süreli konservatif tedavi yerine erken cerrahi müdahale düşünülmelidir.

Anahtar Kelimeler: Akut kolesistit, kolesistektomi, hasta yeniden yatışı, diyabetes mellitus, risk faktörleri

Corresponding Author: Kazım Duman ✉ dr.kazimduman@gmail.com **Received:** 05.12.2025 **Revision:** 15.12.2025 **Accepted:** 21.12.2025

Copyright © 2025 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Male gender and diabetes mellitus are independent predictors of early readmission in patients awaiting interval cholecystectomy.
- Diabetic patients have an 8.5-fold increased risk of readmission compared to non-diabetic patients.
- Male patients show a 5-fold higher risk of readmission compared to female patients.
- Age, C-reactive protein, white blood cell count, and gallbladder stone size measured at initial admission did not show a significant association with readmission.
- Early cholecystectomy should be considered instead of delayed surgery in high-risk patients.

INTRODUCTION

Acute cholecystitis is a major cause of emergency department visits and the third leading cause of abdominal complications requiring hospitalization. The incidence of the disease increases with age, and approximately 10-15% of the adult population in Western societies has gallstones. The optimal timing of surgery for acute cholecystitis remains a controversial issue. Although the Tokyo Guidelines 2018 (TG18) criteria define treatment strategies based on disease severity, the applicability and outcomes of these recommendations in clinical practice vary (1,2).

There are two basic approaches to the treatment of acute cholecystitis: early cholecystectomy, performed within 72 hours of symptom onset, and interval (delayed) cholecystectomy, performed after 6-12 weeks of conservative treatment. Although randomized controlled trials have shown that early cholecystectomy shortens hospital stays and is cost-effective, many centers prefer interval cholecystectomy, particularly in high-risk patients, due to technical difficulties and concerns about complications (2-4).

Patients who are successfully discharged with conservative treatment and scheduled for interval cholecystectomy may develop significant complications during the waiting period. The literature reports that rehospitalization rates during this period range from 15% to 40% due to complications such as recurrent cholecystitis, acute cholangitis, biliary pancreatitis, and gallbladder perforation. These readmissions not only increase patient morbidity, but also place additional burden on the healthcare system and increase the risk of complications by requiring planned surgery to be performed in emergency settings (5).

Identifying patients at high risk for complications and early readmission during the interval period is critical for individualizing treatment strategies. Current literature suggests that factors such as advanced age, male gender, the presence of diabetes mellitus, elevated CRP and white blood cell counts, gallbladder wall thickness, and the presence of pericholecystic effusion may be associated with the risk of readmission. However, there is no consensus on the predictive value of these factors and in which patient groups early cholecystectomy is absolutely indicated (5,6).

Recent systematic reviews and meta-analyses highlight the lack of reliable predictive models for risk stratification in patients awaiting interval cholecystectomy (2,5,7). This retrospective study aims to identify factors affecting early readmission in patients undergoing conservative treatment for acute cholecystitis and scheduled for interval cholecystectomy.

MATERIAL and METHODS

Study Design

The study was approved by the Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital Scientific Research Ethics Committee (Decision No: 2025/436). The study then began by reviewing the data of patients with a history of a single hospitalization for acute cholecystitis at the same institution. Due to the retrospective nature of the study, patient consent was not obtained. Patients hospitalized with a preliminary diagnosis of acute cholecystitis between January 2023 and December 2024 were screened for the study.

Inclusion Criteria

To ensure a homogeneous patient cohort, specific criteria were established for inclusion in the study. These criteria included emergency hospital admission with only a preliminary diagnosis of acute cholecystitis, diagnosis supported by abdominal ultrasonography, age 18 years or older and younger than 80 years, initial follow-up with standard daily treatment of 2 g ceftriaxone and 3 x 0.5 g metronidazole for at least 3 days, and discharge after planned cholecystectomy following medical follow-up.

Exclusion Criteria

To reduce the variability of the study, specific exclusion criteria were established. These included the presence of acute biliary pancreatitis or cholangitis at presentation, the

presence of an additional inflammatory or neoplastic disease of hepatobiliary origin, being younger than 18 or older than 70 years, a history of prior hospitalization for acute cholecystitis, having undergone non-invasive procedures or direct cholecystectomy at initial hospitalization, receiving treatment other than standard antibiotics at admission, or being allergic to metronidazole or penicillin.

Patient Follow-up Protocols

The study began by scanning hospital records for patients admitted to the hospital with a diagnosis of acute cholecystitis. Patients who were admitted to hospital with a preliminary diagnosis of acute cholecystitis were excluded. Their hemogram and biochemistry parameters were evaluated during admission and at discharge. Ultrasound findings from the initial admission were also included. The ultrasound findings assessed increased gallbladder wall thickness and stone size. Stone sizes were recorded separately as millimeters and as large as 1 cm. Comorbidities were also assessed based on patient demographics.

All patients were admitted to the ward with oral feedings off. In addition to intravenous (IV) antibiotics, analgesia, antiemetics, and crystalloids and fluid replacement were administered. Oral feeding was gradually initiated in patients who responded clinically and laboratory tests and who did not undergo early surgery. Patients who received at least 3 days of treatment and were able to receive oral feedings were then discharged with elective surgery scheduled for at least 6 weeks later. Length of stay was recorded. Upon discharge, patients were prescribed two 1-gram ceftriaxone tablets with analgesia. Patients who presented earlier after discharge were grouped as the “Early Presentation Group (EPG)” and the “Normal Presentation Group (NPG).” The extracted demographic and clinicopathological data were compared between the two groups.

Statistical Analysis

SPSS 27.0 (SPSS Inc., Chicago, IL) was used for statistical analyses. The Kolmogorov–Smirnov test was used for the distribution of numerical data. The Mann–Whitney U test was used to evaluate non-parametric data. The Pearson Chi-Square test was used for the distribution of categorical data. Fisher’s exact test was applied when any expected cell count was less than 5. Parameters found to be significant as risk factors for early hospitalization were subjected to regression analysis. Non-parametric variables were processed as median (25th–75th percentile). Categorical data are presented as frequencies (%). p-value <0.05 was considered statistically significant.

RESULTS

There were total of 110 patients included in the study. NPG group consisted 78; EPG group consisted 32 patients. In the comparison of demographic, clinical, and laboratory variables; two variables demonstrated statistically significant differences between groups. Regarding gender distribution, a notably higher proportion of female patients was observed in the EPG compared to the NPG (56.3% vs 19.3%, p<0.001), suggesting that women tend to seek medical attention earlier following symptom onset. Similarly, diabetes mellitus showed a striking association with early presentation, with 62.5% of EPG patients having diabetes compared to only 17.9% in the NPG (p<0.001) (Figure 1). This finding may reflect heightened health awareness or closer

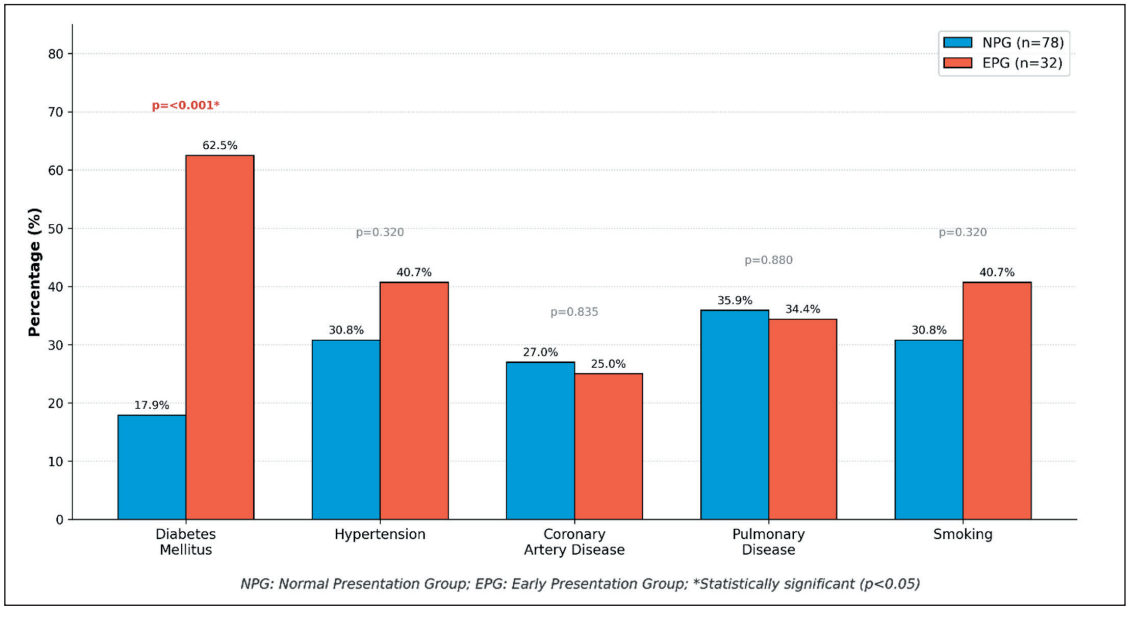


Figure 1: Distribution of Comorbidities Between Study Groups

Table 1: Variable comparison based on early admission

		NPG (n = 78)	EPG (n = 32)	p value •
		n (%)	n (%)	
Gender	Male	63 (80.7)	14 (43.7)	<0.001
	Female	15 (19.3)	18 (56.3)	
DM	No	64 (82.1)	12 (37.5)	<0.001
	Yes	14 (17.9)	20 (62.5)	
HT	No	54 (69.2)	19 (59.3)	0.320
	Yes	24 (30.8)	13 (40.7)	
CAD	No	57 (73.0)	24 (75.0)	0.835
	Yes	21 (27.0)	8 (25.0)	
Pulmonary Disease	No	50 (64.1)	21 (65.6)	0.880
	Yes	28 (35.9)	11 (34.4)	
Smoking	No	54 (69.2)	19 (59.3)	0.320
	Yes	24 (30.8)	13 (40.7)	
USG	Mm GB Stones	68 (87.1)	27 (84.3)	0.762
	>1 cm GB Stones	10 (12.9)	5 (15.7)	
		Median (IQR)		p value ••
Age	Years	43.00 (42.00-46.00)	45.00 (40.00-56.00)	0.600
Lymphocyte Count	µl	13.50 (10.90-16.10)	12.45 (10.40-15.80)	0.838
Crp	mg/dL	9.70 (7.90-11.00)	7.80 (6.50-10.70)	0.730
Alt	(U/L)	20.80 (17.90-23.70)	21,65 (18.20-26.90)	0.359
Ast	(U/L)	35.70 (32.30-38.70)	36.65 (33.20-41.90)	0.342
Ggt	U/L	51.65 (45.20-57.70)	51.05 (43.80-60.90)	0.703
Alp	U/L	97.95 (82.90-112.10)	105.10 (93.80-113.00)	0.435
PA Complaint Period	Days	6.00 (6.00-7.00)	6.50 (5.00-7.00)	0.021
Admission Period	Days	2.00 (2.00-3.00)	2.00 (2.00-3.00)	0.474

NPG: Normal Presentation Group, **EPG:** Early Presentation Group, **DM:** Diabetes, **HT:** Hypertension, **CAD:** Coronary Artery Disease, **USG:** Ultrasonography, **CRP:** C Reactive Protein, **Ast:** Aspartate Aminotransferase, **Alt:** Alanine Aminotransferase, **GGT:** Gamma-Glutamyltransferase, **Alp:** Alkaline Phosphatase, **PA:** Pre-application, **GB:** Gallbladder, **IQR:** Interquartile Range.
•: Pearson Chi-square test, **••:** Mann-Whitney U test.

medical follow-up among diabetic patients due to their existing chronic condition. The pre-application complaint period also differed significantly between groups (median 6.5 vs 6.0 days, p=0.021), though the clinical relevance of this difference appears limited given the overlapping interquartile ranges.

Conversely, no statistically significant differences were identified for hypertension (p=0.320), coronary artery disease (p=0.835), pulmonary disease (p=0.880), smoking status (p=0.320), gallbladder stone size on ultrasonography (p=0.697), age (p=0.600), or any of the laboratory parameters including white blood cell count, C-reactive protein, and liver enzymes (all p>0.05). The admission period was also comparable between groups (median 2 days for both, p=0.474) (Table 1).

In univariate logistic regression analysis, three variables demonstrated statistically significant associations with early readmission following discharge. Male gender was associated with a 5.4-fold increased odds of early readmission compared to females (OR 5.400, 95% CI 2.202-13.244, p<0.001), while the presence of diabetes mellitus conferred the highest risk with approximately 7.6 times greater odds (OR 7.619, 95% CI 3.036-19.121, p<0.001). Additionally, admission period showed a modest but significant association, with each additional day of hospitalization increasing the odds of early readmission by approximately 46% (OR 1.458, 95% CI 1.054-2.018, p=0.023).

After adjusting for potential confounders in multivariate analysis, both male gender and diabetes mellitus retained their statistical significance as independent predictors of early readmission. Male patients demonstrated a five-fold

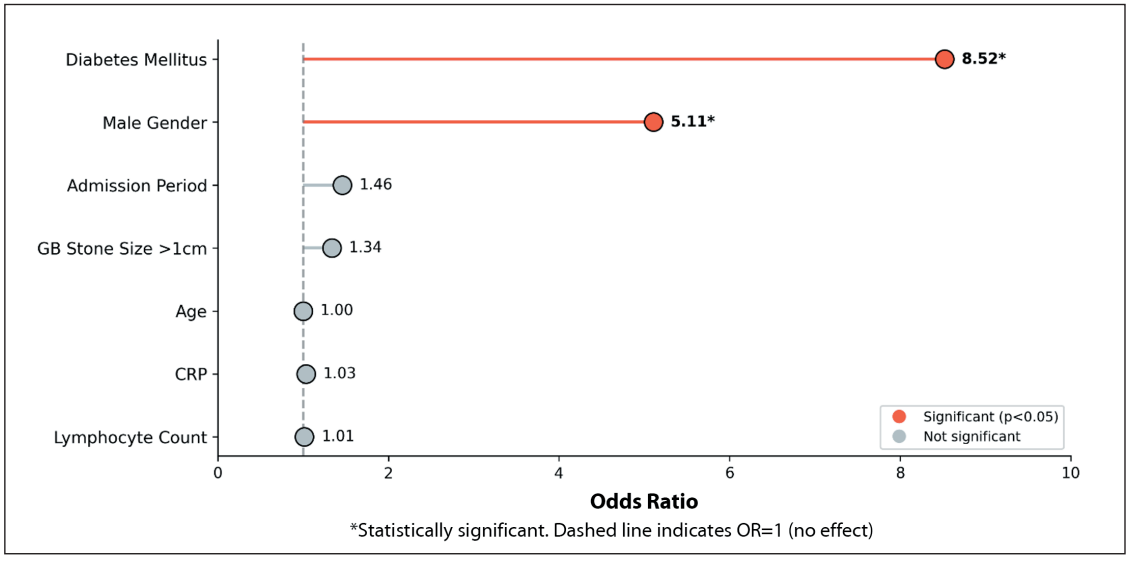


Figure 2: Multivariate Analysis: Predictors of Early Readmission

Table 2: Regression analysis on Early Admission After Discharge

		Univariate			Multivariate		
		OR	%95 CI	p	%95 CI	IQR	p
Lymphocyte count	µl	0.993	0.926-1.065	0.849	1.015	0.929-1.109	0.737
Crp	mg/dL	0.981	0.891-1.080	0.694	1.033	0.916-1.165	0.595
Gender	Male	5.400	2.202-13.244	<0.001	5.108	1.758-14.852	0.003
Age	Years	1.009	0.974-1.045	0.622	1.001	0.960-1.043	0.976
DM	Yes	7.619	3.036-19.121	<0.001	8.521	2.933-24.756	<0.001
GB stone size	>1 cm	1.259	0.394-1.027	0.698	1.338	0.310-5.770	0.696
Admission Period	Days	1.458	1.054-2.018	0.023	1.460	0.981-2.173	0.062

NPG: Normal Presentation Group, **EPG:** Early Presentation Group, **CRP:** C Reactive Protein, **DM:** Diabetes, **GB:** Gallbaldder, **OR:** Odds Ratio, **CI:** Confidence Interval

increased risk compared to females (OR 5.108, 95% CI 1.758-14.852, p=0.003), while diabetic patients showed an 8.5-fold increased odds of early readmission (OR 8.521, 95% CI 2.933-24.756, p<0.001) (Figure 2). The admission period, however, lost its statistical significance after adjustment (OR 1.460, 95% CI 0.981-2.173, p=0.062), suggesting that its univariate association may have been influenced by confounding factors, although the borderline significance indicates a potential trend warranting further investigation. Lymphocyte count, C-reactive protein, age, and gallbladder stone size demonstrated no significant associations in either univariate or multivariate analyses (Table 2).

DISCUSSION

The present study identified male gender and diabetes mellitus as independent predictors of early readmission in patients awaiting interval cholecystectomy following conservative management of acute cholecystitis. In multivariate analysis, male patients demonstrated a five-fold increased risk of readmission within six weeks of discharge, while diabetic patients exhibited an even more pronounced association with an 8.5-fold increased odds of early readmission. These findings carry significant clinical implications, as they suggest that a substantial proportion of patients managed conservatively may benefit from closer surveillance or earlier surgical intervention based on readily identifiable risk factors assessed at the time of initial presentation.

The management of acute cholecystitis remains a subject of ongoing debate, with interval cholecystectomy representing a widely adopted strategy particularly for patients deemed high-risk for immediate surgical intervention. This delayed approach, typically performed six to twelve weeks following the initial episode, offers several theoretical advantages including resolution of acute inflammation, optimization of comorbid conditions, and the opportunity for elective surgical planning in a controlled setting. However, the waiting period

is not without consequences, as a considerable proportion of patients experience recurrent biliary symptoms or complications necessitating unplanned hospital readmission. These early readmissions are clinically concerning not only because they increase patient morbidity and healthcare costs, but also because patients presenting with recurrent or complicated cholecystitis often require emergency surgery under less favorable conditions than originally planned. Therefore, accurate identification of patients at heightened risk for early readmission is essential to guide individualized treatment decisions and determine which patients may benefit from expedited surgical intervention rather than prolonged conservative management (8,9).

In our cohort, males demonstrated a five-fold increased risk of readmission compared to females. Köstenbauer et al. reported similar results in their population-based study of over 47,000 cholecystectomies, where male sex was significantly associated with delayed surgical intervention and poorer outcomes. The underlying mechanisms may involve later presentation with more severe disease, differences in pain perception, or higher prevalence of antiplatelet and anticoagulant use among males. However, our finding contrasts with the pediatric literature, where Pathak et al. reported that female patients had higher odds of 30-day readmission. This discrepancy likely reflects the distinct characteristics of our adult cohort, in which male patients tended to be older and had a substantially higher prevalence of diabetes mellitus, both of which may have contributed to the observed gender-related differences in readmission risk (9,10).

Diabetes mellitus emerged as the strongest independent predictor of early readmission in our study, with diabetic patients demonstrating an 8.5-fold increased risk compared to non-diabetic individuals. This finding is consistent with the established literature indicating that acute cholecystitis follows a more aggressive clinical course in diabetic patients. Several pathophysiological mechanisms may explain this association. Diabetic autonomic neuropathy has been shown to impair gallbladder motility, resulting in increased fasting gallbladder volume and reduced ejection fraction, which promotes bile stasis and facilitates stone formation and recurrent inflammation. Furthermore, hyperglycemia itself directly reduces gallbladder contractility in response to cholecystokinin stimulation, independent of neuropathy. Diabetic patients also exhibit impaired immune responses, with altered neutrophil function and compromised microvascular circulation, predisposing them to more severe infectious complications and rapid disease progression. Additionally, the metabolic derangements associated with diabetes, including hypertriglyceridemia and altered bile composition with cholesterol supersaturation, create a favorable environment for gallstone formation and recurrence. These factors collectively explain why diabetic patients managed conservatively may be at substantially higher risk of developing recurrent biliary symptoms necessitating unplanned readmission during the interval period, supporting consideration of expedited surgical intervention in this population (11-13).

Interestingly, several variables traditionally considered relevant in acute cholecystitis did not demonstrate significant associations with early readmission in our analysis. Age, white blood cell count, C-reactive protein levels, liver enzymes, and gallbladder stone size showed no predictive value for readmission within six weeks. This finding contrasts with studies examining disease severity at initial presentation, where inflammatory markers such as CRP and neutrophil-to-lymphocyte ratio have been established as reliable predictors of complicated cholecystitis and conversion to open surgery. However, these markers primarily reflect the acute inflammatory state at the time of admission rather than the likelihood of recurrent symptoms during the interval period. Similarly, while advanced age has been consistently associated with delayed surgical intervention and increased perioperative morbidity, our data suggest that age alone does not predict recurrence risk in conservatively managed patients. Regarding gallbladder stone size, although some studies have reported that larger stones may be associated with late readmissions due to mechanical obstruction, our findings align with recent evidence suggesting that stone characteristics are inconsistent predictors of gallstone disease recurrence. The RELAPSTONE multicenter study similarly found that stone-related parameters were not reliable determinants of relapse, whereas patient-specific factors such as comorbidities played a more significant role. These observations suggest that the pathophysiology driving recurrent cholecystitis during the interval period may differ from that determining initial disease severity, with metabolic and systemic factors such as diabetes exerting greater influence than acute inflammatory parameters (14-17).

This study has several limitations that should be acknowledged. First, the retrospective design inherently limits the ability to establish causal relationships and introduces potential selection bias. Also our cohort size was relatively small, which may have limited statistical power to detect associations with smaller effect sizes. Due to the retrospective nature of data collection, certain clinically relevant variables could not be consistently evaluated across all patients. Specifically, fever at presentation, gallbladder wall thickness measurements, and the presence of additional pathological findings such as pericholecystic abscess or perforation were not uniformly documented in medical records and therefore could not be included in the analysis.

Our cohort may not capture all patients who experienced recurrent symptoms during the interval period, as some patients may have sought care at different institutions or been

lost to follow-up prior to their scheduled cholecystectomy. This limitation potentially underestimates the true readmission rate and may introduce bias if patients lost to follow-up differ systematically from those retained in the study. Finally, as a single-center study, the generalizability of our findings to other healthcare settings and patient populations may be limited. Future prospective multicenter studies with standardized data collection protocols and longer follow-up periods are warranted to validate these findings and identify additional predictors of early readmission in this patient population.

Despite these limitations, our study has notable strengths. We included a homogeneous cohort of patients who received standardized conservative treatment with uniform antibiotic regimens and discharge protocols, minimizing treatment-related variability. The clearly defined endpoint of readmission within six weeks prior to scheduled cholecystectomy provides clinically relevant data for surgical planning. From a practical standpoint, the identified predictors, male gender and diabetes mellitus, are easily ascertainable at the time of initial presentation without requiring additional diagnostic workup or cost. This allows for immediate risk stratification at the bedside, enabling clinicians to identify high-risk patients who may benefit from expedited cholecystectomy during the index admission or early in the interval period rather than standard delayed surgery. Implementation of such risk-based triage could potentially reduce emergency readmissions, decrease healthcare costs, and improve patient outcomes by avoiding surgery under unfavorable emergency conditions.

Conclusion

In conclusion, this study identified male gender and diabetes mellitus as independent predictors of early readmission in patients awaiting interval cholecystectomy following conservative management of acute cholecystitis. Diabetic patients demonstrated the highest risk with an 8.5-fold increased odds of readmission, while male patients exhibited a five-fold increased risk compared to their counterparts. These findings suggest that patients with these readily identifiable risk factors may benefit from expedited surgical intervention rather than prolonged conservative management. Clinicians should consider prioritizing early cholecystectomy in diabetic and male patients to reduce the burden of recurrent biliary complications during the interval period. Further prospective studies are needed to develop and validate risk stratification tools that can guide individualized treatment decisions in this patient population.

Author Contributions

Study design and conceptualization: **Kazım Duman,** data collection: **Kazım Duman,** analysis and interpretation of results: KD; preparation of the draft article: **Kazım Duman.** The author reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors have no confict of interest to declare.

Financial Support

This study received no financial support.

Ethical Approval

This study was approved by the Scientific Research Ethics Committee of Sancaktepe Martyr Prof. Dr. İlhan Varank Training and Research Hospital (Date: 26.11.2025, Decision No: 2025/436). Due to the retrospective nature of the study, patient consent was not obtained.

REFERENCES

1. Yokoe M, Hata J, Takada T, Strasberg SM, Asbun HJ, Wakabayashi G, et al. Tokyo Guidelines 2018: diagnostic criteria and severity grading of acute cholecystitis (with videos). J Hepatobiliary Pancreat Sci. 2018;25(1):41-54. https://doi.org/10.1002/jhbp.515
2. Gallaher JR, Charles A. Acute Cholecystitis: A Review. JAMA. 2022;327(10):965-975. https://doi.org/10.1001/jama.2022.2350
3. Gurusamy KS, Davidson C, Gluud C, Davidson BR. Early versus delayed laparoscopic cholecystectomy for people with acute cholecystitis. Cochrane Database Syst Rev. 2013;(6):CD005440. https://doi.org/10.1002/14651858.CD005440.pub3
4. Gutt CN, Encke J, Köninger J, Harnoss JC, Weigand K, Kipfmüller K, et al. Acute cholecystitis: early versus delayed cholecystectomy, a multicenter randomized trial (ACDC study, NCT00447304). Ann Surg. 2013;258(3):385-393. https://doi.org/10.1097/SLA.0b013e3182a1599b
5. Borzellino G, Khuri S, Pisano M, Mansour S, Allievi N, Ansaloni L, et al. Timing of early laparoscopic cholecystectomy for acute calculous cholecystitis revised: Protocol of a systematic review and meta-analysis of results. World J Emerg Surg. 2020;15:1. https://doi.org/10.1186/s13017-019-0285-7
6. Blohm M, Österberg J, Sandblom G, Lundell L, Hedberg M, Enochsson L. The Sooner, the Better? The Importance of Optimal Timing of Cholecystectomy in Acute Cholecystitis: Data from the National Swedish Registry for Gallstone Surgery, GallRiks. J Gastrointest Surg. 2017;21(1):33-40. https://doi.org/10.1007/s11605-016-3223-y
7. Cao AM, Eslick GD, Cox MR. Early Cholecystectomy Is Superior to Delayed Cholecystectomy for Acute Cholecystitis: a Meta-analysis. J Gastrointest Surg. 2015;19(5):848-857. https://doi.org/10.1007/s11605-015-2747-x
8. Mou D, Tesfasilassie T, Hirji S, Ashley SW. Advances in the management of acute cholecystitis. Ann Gastroenterol Surg. 2019;3(3):247-253. https://doi.org/10.1002/ags3.12240
9. Pathak SJ, Ji H, Nijagal A, Avila P, Dai SC, Arain MA, et al. Index admission cholecystectomy for acute cholecystitis reduces 30-day readmission rates in pediatric patients. Surg Endosc. 2024;38(3):1351-1357. https://doi.org/10.1007/s00464-023-10632-7

10. Köstenbauer JK, Gandy RC, Close J, Harvey L. Factors Affecting Early Cholecystectomy for Acute Cholecystitis in Older People-A Population-Based Study. World J Surg. 2023;47(7):1704-1710. https://doi.org/10.1007/s00268-023-06968-9
11. Hahm JS, Park JY, Park KG, Ahn YH, Lee MH, Park KN. Gallbladder motility in diabetes mellitus using real time ultrasonography. Am J Gastroenterol. 1996;91(11):2391-2394.
12. Cho JY, Han HS, Yoon YS, Ahn KS. Risk factors for acute cholecystitis and a complicated clinical course in patients with symptomatic cholelithiasis. Arch Surg. 2010;145(4):329-333. https://doi.org/10.1001/archsurg.2010.35
13. Ratheesh R, Ulrich MT, Ghozy S, Al-Jaboori M, Nayak SS. The association between diabetes and gallstones: a nationwide population-based cohort study. Prz Gastroenterol. 2023;18(3):292-299. https://doi.org/10.5114/pg.2023.131395
14. Mahmood F, Akingboye A, Malam Y, Thakkar M, Jambulingam P. Complicated Acute Cholecystitis: The Role of C-Reactive Protein and Neutrophil-Lymphocyte Ratio as Predictive Markers of Severity. Cureus. 2021;13(2):e13592. https://doi.org/10.7759/cureus.13592
15. Shabanzadeh DM, Sørensen LT, Jørgensen T. Determinants for symptomatic gallstone disease readmissions - results from a cohort with screen-detected gallstone disease. J Visc Surg. 2019;156(5):387-396. https://doi.org/10.1016/j.jviscsurg.2019.02.005
16. Velamazán R, López-Guillén P, Martínez-Domínguez SJ, Abad Baroja D, Oyón D, Arnau A, et al. Symptomatic gallstone disease: Recurrence patterns and risk factors for relapse after first admission, the RELAPSTONE study. United European Gastroenterol J. 2024;12(3):286-298. https://doi.org/10.1002/ueg2.12544
17. Salama A, Calpin GG, Fuller R, Hill ADK. Clinical predictors of recurrent cholecystitis in non-operative management: A systematic review & meta-analysis. Surgeon. 2025;23(2):106-113. https://doi.org/10.1016/j.surge.2024.11.004

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):39-48
DOI: 10.29058/mjwbs.1754110

Investigation of the effects of odontoid bone anatomical location and structure on cervical lordosis angle in cervical CT examinations: A clinical study

Servikal BT incelemelerinde odontoid kemik anatomik yerleşiminin ve yapısının servikal lordoz açısı üzerine olan etkilerinin araştırılması: Bir klinik araştırma

Alemiddin Özdemir[1], Ahmet Melih Erdoğan[1], Özge Sevimoğlu[1], Buket Oğuz[2], Buse Kaymakcı[3], Selma Çalışkan[2], Bülent Bakar[1]

[1]Department of Neurosurgery, Kırıkkale University, Faculty of Medicine, Kırıkkale, Türkiye
[2]Department of Anatomy, Yıldırım Beyazıt University, Ankara, Türkiye
[3]Department of Radiology, Bağcılar Training and Research Hospital, Istanbul, Türkiye

Cite this article as: *Özdemir A et al. Investigation of the effects of odontoid bone anatomical location and structure on cervical lordosis angle in cervical CT examinations: A clinical study. Med J West Black Sea. 2026;10(1):39-48*

This study was presented as an oral presentation (SS-056) at the 37th Scientific Congress of the Turkish Neurosurgical Society, held at Pine Beach Congress Center, Belek, Antalya, on April 18–21, 2024.

ABSTRACT

Aim: In literature, anatomical variations of the odontoid process in the sagittal plane have been suggested to influence the magnitude of cervical lordosis. This study investigated the effects of odontoid process morphology on the degree of cervical lordosis.

Material and Methods: This retrospective study included individuals who underwent cervical computed tomography (CT) scans and had no cervical vertebral disease. The age and gender of these individuals were recorded. Their CT images were analyzed for the morphologic relationships and morphologic measurements of the C2 vertebra and other cervical vertebrae in relation to their adjacent formations.

Results: Thirty-five subjects (male=23, female=12) were included in the study. Correlation analysis revealed a positive correlation between C2-C7 Cobb angle and age, sagittal odontoid height, and sagittal C2 vertebral height values. In addition, positive correlations were found between sagittal C2-C7 height values and sagittal odontoid mid-thickness, sagittal odontoid base thickness, and C2 vertebral height values. Linear Regression analysis revealed that the T1 slope angle, sagittal C2 vertebral height, and sagittal C2-C7 height values could predict the C2-C7 Cobb angle.

Conclusion: These findings suggest that cervical lordosis may decrease over time when the T1 slope angle, sagittal odontoid height, sagittal C2 vertebral height, and sagittal C2-C7 height are measured low, and thus, those individuals may be more exposed to cervical vertebral alignment disorders.

Keywords: Odontoid process, C2 vertebra, cervical lordosis, cobb angle

ÖZ

Amaç: Literatürde, sagittal düzlemdeki odontoid çıkıntının anatomik varyasyonlarının servikal lordozun derecesini etkilediği öne sürülmüştür. Bu çalışmada odontoid çıkıntı morfolojisinin servikal lordoz derecesi üzerindeki etkileri araştırılmıştır.

Gereç ve Yöntemler: Bu retrospektif çalışmaya servikal bilgisayarlı tomografi (BT) taramaları yapılan ve servikal disk hernisi, servikal dar kanal ya da servikal vertebra hastalığı olmayan bireyler dahil edilmiştir. Katılımcıların yaşları ve cinsiyetleri kaydedilmiştir. Ayrıca bireylerin servikal BT görüntülerinde C2 vertebra ve diğer servikal vertebraların kendi aralarında ve komşu oluşumlarla olan morfolojik ilişkileri ve ölçümleri analiz edilmiştir.

Bulgular: Çalışmaya 35 birey (erkek=23, kadın=12) dahil edildi. Korelasyon analizi, C2-C7 Cobb açısı ile yaş, sagittal odontoid yüksekliği ve sagittal C2 vertebra yüksekliği değerleri arasında pozitif bir korelasyon olduğunu ortaya koydu. Ek olarak, sagittal düzlemde C2-C7 yükseklik değerleri ile sagittal düzlemde odontoid çıkıntı orta kalınlığı, sagittal düzlemde odontoid çıkıntı taban kalınlığı ve sagittal düzlemde C2 vertebra yüksekliği değerleri arasında pozitif korelasyon bulundu. Doğrusal Regresyon analizi, T1 eğim açısı, sagittal düzlemde C2 vertebral yüksekliği ve sagittal düzlemde C2-C7 yüksekliği değerlerinin C2-C7 Cobb açısını tahmin edebileceğini ortaya koydu.

Sonuç: Bu bulgular, T1 eğim açısı, sagittal düzlemde odontoid çıkıntı yüksekliği, sagittal düzlemde C2 vertebra yüksekliği ve sagittal düzlemde C2-C7 yüksekliği düşük ölçüldüğünde servikal lordozun zamanla azalabileceğini ve dolayısıyla bu bireylerin servikal vertebral hizalama bozukluklarına daha fazla maruz kalabileceğini düşündürdü.

Anahtar Kelimeler: Odontoid çıkıntı, C2 vertebra, servikal lordoz, cobb açısı

Corresponding Author: Alemiddin Özdemir ✉ alemozdemir@yahoo.com **Received:** 15.09.2025 **Revision:** 24.09.2025 **Accepted:** 07.10.2025

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Odontoid and C2 vertebral morphology constitute key structural determinants of cervical lordosis.
- T1 slope, sagittal odontoid height, and sagittal C2 vertebral height show significant positive correlations with the C2–C7 Cobb angle.
- Reduced values of these morphometric parameters may indicate a predisposition to decreased cervical lordosis and progressive kyphotic alignment.
- Age-related changes in cervical alignment appear closely linked to the morphometric characteristics of the odontoid–C2 complex.
- These findings highlight the clinical importance of incorporating odontoid and C2 measurements into routine cervical spine assessments to enhance predictive accuracy.

INTRODUCTION

Recent studies have shown that cervical alignment is influenced by global body sagittal alignment via compensatory mechanisms that maintain an upright posture and horizontal gaze (1–3). However, studies investigating the relationship between cervical alignment and thoracolumbar spine and pelvic alignment have failed to establish a direct correlation between thoracic kyphosis, lumbar lordosis, and cervical lordosis. These studies have reported that lumbar lordosis influences thoracic kyphosis, but has only an indirect effect on cervical alignment (4-6). Unfortunately, no valid definition of 'normal' cervical alignment has been provided. Although lordosis is considered a component of normal cervical alignment, it is well known that a significant proportion of asymptomatic cervical spines are not lordotic (7, 8).

Studies have suggested that the C2 vertebra marks the boundary between the subaxial and craniovertebral junctions, and that odontoid process morphology is related to sagittal cervical spine alignment (9). Some even suggest that the C2 vertebra acts as the base of the subaxial and craniovertebral junction (CVJ), contributing to the orientation and structural geometry of physiologic sagittal alignment, with a correlation between the structural geometry of the C2 vertebra, the odontoid process, and the sagittal alignment of the cervical spine (9, 10). Although the morphometry of CVJ bony structures has been defined in the paediatric population, studies investigating the impact of odontoid process morphology on cervical lordosis in adults remain limited (11). Therefore, little information was found in the literature regarding this anatomical configuration.

This study aimed to analyse the relationship between the structural geometry of the odontoid process and the sagittal alignment of the cervical spine in an asymptomatic population, and to determine the potential impact of these factors on cervical lordosis.

MATERIAL and METHODS

This retrospective clinical study was conducted after obtaining the approval of the Kırıkkale University Non-Interventional Research Ethics Committee (Decision date: 10.01.2024, decision number: 2024.01.09).

Participants

The study included individuals who were admitted to the outpatient clinic between July 2021 and December 2023, underwent cervical CT scans, and had no cervical vertebrae or disc disease. Data on these individuals were obtained from the hospital's digital patient record database. The individuals were then grouped according to their age as follows:

- Age <50 years group (n=20).
- Age >50 years group (n = 15).

Participants were also grouped according to gender as follows:

- Male patients (n = 23)
- Female patients (n = 12).

Furthermore, the age groups were divided into gender groups.

Participants with a cervical spine fracture or trauma, spinal surgery, infection, primary or metastatic vertebral tumour, cervical disc herniation, rheumatological disease, spondylosis, spondylolisthesis, spinal deformity or congenital anomaly, or cervical spinal stenosis were excluded from the study. The paediatric age group was also excluded from the study.

Morphological Measurements

The following morphologic measurements were made of the C2 vertebra and other cervical vertebrae, as well as their relationships with the adjacent formations, especial-

ly the clivus and T1 vertebra, in the sagittal, coronal, and axial CT images: Clivus posterior wall–odontoid posterior wall angle; clivus anterior wall–odontoid posterior wall angle; clivus posterior wall–odontoid tip–occipital bone angle; clivus anterior wall–odontoid tip–occipital bone angle; C2 slope angle; odontoid tip–C2 vertebral lower endplate angle; C2–C7 Cobb angle; T1 slope angle; and occipital condyles–odontoid tip angle (Figure 1) (12). Additionally, the following length measurements were taken from the CT images: odontoid sagittal mid-thickness, odontoid sagittal base thickness, odontoid coronal mid-thickness, odontoid coronal base thickness, odontoid sagittal height, C2 sagittal height, C2–C7 sagittal height, the axial thickness of the joint between C1 and the odontoid process and the length between the clivus and the occipital bone (Figure 2).

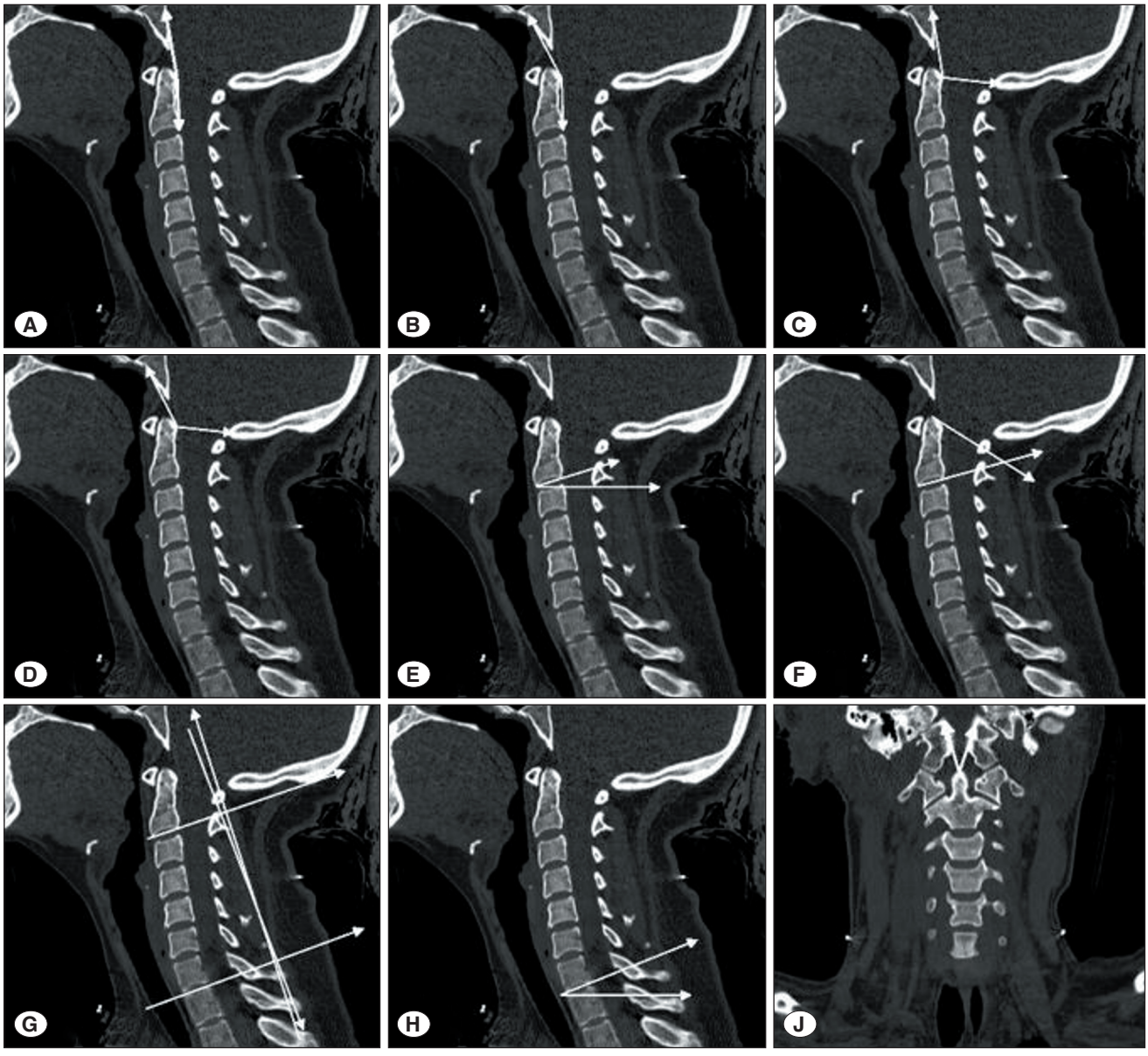


Figure 1: The pictures show the measurement techniques for the following angles: **(A)** the angle between the posterior clivus and the posterior odontoid; **(B)** the angle between the anterior clivus and the posterior odontoid; **(C)** the angle between the posterior clivus and the occipital bone's inferior tip; **(D)** the angle between the anterior clivus and the odontoid crest and the inferior tip of the occipital bone; **(E)** the angle between the C2 lower endplate and the horizontal plane (C2 slope); **(F)** the angle between the odontoid crest and the C2 lower endplate; **(G)** the Cobb angle between the C2 and C7 vertebrae; **(H)** angle between the T1 upper end-plate and the horizontal plane (T1 slope) and **(J)** angle between the occipital condyle-odontoid tip-occipital condyle.

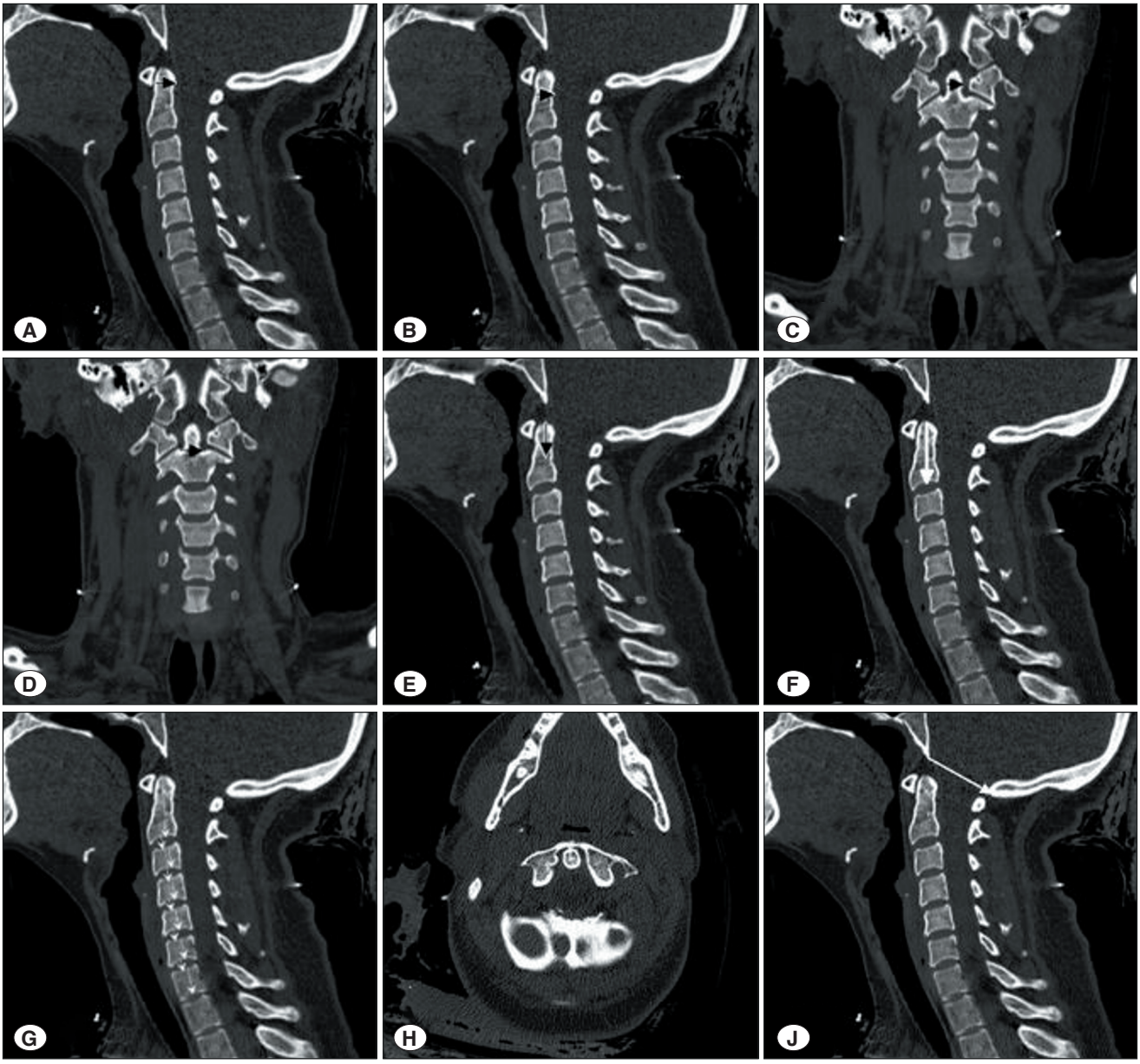


Figure 2: The pictures show the measurement techniques for **(A)** odontoid mid-thickness in the sagittal plane; **(B)** odontoid base thickness in the sagittal plane; **(C)** odontoid mid-thickness in the coronal plane; **(D)** odontoid base thickness in the coronal plane; **(E)** odontoid sagittal height; **(F)** C2 vertebra sagittal height; **(G)** C2–C7 vertebrae sagittal height; **(H)** width of the joint between C1 and the odontoid in the axial plane; and **(J)** length between the lower end of the clivus and the lower end of the occipital bone in the sagittal plane.

Statistical Analysis

A power analysis was performed on the study results using G*Power 3.1 software. It was concluded that the number of individuals included in the study constituted an adequate sample size (effect size dz=0.750, critical t=2.060, actual power=0.957, total sample size=26).

The Pearson chi-squared test was used to evaluate differences in categorical data between groups ($p < 0.05$). The *Mann–Whitney U test* was used to compare non-parametric data between groups ($p < 0.05$). An *independent samples t-test* was used to evaluate differences in parametric data between groups ($p < 0.05$). *Spearman's rho correlation* test was used to determine correlations between parameters ($p < 0.05$). A *linear regression* was performed to determine which study parameters could predict a decrease in the cervical lordosis angle ($p < 0.05$).

RESULTS

A total of 35 participants (23 males and 12 females) were included in the study (Table 1). C2-C7 Cobb angle (t=-2.250, p=0.024), odontoid coronal base thickness (t=2.185, p=0.036), C2-C7 sagittal height (t=2.685, p=0.011), and axial thickness of the joint between C1 and odontoid process (t=2.369, p=0.024) were found to be different between age groups(Table 2).

In addition, when participants under 50 years of age were grouped by gender, clivus anterior wall-odontoid tip-occipital bone angle (t=-2.701, p=0.015) and C2-C7 sagittal height values (t=3.520, p=0.002) were found to be different between male and female participants. On the other hand, when people over 50 years of age were grouped according to gender, the angle between the occipital condyles and the odontoid tip (t=2.671, p=0.020) measured in the coronal plane, and the odontoid base thickness measured in the sagittal plane (t=2.782, p=0.016) were different between the genders (Table 3).

The clivus anterior wall-odontoid tip-occipital bone angle values (t=-2.462, p=0.019), odontoid sagittal mid-thickness (t=2.541, p=0.016) and odontoid sagittal base thickness values (t=3.219, p=0.003), C2 vertebra sagittal height values (t=2.363, p=0.018) and C2-C7 sagittal height values (t=3.654, p=0.001) were different between the male and female participants (Table 4).

Correlation analysis revealed a positive correlation between odontoid sagittal base thickness and C2 vertebra sagittal height values (r=0.402, p=0.017), and between odontoid sagittal base thickness and C2-C7 sagittal height values (r=0.400, p=0.017), and a negative correlation between odontoid sagittal base thickness and C2 slope angle values (r=-0.457, p=0.006) and between odontoid sagittal base thickness values and gender (r=-0.432, p=0.010). In addition, there was a positive correlation between odontoid sagittal mid-thickness and C2-C7 sagittal height values (r=0.435, p=0.009) and a negative correlation between odontoid sagittal mid-thickness values and gender (r=-0.408, p=0.015). On the other hand, the odontoid coronal base thickness values were positively correlated with C2 vertebral sagittal height values (r=0.349, p=0.040) and C2-C7 sagittal height values (r=0.378, p=0.025). There was a positive correlation between sagittal odontoid height and C2 vertebral height

Table 1: Descriptive table of data for all participants.

Variable		Min	Max	Mean/ Median/ N (%)	SD
Age (year)		19.00	80.00	47.20*	16.54
Sex	Male			23 (65.7%)‡	
	Female			12 (34.3%)‡	
Clivus posterior wall-odontoid posterior wall angle		142.50	194.80	163.35*	10.04
Clivus anterior wall-odontoid posterior wall angle		115.70	174.60	144.44*	12.30
Clivus posterior wall-odontoid tip-occipital bone angle		100.00	146.20	123.51*	10.93
Clivus anterior wall-odontoid tip-occipital bone angle		88.30	119.90	105.11*	7.90
C2 slope angle		1.10	42.10	13.40†	10.05
Odontoid tip-C2 vertebral lower end-plate angle		45.60	85.00	65.58*	9.81
C2-C7 Cobb angle		0.10	57.70	14.70†	12.23
T1 slope angle		0.70	54.00	28.18*	10.30
Occipital condyle-odontoid tip angle		26.40	112.15	62.66*	27.57
Odontoid sagittal mid-thickness		7.20	14.03	10.29*	1.33
Odontoid sagittal base thickness		7.38	12.86	10.87*	1.24
Odontoid coronal mid-thickness		6.67	10.84	8.79*	1.11
Odontoid coronal base thickness		5.63	12.82	9.10*	1.55
Odontoid sagittal height		11.40	21.90	16.42*	2.45
C2 vertebra sagittal height		30.01	41.80	33.91†	2.77
C2-C7 vertebrae sagittal height		97.27	132.37	114.31*	9.03
Axial thickness of the joint between C1 and odontoid		0.39	1.95	1.01*	0.38
Length between the clivus and occipital bone		24.60	42.73	35.39†	3.65

(*) mean value, (†) median value, (‡) number of patients

Min: minimum, **max:** maximum, **SD:** standard deviation

Table 2: Descriptive and comparative table of participants' data according to age group.

Variable		Age <50 Mean ± SD/ Median (min-max)/ n (%)	Age >50 Mean ± SD/ Median (min-max)/ n (%)	t/ Z/ X_2	p
Sex	Male	13 (37.1%)	10 (28.6%)	0.011‡	0.918
	Female	7 (20.0%)	5 (14.3%)		
Clivus posterior wall-odontoid posterior wall angle		161.49±8.27	165.83±11.85	-1.279*	0.210
Clivus anterior wall-odontoid posterior wall angle		143.12±11.94	146.21±12.97	-0.731*	0.470
Clivus posterior wall-odontoid tip-occipital bone angle		120.63±10.58	127.35±10.52	-1.866*	0.071
Clivus anterior wall-odontoid tip-occipital bone angle		104.54±7.83	105.87±8.20	-0.489*	0.628
C2 slope angle		12.85 (1.40-164.70)	15.90 (1.10-167.70)	-0.067†	0.947
Odontoid tip-C2 vertebral lower end-plate angle		65.14±9.78	66.15±10.16	-0.297*	0.768
C2-C7 Cobb angle		10.60 (0.10-57.709	18.70 (4.20-46.80)	-2.250†	**0.024**
T1 slope angle		27.82±7.95	28.65±12.98	-0.230*	0.819
Occipital condyle-odontoid tip angle		60.26±28.43	66.10±26.96	-0.602*	0.551
Odontoid sagittal mid-thickness		10.42±1.47	10.11±1.14	0.678*	0.503
Odontoid sagittal base thickness		10.97±1.14	10.74±1.39	0.541*	0.592
Odontoid coronal mid-thickness		8.74±1.24	8.85±0.93	-0.297*	0.769
Odontoid coronal base thickness		9.57±1.50	8.47±1.43	2.185*	**0.036**
Odontoid sagittal height		16.29±1.74	16.60±3.22	-0.363*	0.719
C2 vertebra sagittal height		33.93 (30.35-40.27)	33.86 (30.01-41.80)	-0.517†	0. 605
C2-C7 vertebrae sagittal height		117.57±9.30	109.96±6.70	2.685*	**0.011**
Axial thickness of the joint between C1 and odontoid		1.13±0.38	0.84±0.32	2.369*	**0.024**
Length between the clivus and occipital bone		35.40 (26.90-42.73)	35.38 (24.60-39.20)	-0.633†	0. 527

(*) t value, Independent Samples t test; (†) Z value, Mann Whitney U test; (‡) X^2 value, Pearson Chi-square test; p<0.05

min: minimum, **max:** maximum, **n:** number of participants, **SD:** standard deviation

values (r=0.420, p=0.012), and between sagittal odontoid height and C2-C7 Cobb angle values (r=0.391, p=0.020), and a negative correlation between sagittal odontoid height and C2 slope angle values (r=-0.457, p=0.006). There was also a positive correlation between sagittal C2 vertebral height and C2-C7 vertebrae sagittal height values (r=0.540, p=0.001) and between sagittal C2 vertebral height and C2-C7 Cobb angle values (r=0.339, p=0.046), and a negative correlation between C2 vertebral height values and gender (r=-0.405, p=0.016). Finally, C2-C7 height values were negatively correlated with age (r=-0.335, p=0.049) and gender (r=-0.528, p=0.001).

Linear Regression analysis revealed that the T1 slope angle values (B=0.448, SE=0.170, beta=0.377, t=2.628, p=0.013), C2 vertebra sagittal height values (B=3.122, SE=0.731, beta=0.717, t=4. 269, p<0.001) and the C2-C7 vertebra sagittal height values (B=-0.891, SE=0.247, beta=-0.629, t=-3.614, p=0.001) could predict the cervical lordosis angle (Table 5).

DISCUSSION

The cervical spine supports the skull and keeps a horizontal gaze (13). Therefore, changes that maintain horizontal gaze can alter cervical and sagittal spine alignment, resulting in conditions such as thoracic kyphosis and sacral tilt (14). These changes may cause neck and back muscle fatigue, as well as neck and back pain over time (15). Conversely, studies suggest that age-related increases in C2-C7 lordosis and thoracic kyphosis may lead to changes in the T1 slope angle (16, 17).

This study found that odontoid mid-thickness and base thickness values, as well as C2-C7 vertebral height and C2 sagittal height values measured in the sagittal plane, were higher in male participants than in females. These findings may originate from androgenic factors in males. Additionally, when participants were divided into two age groups, it was observed that the angle between the occipital condyles and the odontoid tip, as well as the joint axial thickness measured between C1 and the odontoid process, were

Table 3: Descriptive and comparative table of data for age groups, classified by gender.

	Male (n=23)	Female (n=12)		
Variable	**Mean ± SD/ Median (min-max)**	**Mean ± SD/ Median (min-max)**	**t/ Z**	**p**
Age (year)	48.48±16.32	44.75±17.39	0.627*	0.535
Clivus posterior wall-odontoid posterior wall angle	163.52±8.37	163.03±13.09	0.136*	0.893
Clivus anterior wall-odontoid posterior wall angle	145.57±11.76	142.28±13.54	0.747*	0.460
Clivus posterior wall-odontoid tip-occipital bone angle	121.33±11.31	127.68±9.22	-1.671*	0.104
Clivus anterior wall-odontoid tip-occipital bone angle	102.90±7.39	109.35±7.32	-2.462*	**0.019**
C2 slope angle	11.90 (1.10-167.70)	15.75 (6.20-160.70)	-0.730†	0.466
Odontoid tip-C2 vertebral lower end-plate angle	66.66±8.50	63.51±12.06	0.899*	0.375
C2-C7 Cobb angle	15.50 (0.10-57.70)	11.95 (4.20-33.20)	-0.817†	0.414
T1 slope angle	29.85±10.10	25.12±10.38	1.295*	0.205
Occipital condyle-odontoid tip angle	64.58±28.38	59.14±26.86	0.544*	0.590
Odontoid sagittal mid-thickness	10.67±1.14	9.56±1.40	2.541*	**0.016**
Odontoid sagittal base thickness	11.30±0.87	10.04±1.44	3.219*	**0.003**
Odontoid coronal mid-thickness	8.99±1.00	8.39±1.23	1.574*	0.125
Odontoid coronal base thickness	9.38±1.46	8.55±1.63	1.540*	0.133
Odontoid sagittal height	16.56±2.59	16.16±2.24	0.448*	0.657
C2 vertebra sagittal height	34.28 (31.03-41.80)	33.19 (30.01-34.65)	-2.363†	**0.018**
C2-C7 vertebrae sagittal height	117.78±8.17	107.70±6.76	3.654*	**0.001**
Axial thickness of the joint between C1 and odontoid	1.02±0.33	0.99±0.47	0.190*	0.851
Length between the clivus and occipital bone	35.68 (26.90-40.38)	34.47 (24.60-42.73)	-1.529†	0.126

(*) t value, Independent Samples t test; (†) Z value, Mann Whitney U test; p<0.05

Min: minimum, **max:** maximum, **SD:** standard deviation

both lower in participants over 50 years old; however, the C2-C7 Cobb angle was higher. These results imply that the craniovertebral junction and joint spaces between C1 and the odontoid process may narrow slightly with age, while the C2-C7 Cobb angle increases, resulting in reduced cervical lordosis. However, contrary to previous studies, there was no difference in C2 slope, T1 slope, or C2-C7 vertebral sagittal height measurements between the two age groups. Furthermore, when the age groups were subdivided by gender, it was found that C2-C7 vertebral sagittal height measurements were significantly greater in men under 50 than in women of the same age. Additionally, odontoid base thickness values and the angle between the occipital condyles and the odontoid tip were greater in males over 50 than in females. Nevertheless, no significant differences were observed between the groups regarding C2-C7 Cobb angles or other morphological measurements of the odontoid process.

The correlation analysis revealed that high C2 vertebral height and C2-C7 vertebral sagittal height values might be indicative of high odontoid base thickness values measured in the sagittal and coronal planes. Additionally, it was hypothesised that a greater odontoid base thickness in the sagittal plane could result in a lower C2 slope angle. Conversely, it was hypothesised that the odontoid base thickness might be thinner in women. Furthermore, it was observed that the sagittal height of the C2-C7 vertebrae might be high if greater odontoid mid-thickness values were measured in the sagittal plane. Interestingly, it was hypothesised that C2 vertebral sagittal height and C2–C7 Cobb angle values would be high when odontoid height values measured in the sagittal plane were high. It was also thought that C2-C7 vertebral sagittal height values would be high when C2 vertebral sagittal height values were high, but low in women. Therefore, it was suggested that the thickness and height of the odontoid process could affect the sagittal height of the C2 vertebra, the C2-C7 Cobb angle, and the sagittal height of the C2-C7 vertebrae. In conclusion, it was argued that the thickness of the odontoid process, the sagittal height of the C2 vertebra, and the sagittal height of the C2-C7 vertebrae could directly or indirectly affect the cervical lordosis angle.

Linear regression analysis showed that C2 vertebral sagittal height, T1 slope angle, and C2–C7 vertebral sagittal height values could predict progression to cervical lordosis or kyphosis. It was suggested that when these values were

Table 4: Descriptive and comparative table of participants' data according to gender groups.

		Male	Female		
	Variable	**Mean ± SD/ Median (min-max)**	**Mean ± SD/ Median (min-max)**	**t/ Z**	**p**
Age <50	Clivus posterior wall-odontoid posterior wall angle	163.16±7.03	158.37±10.03	1.253*	0.226
	Clivus anterior wall-odontoid posterior wall angle	144.76±12.71	140.06±10.53	0.834*	0.415
	Clivus posterior wall-odontoid tip-occipital bone angle	118.35±12.14	124.84±5.26	-1.335*	0.198
	Clivus anterior wall-odontoid tip-occipital bone angle	101.53±7.51	110.13±5.04	-2.701*	**0.015**
	C2 slope angle	12.30 (1.40-35.35)	13.40 (6.20-42.10)	-0.277†	0.782
	Odontoid tip-C2 vertebral lower end-plate angle	67.25±9.00	61.24±10.65	1.336*	0.198
	C2-C7 Cobb angle	9.50 (0.01-57.70)	11.70 (5.80-16.10)	-0.357†	0.721
	T1 slope angle	27.75±7.94	27.92±8.60	-0.046*	0.964
	Occipital condyle-odontoid tip angle	55.35±27.58	69.37±29.82	-1.055*	0.305
	Odontoid sagittal mid-thickness	10.85±1.18	9.63±1.70	1.887*	0.075
	Odontoid sagittal base thickness	11.29±0.80	10.37±1.47	1.804*	0.088
	Odontoid coronal mid-thickness	8.92±1.04	8.40±1.59	0.871*	0.395
	Odontoid coronal base thickness	9.82±1.73	9.10±0.81	1.028*	0.317
	Odontoid sagittal height	16.36±1.56	16.16±2.15	0.240*	0.813
	C2 vertebra sagittal height	34.29 (32.14-40.27)	33.52 (30.35-34.56)	-1.704†	0.088
	C2-C7 vertebrae sagittal height	121.82±7.85	109.68±6.23	3.520*	**0.002**
	Axial thickness of the joint between C1 and odontoid	1.12±0.32	1.15±0.51	-0.179*	0.860
	Length between the clivus and occipital bone	35.77 (26.90-42.39)	34.96 (31.29-42.73)	-0.357†	0.721
Age >50	Clivus posterior wall-odontoid posterior wall angle	163.98±10.26	169.54±15.14	-0.848*	0.412
	Clivus posterior wall-odontoid anterior wall angle	146.62±10.96	145.38±17.80	0.168*	0.869
	Clivus posterior wall-odontoid tip-occipital bone angle	125.21±9.30	131.64±12.59	-1.126*	0.280
	Clivus anterior wall-odontoid tip-occipital bone angle	104.67±7.22	108.28±10.34	-0.793*	0.442
	C2 slope angle	10.80 (1.10-39.40)	16.07 (9.90-21.20)	-0.735†	0.462
	Odontoid tip-C2 vertebral lower end-plate angle	65.89±8.22	66.68±14.44	-0.137*	0.893
	C2-C7 Cobb angle	20.45 (11.60-46.80)	14.70 (4.20-33.20)	-0.980†	0.327
	T1 slope angle	32.38±12.17	21.18±12.31	1.674*	0.118
	Occipital condyle-odontoid tip angle	77.92±25.21	44.82±14.49	2.671*	**0.020**
	Odontoid sagittal mid-thickness	10.44±1.10	9.45±1.01	1.683*	0.116
	Odontoid sagittal base thickness	11.32±1.00	9.58±1.41	2.782*	**0.016**
	Odontoid coronal mid-thickness	9.09±0.99	8.36±0.60	1.506*	0.156
	Odontoid coronal base thickness	8.81±0.74	7.78±2.24	1.358*	0.197
	Odontoid sagittal height	16.82±3.60	16.17±2.63	0.356*	0.728
	C2 vertebra sagittal height	34.10 (31.03-41.80)	32.78 (30.01-34.65)	-1.715†	0.086
	C2-C7 vertebrae sagittal height	112.48±5.12	104.92±7.13	2.374*	0.034
	Axial thickness of the joint between C1 and odontoid	0.88±0.32	0.76±0.34	0.671*	0.514
	Length between the clivus and occipital bone	35.63 (32.21-39.20)	29.69 (24.60-36.46)	-1.960†	0.050

(*) t value, Independent Samples t test; (†) Z value, Mann Whitney U test; p<0.05
min: minimum, **max:** maximum, **SD:** standard deviation

Table 5: The linear regression test reveals the study parameters that could predict a decrease in cervical lordosis angle.

						95% CI	
Variable	**B**	**SE**	**Beta**	**t**	**p**	**Lower**	**Upper**
T1 slope angle	0.448	0.170	0.377	2.628	**0.013**	0.100	0.796
C2 vertebra sagittal height	3.122	0.731	0.717	4.269	**<0.001**	1.628	4.615
C2-C7 vertebrae sagittal height	-0.891	0.247	-0.629	-3.614	**0.001**	-1.395	-0.388

Linear Logistic Regression test, p<0.05

measured as low, the cervical Cobb angle and cervical lordosis would decrease, and cervical kyphosis would progress, meaning these individuals might be more susceptible to cervical vertebral alignment disorders. The results also showed that the odontoid process, C2 sagittal height, and T1 slope angle values could directly affect the formation of the cervical lordosis angle, and that morphological changes to these spinal structures could also alter the cervical lordosis angle.

In light of these findings, it was hypothesised that the following parameters may directly or indirectly affect cervical vertebral alignment and be effective in the formation of cervical lordosis or kyphosis angles: age; gender; odontoid height and thickness values; C2 vertebra sagittal height values; C2-C7 vertebrae sagittal height values; C2 slope; and T1 slope angle values.

Limitations

This retrospective study has several limitations. *Firstly*, the study was conducted in a single centre, so the results may not be representative of the general population due to the small number of participants. *Secondly*, this study did not involve taking measurements on standing cervical radiographs. It is well known that simple radiographs are the most important tool for assessing spinal alignment. However, there are still no established radiographic methods to evaluate the sagittal balance of the cervical spine. Changes in arm position can cause significant differences in standing lateral radiographs of the entire spine. Furthermore, the supine position in axial imaging can potentially distort the weight-bearing neutral position of the cervical spine (18). Direct radiography is inadequate for providing three-dimensional visualization and measurement of the cervicothoracic junction, including the alignment of the thoracic inlet. This can be achieved using CT or MRI (19-21). Therefore, CT and/or MR scans provide a better visualisation of the components of the craniocervical and cervicothoracic junctions, the C7 lower-end plate, the T1 upper-end plate, and the tip of the manubrium in midsagittal images than conventional radiography (22). *Thirdly*, because they were outside the study scope, and this study was retrospective, some measurement parameters (such as thoracic kyphosis, lumbar lordosis, sacral slope, and sacral incidence) were neither measured nor compared in this study. Furthermore, normative values of the morphological measurements taken at the end of the study could not be determined. *Fourthly*, radiological measurements of individuals with cervical disc herniation or cervical spondylosis were not included in this study. *Finally*, the preoperative and postoperative measurement results of individuals who underwent surgery to correct cervical lordosis were excluded because they were outside the scope of the study. Despite these limitations, the results of this study are interesting and could inform future research.

Conclusion

Based on these findings, it was hypothesised that the C2-C7 Cobb angle would decrease with age. This would result in a decrease in cervical lordosis and an increase in cervical kyphosis, particularly in cases of low odontoid and C2 vertebral sagittal heights. It was also suggested that C2 and C2-C7 vertebral sagittal height and T1 slope angle values could predict changes in the cervical Cobb angle. It was considered that the cervical Cobb angle and cervical lordosis might decrease over time, especially when T1 slope angle and C2 sagittal height values were low, and that these individuals might be more susceptible to cervical vertebral alignment disorders.

Author Contributions

Study conception and design: **Bülent Bakar, Ahmet Melih Erdoğan, Buket Oğuz, Selma Çalışkan;** data collection: **Ahmet Melih Erdoğan, Özge Sevimoglu, Buse Kaymakcı, Alemiddin Özdemir;** analysis and interpretation of results: **Bülent Bakar, Ahmet Melih Erdoğan, Özge Sevimoglu;** draft manuscript preparation: **Bülent Bakar, Alemiddin Özdemir, Ahmet Melih Erdoğan.** The authors reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Financial Support

The authors declare that the study received no funding.

Ethical Approval

Ethical approval for this retrospective clinical study was obtained from the Kırıkkale University Non-Interventional Research Ethics Committee (Decision Date: January 10, 2024; Decision No: 2024.01.09).

REFERENCES

1. Kim JT, Lee HJ, Choi DY, Shin MH, Hong JT. Sequential alignment change of the cervical spine after anterior cervical discectomy and fusion in the lower cervical spine. Eur Spine J. 2016;25(7):2223-2232. https://doi.org/10.1007/s00586-016-4401-z
2. Le Huec JC, Saddiki R, Franke J, Rigal J, Aunoble S. Equilibrium of the human body and the gravity line: the basics. Eur Spine J. 2011;20 Suppl 5(Suppl 5):558-563. https://doi.org/10.1007/s00586-011-1939-7
3. Protopsaltis T, Bronsard N, Soroceanu A, Henry JK, Lafage R, Smith J, Klineberg E, Mundis G, Kim HJ, Hostin R, Hart R, Shaffrey C, Bess S, Ames C; International Spine Study Group. Cervical sagittal deformity develops after PJK in adult thoracolumbar deformity correction: radiographic analysis utilizing a novel global sagittal angular parameter, the CTPA. Eur Spine J. 2017;26(4):1111-1120. https://doi.org/10.1007/s00586-016-4653-7

4. Hey HWD, Lau ET, Wong GC, Tan KA, Liu GK, Wong HK. Cervical Alignment Variations in Different Postures and Predictors of Normal Cervical Kyphosis: A New Understanding. Spine (Phila Pa 1976). 2017;42(21):1614-1621. https://doi.org/10.1097/BRS.0000000000002160
5. Lee SH, Kim KT, Seo EM, Suk KS, Kwack YH, Son ES. The influence of thoracic inlet alignment on the craniocervical sagittal balance in asymptomatic adults. J Spinal Disord Tech. 2012;25(2):E41-E47. https://doi.org/10.1097/BSD.0b013e3182396301
6. Lee SH, Son ES, Seo EM, Suk KS, Kim KT. Factors determining cervical spine sagittal balance in asymptomatic adults: correlation with spinopelvic balance and thoracic inlet alignment. Spine J. 2015;15(4):705-712. https://doi.org/10.1016/j.spinee.2013.06.059
7. Hardacker JW, Shuford RF, Capicotto PN, Pryor PW. Radiographic standing cervical segmental alignment in adult volunteers without neck symptoms. Spine (Phila Pa 1976). 1997;22(13):1472-1480. https://doi.org/10.1097/00007632-199707010-00009
8. Kim SW, Kim TH, Bok DH, Jang C, Yang MH, Lee S, Yoo JH, Kwak YH, Oh JK. Analysis of cervical spine alignment in currently asymptomatic individuals: prevalence of kyphotic posture and its relationship with other spinopelvic parameters. Spine J. 2018;18(5):797-810. https://doi.org/10.1016/j.spinee.2017.09.008
9. Lee JK, Hyun SJ, Kim KJ. Odontoid Incidence: A Novel Cervical Parameter Influencing Cervical Alignment from Top to Bottom. Neurospine. 2022;19(2):463-471. https://doi.org/10.14245/ns.2244220.110
10. Lee SH, Hyun SJ, Jain A. Cervical Sagittal Alignment: Literature Review and Future Directions. Neurospine. 2020;17(3):478-496. https://doi.org/10.14245/ns.2040392.196
11. Kaya M, Ceylan D, Kaçıra T, Keskin E, Çelenk Y, Yaltırık Bilgin E, Kıtıki Kaçıra Ö. Measuring the shape and dimensions of normal the bony structures in the craniovertebral junction from computed tomography images of the pediatric age group. Ulus Travma Acil Cerrahi Derg. 2022;28(7):997-1007. https://doi.org/10.14744/tjtes.2022.45610
12. Ünsal ÜÜ, Şentürk S, Güvenç Y, Yaman O. Servikal Spinal Koronal-Sagital Balans ve Parametreler [Cervical Spine Coronal-Sagittal Balance and Parameters]. Türk Nöroşir Derg. 2022;32(3):366-372.
13. Chen Y, Luo J, Pan Z, Yu L, Pang L, Zhong J, Li Z, Han Z, Cao K: The change of cervical spine alignment along with aging in the asymptomatic population: a preliminary analysis. Eur Spine J. 2017;26(9):2363-2371. https://doi.org/10.1007/s00586-017-5209-1
14. Makhni MC, Shillingford JN, Laratta JL, Hyun SJ, Kim YJ: Restoration of Sagittal Balance in Spinal Deformity Surgery. J Korean Neurosurg Soc. 2018;61(2):167-179. https://doi.org/10.3340/jkns.2017.0404.013
15. Rajnics P, Templier A, Skalli W, Lavaste F, Illés T: The association of sagittal spinal and pelvic parameters in asymptomatic persons and patients with isthmic spondylolisthesis. J Spinal Disord Tech. 2002;15(1):24-30. https://doi.org/10.1097/00024720-200202000-00004
16. Park MS, Moon SH, Lee HM, Kim SW, Kim TH, Lee SY, Riew KD: The effect of age on cervical sagittal alignment: normative data on 100 asymptomatic subjects. Spine (Phila Pa 1976). 2013;38(8):E458-E463. https://doi.org/10.1097/BRS.0b013e31828802c2
17. Yokoyama K, Kawanishi M, Yamada M, Tanaka H, Ito Y, Kawabata S, Kuroiwa T: Age-related variations in global spinal alignment and sagittal balance in asymptomatic Japanese adults. Neurol Res. 2017;39(5):414-418. https://doi.org/10.1080/01616412.2017.1296654
18. Legaye J, Duval-Beaupere G: Influence of a variation in the position of the arms on the sagittal connection of the gravity line with the spinal structures. Eur Spine J. 2017;26(11):2828-2833. https://doi.org/10.1007/s00586-017-4961-6
19. Xing R, Zhou G, Chen Q, Liang Y, Dong J: MRI to measure cervical sagittal parameters: a comparison with plain radiographs. Arch Orthop Trauma Surg. 2017;137(4):451-455. https://doi.org/10.1007/s00402-017-2639-5
20. Jun HS, Chang IB, Song JH, Kim TH, Park MS, Kim SW, Oh JK: Is it possible to evaluate the parameters of cervical sagittal alignment on cervical computed tomographic scans? Spine (Phila Pa 1976). 2014;39(10):E630-E636. https://doi.org/10.1097/BRS.0000000000000281
21. Singhatanadgige W, Kang DG, Luksanapruksa P, Peters C, Riew KD: Correlation and Reliability of Cervical Sagittal Alignment Parameters between Lateral Cervical Radiograph and Lateral Whole-Body EOS Stereoradiograph. Global Spine J. 2016;6(6):548-554. https://doi.org/10.1055/s-0035-1569462
22. Oshina M, Tanaka M, Oshima Y, Tanaka S, Riew KD: Correlation and differences in cervical sagittal alignment parameters between cervical radiographs and magnetic resonance images. Eur Spine J. 2018;27(6):1408-1415. https://doi.org/10.1007/s00586-018-5550-z

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):49-59
DOI: 10.29058/mjwbs.1783805

Clinical and radiological response assessment after whole brain radiotherapy for patients with non-small cell lung cancer with brain metastases

Küçük hücreli dışı akciğer kanserli beyin metastazlı hastalarda tüm beyin radyoterapisi sonrası klinik ve radyolojik yanıt değerlendirme

Funda Çukurçayır[1], Mert Saynak[2], Canan Çelebi[3], Alaattin Özen[4], Görkem Türkkan[5], Kamuran İbiş[6], Mustafa Akın[1], Zafer Koçak[2]

[1]Department of Radiation Oncology, Balıkesir Atatürk City Hospital, Balıkesir, Türkiye
[2]Department of Radiation Oncology, Trakya University Hospital, Edirne, Türkiye
[3]Department of Neurology, Torbalı State Hospital, İzmir, Türkiye
[4]Department of Radiation Oncology, Başakşehir Çam and Sakura City Hospital, İstanbul, Türkiye
[5]Department of Radiation Oncology, Hisar Hospital Intercontinental, İstanbul, Türkiye
[6]Department of Radiation Oncology, İstanbul University, Faculty of Medicine, İstanbul, Türkiye

Cite this article as: *Çukurçayır F et al. Clinical and radiological response assessment after whole brain radiotherapy for patients with non-small cell lung cancer with brain metastases. Med J West Black Sea. 2026;10(1):49-59.*

ABSTRACT

Aim: This study aims to evaluate the response of brain metastases to whole-brain radiotherapy (WBRT) and the role of this treatment in influencing that response.

Material and Methods: The study involved eighteen patients who received WBRT for brain metastases either at the time of diagnosis or during the course of their disease. Magnetic resonance imaging (MRI) and neurological examinations were conducted 4 to 6 weeks and 8 to 10 weeks after the completion of WBRT.

Results: According to our radiologic assessment, two patients (17%) had a complete response, six patients (50%) had a partial response, and four patients (33%) had stable disease 8-10 weeks after WBRT. The lesion of two patients with solitary brain metastasis vanished, and one lesion of two patients with multiple metastases radiologically disappeared at 8-10 weeks after WBRT. During the initial measurement, the diameter of the biggest lesion was 0.6 to 4.3 cm (average 2.5 cm), 8-10 weeks after whole brain radiotherapy, the corresponding diameter was 0 to 3.5 cm (average 1.6 cm). When the tumor volumes were evaluated, the range was between 78 mm^3 and 623 mm^3. The volumes at diagnosis and first magnetic resonance imaging after radiotherapy were compared, and a significant response rate was obtained (p=0.008). After WBRT, a decrease in intensity or disappearance of the initial neurological symptoms and signs was seen in all patients.

Conclusion: In our study, WBRT provided a neurologic improvement in the patients with brain metastases, and 8-10 weeks after WBRT, a measurable reduction occurred in the size of the lesions in many of them. There is a need for studies with longer follow-ups and more patients focusing on the clinical utility of these findings.

Keywords: Lung cancer, brain metastasis, whole-brain radiotherapy, surgery, stereotactic radiosurgery

ÖZ

Amaç: Bu çalışmanın amacı beyin metastazlarının tüm beyin radyoterapisine (WBRT) yanıtını ve bu tedavinin bu yanıtı etkilemedeki rolünü değerlendirmektir.

Gereç ve Yöntemler: Çalışmaya, tanı anında veya hastalık seyri sırasında beyin metastazları nedeniyle WBRT uygulanan on sekiz hasta dahil edildi. Manyetik rezonans görüntüleme (MRG) ve nörolojik muayeneler, WBRT tamamlandıktan 4-6 hafta ve 8-10 hafta sonra gerçekleştirildi.

Bulgular: Radyolojik değerlendirmede, iki hastada (%17) tam yanıt, altı hastada (%50) kısmi yanıt ve dört hastada (%33) WBRT'den 8-10 hafta sonra stabil hastalık vardı. Soliter beyin metastazı olan iki hastanın lezyonu kayboldu ve multipl metastazı olan iki hastanın bir lezyonu WBRT'den 8-10 hafta sonra radyolojik olarak kayboldu. İlk ölçüm sırasında en büyük lezyonun çapı 0,6 ila 4,3 cm (ortalama 2,5 cm) iken, tüm beyin radyoterapisinden 8-10 hafta sonra buna karşılık gelen çap 0 ila 3,5 cm (ortalama 1,6 cm) idi. Tümör hacimleri değerlendirildiğinde, aralık 78 mm^3 ile 623 mm^3 arasındaydı. Tanı anındaki ve radyoterapi sonrası ilk manyetik rezonans görüntülemedeki hacimler karşılaştırıldığında, anlamlı yanıt oranı elde edildi (p=0,008). WBRT sonrası hastaların tamamında başlangıçtaki nörolojik semptom ve bulguların şiddetinin azaldığı veya kaybolduğu görüldü.

Sonuç: Çalışmamızda, WBRT beyin metastazlı hastalarda nörolojik iyileşme sağlamış ve WBRT'den 8-10 hafta sonra, çoğu hastada lezyonların boyutunda ölçülebilir bir azalma meydana gelmiştir. Bu bulguların klinik faydasına odaklanan, daha uzun takip süreli ve daha fazla hasta içeren çalışmalara ihtiyaç vardır.

Anahtar Kelimeler: Akciğer kanseri, beyin metastazı, tüm beyin radyoterapisi, cerrahi, stereotaktik radyocerrahi

Corresponding Author: Mert Saynak ✉ mertsaynak@yahoo.com
Received: 18.09.2025 **Revision:** 09.10.2025 **Accepted:** 29.11.2025

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Whole-brain radiotherapy (WBRT) produced a significant reduction in both tumor diameter and lesion volume, demonstrating measurable structural regression across follow-up evaluations.
- WBRT achieved a 67% partial or complete radiological response at 8–10 weeks in patients with multiple brain metastases; in those with solitary metastasis, the complete response rate was 50%, indicating a more pronounced treatment effect.
- Radiological assessments at 4–6 and 8–10 weeks confirmed progressive lesion shrinkage, while neurological symptoms improved and sensory–motor functions remained stable without treatment-related decline.
- Combined clinical and radiological findings suggest that WBRT provides meaningful disease control and may render selected responders suitable for subsequent local therapies.

INTRODUCTION

Patients with systemic metastatic disease are experiencing longer survival times due to the more widespread use of new systemic therapies, including targeted therapies and immunotherapies. Recent studies indicate that the increased survival associated with these combined treatments also raises the likelihood of developing brain metastases (1, 2). Notably, brain metastases occur in 40-60% of patients diagnosed with non-small cell lung cancer (NSCLC) at any stage of the disease (2, 3).

Whole brain radiotherapy (WBRT) has long been a fundamental treatment for brain metastases, helping to improve neurological symptoms, enhance quality of life, and reduce fatalities due to neurological issues. Historically, WBRT has been the standard approach for managing brain metastases. Over the past 20 years, the role of surgery or stereotactic radiosurgery (SRS) in conjunction with WBRT has been thoroughly evaluated. Research shows that combining surgery and/or SRS with WBRT, particularly in patients with a limited number of brain metastases and controlled disease outside the brain, can enhance overall survival and improve disease control within the brain. As a result, this combination therapy has emerged as the standard treatment. (4, 5). While the efficacy of WBRT in managing brain disease is well-established for this group of patients, studies show conflicting results regarding its impact on overall survival (5-8).

SRS is a radiotherapy technique that can deliver highly effective doses to small intracranial targets. It is a non-invasive treatment method that patients typically tolerate well. However, there are ongoing controversies regarding patient selection for this procedure. SRS has been shown to provide better outcomes in controlling treated lesions compared to WBRT (9).

However, two-thirds of cases with brain metastases have more than one metastasis; for a significant part of these patients, applying ablative treatments at the time of presentation is not reasonable (2). Our study aimed to evaluate the clinical and radiological response after WBRT in patients with multiple brain metastases diagnosed with NSCLC.

MATERIAL and METHODS

Eighteen patients, each diagnosed with NSCLC and brain metastases, were thoughtfully included in this study, which garnered approval from the local ethics committee. A total of 18 patients were initially enrolled in the study; however, 12 patients who completed both neurological and radiological follow-up evaluations at 4–6 and 8–10 weeks after radiotherapy were included in the final analysis. For these 12 patients, neurological examinations and magnetic resonance imaging (MRI) findings were carefully reviewed to provide a thorough assessment. As a result, only those patients with brain metastases who received a follow-up of at least 10 weeks were selected for inclusion in our study, ensuring a robust and meaningful analysis. Figure 1 summarizes the flow of patient inclusion and exclusion during the study.

Treatment

All patients received an initial dose of 16 mg of dexamethasone upon admission and were gradually tapered off the medication following WBRT. Each patient underwent WBRT delivering a total dose of 30 Gy in 10 fractions. They were positioned supine on the treatment table, immobilized using a thermoplastic mask. A computed tomography (CT) simulation was conducted for each patient included in the study. Radiation was administered using 6 megavolt and 15/18 MV energies with four lateral opposing fields, employing a field-in-field technique to ensure accurate dose distribution.

Three-dimensional dose planning was performed using version 4.6 of the XiO CMS model planning system (Computerized Medical Systems, St. Louis, MO, USA) or version 6.5 of the Eclipse planning system. In addition, an extra dose of 9 Gy in 3 fractions was delivered to the tumor region for 4 patients with solitary brain metastases. Chemotherapy was

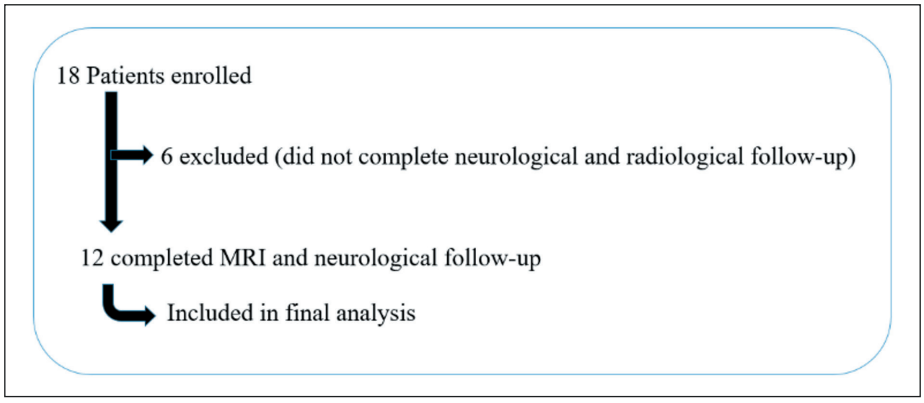


Figure 1: Flow diagram of patient inclusion and analysis

Table 1: Patient demographics and clinical characteristics

Characteristic	n	%
Male	12	100
Median Age (range, minimum-maximum)	61.5 (49-72)	
Median follow-up time (months) (minimum-maximum)	7.0 (2.8-12.9)	
Pathologic diagnosis		
Adenocarcinoma	7	58
Squamous cell carcinoma	2	17
Unknown	3	25
T Stage		
T1	1	8
T2	4	34
T3	1	8
T4	6	50
N Stage		
N0	1	8
N1	2	17
N2	7	67
N3	1	8
Number of Brain Metastases		
1	4	33
2	6	50
≥3	2	17
Largest Metastasis Diameter at Presentation (cm)	0.6-4.3 (median 2.5)	
Largest Metastasis Diameter at 8-10 weeks (cm)	0-3.5 (median 1.6)	

administered to 10 patients, accounting for 83% of the study group. This study focused solely on patients who were not receiving targeted or immunotherapy agents. Treatment was carried out using either 6 MV and 15 MV photon energies from the Elekta Synergy Platform or 6 MV and 18 MV photon energies from the Varian linear accelerator.

The T and N stages of the patients involved in the study are summarized in Table 1. Among the nine patients who underwent thoracic radiotherapy (RT), two received palliative RT, while the remaining seven were treated with curative or near-curative RT using either a conventional or hypofractionated regimen. Additionally, three patients had metastases in other sites.

The patient's performance status was assessed using the Eastern Cooperative Oncology Group Performance Scale (10).

Radiological Evaluation

Radiological evaluation was performed with a 1 Tesla superconductive magnet (Magnetom impact-Expert; Siemens, Erlangen, Germany) MRI device with a gradient power of 20 mT/m.

Lesion dimensions were measured in three planes length, width, and height using T1-weighted MR images. Lesion volumes were calculated using the formula "volume= π/6 x R1 x R2 x R3", R1, width; R2, depth and R3, height, or r1= R1/2, r2= R2/2 and r3= R3/2, volume = 4/3 x π x r1 x r2 x r3 formula to calculate elliptical shaped volumes.

Neurological Evaluation

Neurological examinations of the patients included in the study were conducted upon admission and during the fourth and eighth weeks following WBRT. These examinations were performed by a radiation oncologist and a neurologist. The findings were graded according to the RTOG (Radiation Therapy Oncology Group) Central Nervous System (CNS) Toxicity Criteria. (11).

Statistical Method

The statistical analysis was conducted using the SPSS 20.0 software. Data were analyzed, and survival times were esti-

mated with Kaplan-Meier analysis. Categorical (qualitative) variables were compared using Fisher's exact test. For the comparison of repeated measurements obtained at three time points (pre-radiotherapy, first control, and second control), the non-parametric Friedman test was applied. When statistically significant differences were found, post-hoc pairwise comparisons were conducted using the Wilcoxon signed-rank test with Bonferroni correction to adjust for multiple comparisons. A p-value < 0.05 was considered statistically significant for all analyses.

To enhance methodological transparency, an approximate a priori power analysis was conducted based on the differences between paired measurements obtained before and after WBRT. This analysis focused on the comparison between pre-radiotherapy (Pre-RT) and the first (4–6 weeks after WBRT) and, second (8-10 weeks) to estimate the magnitude of expected change across the three evaluation points. For a two-sided test, the significance level (α) was set at 0.05, and the statistical power (1–β) at 0.80. Assuming a large paired effect size ($d_x \approx 0.8$) as reported in similar radiotherapy response studies, the required sample size was approximately n ≈ 13. Accordingly, the present 12-patient analysis cohort was close to this threshold, while the full cohort of 18 patients exceeded it. Therefore, the study was considered statistically adequate for pilot-level exploratory analyses.

RESULTS

The treatment response of 12 patients diagnosed with NSCLC and brain metastases (ranging from 49 to 72 years) was compared in MRI examinations before WBRT and 8-10 weeks after irradiation.

Among the patients, 7 (58%) were diagnosed with adenocarcinoma, 2 (17%) with squamous cell carcinoma (SCC), and the pathological subtype could not be determined for 3 patients (25%). Before WBRT, a median of 2 brain lesions (ranging from 1 to 4 lesions) was detected in the MRI examinations. Detailed patient characteristics are summarized in Table 1.

Headache was present in 4 patients (33%), seizures in 2 patients (17%), coordination disorder in 2 patients (17%), weakness in the ipsilateral extremities in 2 patients (17%), and personality changes in one patient (8%). One patient (8%) did not have any symptoms at presentation. The patient's symptoms and neurological examination findings are shown in Table 2.

Neurological scores were graded between 1 (normal) and 4 (severe deficit) based on the Radiation Therapy Oncology Group (RTOG) toxicity criteria, Pre-RT: Before radiotherapy (baseline evaluation), None: No predominant lesion. Statistical comparison among the three time points was performed using the Friedman test, followed by Wilcoxon signed-rank tests with Bonferroni correction. No statistically significant difference was observed for motor (p = 0.11) or sensory (p = 0.24) scores across evaluations.

Motor and sensory functions were evaluated in the neurological examination. While motor and sensory deficits were not detected in 10 of 12 patients (83%), sensory score grade 2 was found in 2 patients (17%), and motor score

Table 2: Predominant neurological symptoms and their graded motor and sensory scores before and after radiotherapy (according to RTOG criteria)

Patient	Predominant Symptom	RTOG Neurological Toxicity Score					
		Initial (Pre-RT)		4-6 weeks		8-10 weeks	
		Motor	Sensory	Motor	Sensory	Motor	Sensory
1	Coordination disorder	1	1	1	1	1	1
2	Extremity weakness	1	1	1	1	1	1
3	None	1	1	1	1	1	1
4	Headache	1	1	1	1	1	1
5	Extremity weakness	1	1	1	1	1	1
6	Coordination disorder	1	1	1	1	1	1
7	Headache	1	1	1	1	1	1
8	Personality change	1	1	1	1	1	1
9	Seizure	3	2	2	2	2	2
10	Seizure	1	1	1	1	1	1
11	Headache	1	1	1	1	1	1
12	Headache	3	2	1	1	1	1

Pre-RT: Pre-radiotherapy

grade 3 was found in the same two patients. In all patients who had symptoms in the presentation, there was an improvement with WBRT. One patient with motor and sensory deficits showed a complete response, but the other patient had a motor score of 2 in the control evaluations and was considered a partial response.

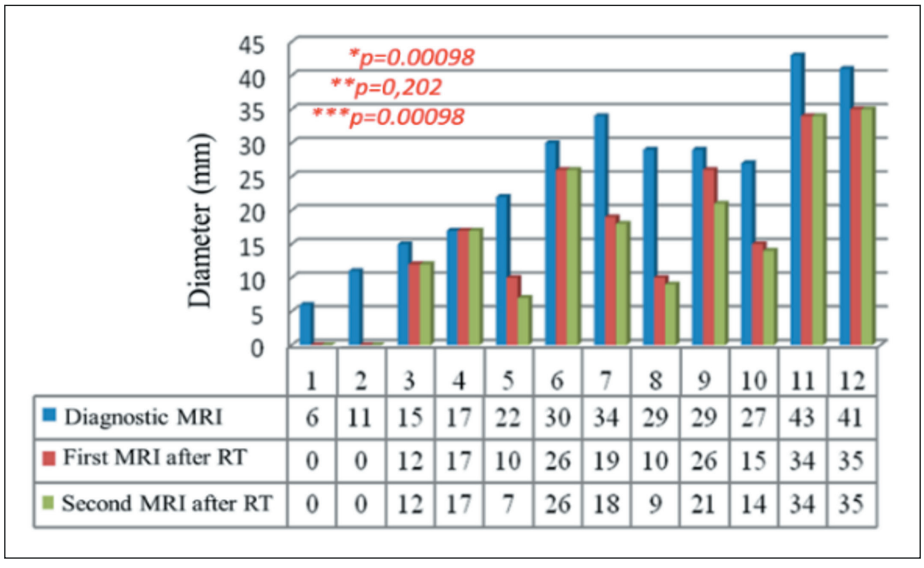
In the evaluation performed 4-6 weeks after the treatment, radiologically complete response was determined in 2 patients (17%), partial response in 6 patients (50%), and stable disease in 4 patients (33%). 4-6 weeks after WBRT. When the 4-6 week and 8-10 week imaging were compared, stable disease was found in 11 patients (92%), and partial response was found in 1 patient (8%).

Initially, the size of the largest metastasis ranged from 0.6 cm to 4.3 cm, with a median of 2.5 cm. After 8 to 10 weeks, the measurements decreased to a range of 0 to 3.5 cm, with a median of 1.6 cm.

Tumor diameters measured at pre-radiotherapy, first follow-up, and second follow-up were compared in millimeters (mm). As the data did not meet the assumptions required for parametric testing, the non-parametric Friedman test was used for analysis, and the Wilcoxon signed-rank test was applied for post-hoc pairwise comparisons when significant differences were observed. According to the Friedman test, there was a statistically significant difference among the three time points ($\chi^2(2) = 15.05$, $p = 0.00054$). To identify which time points contributed to this difference, post-hoc Wilcoxon signed-rank tests were conducted, and p-values were adjusted using the Bonferroni correction for multiple comparisons (adjusted significance level $\alpha = 0.05/3 = 0.0167$). After correction, significant reductions were observed between pre-radiotherapy and the first follow-up ($p = 0.00098 < 0.0167$) and between pre radiotherapy and the second follow-up ($p = 0.00098 < 0.0167$). However, there was no significant difference between the first and second follow-up measurements ($p = 0.202 > 0.0167$). These findings indicate that a significant reduction in tumor size occurred in the early post-radiotherapy period and that this reduction was maintained until the second follow-up evaluation (Figure 2).

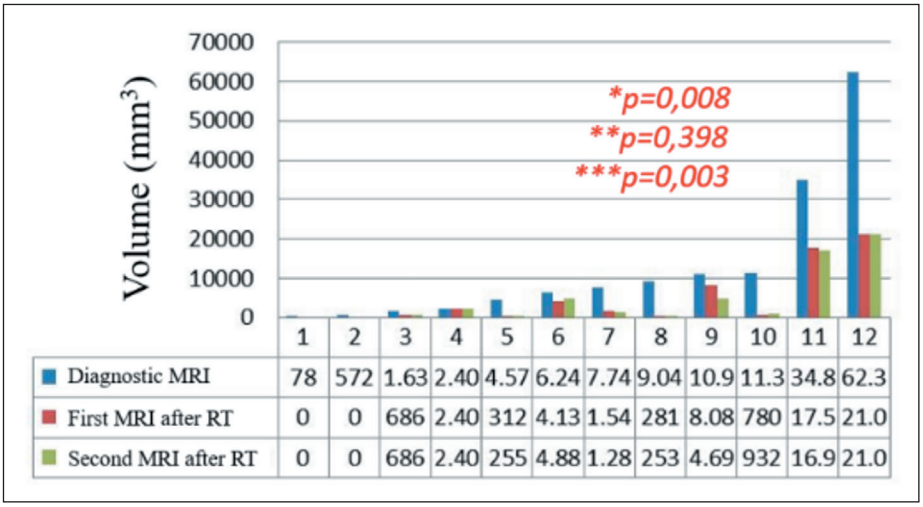
Similar results were confirmed by volumetric tumor assessment. When comparing pre-radiotherapy and first follow-up MRI scans, a significant difference in tumor volume was found ($p = 0.008$); a significant volumetric difference was also observed between the pre-radiotherapy and second follow-up MRIs ($p = 0.003$). However, no significant difference in tumor volume was detected between the first and second follow-up MRIs ($p = 0.398$) (Figure 3).

Sensory and motor (neurological) grade scores were also evaluated at pre-radiotherapy, first follow-up, and second follow-up, using a grading scale from 1 to 4, and their temporal changes were statistically analyzed. As the variables were ordinal in nature, the non-parametric Friedman test was used, followed by Wilcoxon signed-rank tests for post-hoc pairwise comparisons when relevant. No statistically significant differences were detected among the three time points in either the sensory or motor grade scores (sensory: $\chi^2(2) = 2.00$, $p > 0.05$; motor: $\chi^2(2) \approx 3.00$, $p > 0.05$). Post-hoc Wilcoxon tests also revealed no significant pairwise differences. Most patients maintained stable sensory and motor scores (predominantly Grade 1), with only mini-



	1	2	3	4	5	6	7	8	9	10	11	12
Diagnostic MRI	6	11	15	17	22	30	34	29	29	27	43	41
First MRI after RT	0	0	12	17	10	26	19	10	26	15	34	35
Second MRI after RT	0	0	12	17	7	26	18	9	21	14	34	35

Figure 2: Tumor diameters measured at baseline and at subsequent follow-up assessments are presented .
MRI: Magnetic Resonance Imaging, **RT:** Radiotherapy, *p: Comparison of pre-radiotherapy and first follow-up MRI; **p: Comparison of first and second follow-up MRI, ***p: Comparison of pre-radiotherapy and second follow-up MRI.



	1	2	3	4	5	6	7	8	9	10	11	12
Diagnostic MRI	78	572	1.63	2.40	4.57	6.24	7.74	9.04	10.9	11.3	34.8	62.3
First MRI after RT	0	0	686	2.40	312	4.13	1.54	281	8.08	780	17.5	21.0
Second MRI after RT	0	0	686	2.40	255	4.88	1.28	253	4.69	932	16.9	21.0

Figure 3: Tumor volumes measured at baseline and at subsequent follow-up assessments are presented
MRI: Magnetic Resonance Imaging; **RT:** Radiotherapy; *p: Comparison of pre-radiotherapy and first follow-up MRI; **p: Comparison of first and second follow-up MRI; ***p: Comparison of pre-radiotherapy and second follow-up MRI.

mal improvement observed in a few cases. From a clinical perspective, considering that patients had no significant neurological deficits before radiotherapy and their baseline scores were low, the absence of significant change during follow-up suggests that radiotherapy did not cause additional neurological or sensory impairment. These findings support the preservation of neurological stability after treatment.

Our study obtained a partial or complete response in 8 (67%) of 12 patients at 8-10 weeks of follow-up. Complete response was detected radiologically in 2 (50%) of 4 patients with solitary lesions. The median response rate of responders was 86% (58-100%). The pathological diagnosis of two patients with a complete response was adenocarcinoma, and the initial diameter of the metastasis was 0.6 cm and 1.1 cm. The pathological diagnosis of a patient with solitary brain metastases in whom a complete response could not be obtained was SCC, whose initial diameter was 4.1 cm. In this patient, the diameter of metastasis measured in the 8 weeks following WBRT was 3.5 cm in the brain MRI examination. SRS was applied to the other patient with solitary 1.8 cm metastasis, and stable disease was detected in the examinations after WBRT. One of the lesions disappeared in 2 (25%) of 8 patients with multiple metastases, and partial response was obtained in the remaining lesions. The dimensions of the disappearing lesion were 0.8 cm in diameter in both patients. Partial response was determined in 5 (63%) of the other eight patients who presented with multiple brain metastases, and stable disease was defined in 3 (37%) patients.

The median follow-up period of the patients from the presentation was 12.3 months (5.4-35.4 months). The median overall survival was 12.9 months (95% CI, 6.2-36.4 months). When evaluated specifically for cranial disease, the median follow-up time following WBRT was seven months (3.2-14.4 months), and the median survival time was 7.7 months (95% CI, 3.7-15.1 months). Eleven patients had brain metastases at presentation.

Six of 12 patients (50%) died during follow-up. Three patients (25%) died due to the progression of primary disease, two patients (17%) due to the progression of the cranial disease, and one patient (13%) due to non-disease reasons. Cranial progression was observed 6.2 months after WBRT in the patient who died due to cranial disease. Both local and systemic progression was detected 2.5 months before cranial progression in this patient.

DISCUSSION

Lung cancer is a leading cause of brain metastasis. Approximately 10–20% of patients with NSCLC present with brain metastasis at diagnosis, and about 40% will develop it during the course of the disease (2, 3). The incidence of brain metastases is rising, largely due to improvements in imaging techniques and advancements in the treatment of extracranial disease (12). Given the prevalence of lung cancer and the significant risk of brain metastases among these patients, it is crucial to select the most effective treatment method at the appropriate time. Treatment options for brain metastases include WBRT, surgical excision, SRS, supportive care, and various combinations of these approaches. Generally, WBRT is the standard treatment for patients with multiple brain metastases. Additionally, corticosteroids and anticonvulsant medications are commonly used to manage symptoms (13).

Dexamethasone is a steroid drug that has a palliative effect by reducing edema caused by brain metastases. It has been in use since the 1960s and has a high rate of palliative effectiveness (14). Although there is no established consensus on the optimal dosage, it is generally administered at a range of 4 to 16 mg per day (15). In our study, all patients received a 16 mg intravenous loading dose of dexamethasone, followed by 8 mg given either orally, intramuscularly, or intravenously every 12 hours during radiotherapy. An initial neurologic evaluation was conducted after at least 24 hours of dexamethasone treatment to exclude any effects of the anti-edema therapy on the study results.

Brain metastasis is accepted as a precursor of short survival time in cancer patients (16), and these patients are treated with palliation-oriented treatment methods. WBRT has historically been the only treatment for brain metastases. While this treatment alleviates the patient's symptoms by shrinking the radiologically visible lesions or preventing the growth of these lesions, it also contributes to disease control in the brain in the long term by affecting the microscopic disease foci that cannot be visualized in the brain tissue. However, there hasn't been enough research done to determine the likelihood of an objective response to this treatment. In recent years, information based on comparative studies shows that patients with limited metastases may benefit from more aggressive treatment methods. Surgery and/or SRS are essential in treating patients with 1-4 brain metastases, especially with favorable prognostic factors (5, 6).

Although adding WBRT to local ablative therapy such as surgery or SRS increases the probability of brain control in oligometastatic brain disease, it does not appear to contribute to survival. On the other hand, it has been reported that WBRT may negatively affect neurocognitive brain functions. Aoyama et al. applied SRS and WBRT or SRS alone to 132 patients with 1-4 brain metastases of 3 cm or less. The recurrence rate was found to be significantly lower in the WBRT group (46.8% vs. 76.4%, $p<0.001$); but, no difference was found in survival time (median 7.5 months vs. 8 months) and preservation of neurological functions (7). Chang et al. published a prospective randomized study of 58 patients with treatment arms that added WBRT to SRS versus SRS alone. Researchers reported that patients who received WBRT had better disease control in the brain, while overall survival was worse in these patients (8). The European Organization for Research and Treatment of Cancer (EORTC) conducted a study on 359 patients with 1-3 brain metastases to evaluate the same subject. The patients were randomly assigned to receive local therapy (surgery or SRS) and/or WBRT. There was no difference in survival, but the probability of brain recurrence with WBRT decreased from 59% to 27% in the local ablative treatment group and from 31% to 19% in the SRS group (17). Patchel et al. randomized 95 patients to surgery with or without WBRT arms; with 50.4 Gy/28fr WBRT after resection, the recurrence rate occurring in any region of the brain (18% vs. 70%), the risk of progression in the initial metastasis region (10% vs. 46%) and the risk of death due to metastatic disease in the brain decreased (14% vs. 44%). However, there was no significant difference in overall survival time (median 48 vs. 43 weeks). As a result, they reported that WBRT was associated with higher local control and less neurological death but did not contribute to overall survival (18).

In the analyses made considering the technological level reached, it has been revealed that the volume of the lesions is a more important criterion than the number of lesions for SRS applications. In a prospective observational study of 1174 patients with 76% NSCLC with brain metastases treated with SRS for up to 10 lesions, survival rates were similar in patients with 2-4 or 5-10 metastases. No additional toxicity due to SRS was identified in patients with >4 metastases (19).

Systemic treatments cannot reach the concentration that can be effective in the tumor tissue in the brain by crossing the blood-brain barrier. The passage of molecules larger than 180dalton is blocked by the blood-brain barrier. The molecular weight of most chemotherapy (CT) drugs (such as vincristine, vinblastine, paclitaxel, and etoposide) exceeds 400 daltons. In addition, many CT drugs (such as chlorambucil, etoposide, melphalan, and paclitaxel) are more than 90% bound to plasma proteins. In contrast, crossing drugs across the blood-brain barrier is related to the concentration of free drugs in the blood (20). Although some hydrophilic drugs, such as methotrexate, have relatively better diffusion into tumor tissue, animal studies show that the ratio of brain concentration to serum concentration is about 0.2 even for this drug (21, 22). CT drugs such as cisplatin, carboplatin, tegafur, and temozolomide have also been tried in treating patients with brain metastases diagnosed with NSCLC. Although significant response rates have been obtained in some studies, these drugs did not contribute to survival time (23-25).

Patients with multiple brain metastases or extensive extracranial disease are treated with WBRT. In this kind of treatment, a dose of 30 Gy in 10 fractions or 20 Gy in 5 fractions is usually administered to the whole brain. The use of higher fraction doses increases toxicity (26). This treatment's median expected survival time is 4-5 months, and the 1-year survival is 10-15% (27). In the QUARTZ randomized trial, WBRT and best supportive care were compared in patients with NSCLC with brain metastases unsuitable for resection or stereotactic radiotherapy. No benefit was found in terms of overall survival or quality of life (28).

Many studies have examined the side effects of WBRT, SRS, and both. For example, according to De Angelis et al, in a trial conducted, it was reported that WBRT causes dementia. But this effect is observed only when doses higher than 300cGy/fraction are used (29). In the NRG Oncology CC001 phase III randomized study, it was determined that reducing the hippocampus dose at the same time while limiting the maximum doses in the irradiated volume protects neurocognitive functions (in tests performed 6 months after RT) in patients receiving 30 Gy RT in 10 fractions to the whole brain region due to metastatic disease (30). The RTOG conducted a placebo-controlled, double-blind, randomized study to evaluate the potential protective effect of memantine, a drug used to treat Alzheimer's disease, on neurocognitive function in patients receiving WBRT (31).

Although stereotactic radiotherapy substantially improves tumor control, the management of treatment-related necrosis has become increasingly critical. Bevacizumab is widely employed off-label in the management of SRS-associated brain necrosis, where it has demonstrated improvements in neurological symptoms, reductions in corticosteroid dependence, and marked decreases in both contrast-enhancing lesion volume and perilesional edema on MRI. Additionally, exploratory clinical studies have examined the potential therapeutic role of Boswellia serrata, whereas Annona muricataleaf extract has shown promising neuroprotective effects in preclinical animal models (32-34).

Typically, the literature explores the responsiveness of brain metastases to WBRT to determine the potential for symptom relief. However, evaluating lesions' quantitative response has not been a primary goal (6). Since the survival expectancy of patients with brain metastases is relatively low, less attention has been paid to the radiological response. While there have been varied evaluation criteria used in the literature, the issue of symptomatic response to WBRT in patients with brain metastases has been more extensively researched than radiological response. Various series have reported symptom responses ranging from 60% to 90% (35-37). All patients with symptoms in the presentation showed improvement with WBRT in our study. Some of the existing limited studies were conducted earlier and relied on CT scans for evaluation (38). We used MRI, which is more sensitive and specific than CT, to assess brain metastases' number and size. Knowing when to evaluate the response to radiotherapy is key to increasing the number of brain metastases patients who can benefit from curative treatments. Unfortunately, a study to determine this period has yet to be published.

The recent discovery of targetable driver mutations, such as epidermal growth factor receptor (EGFR), anaplastic lymphoma kinase (ALK), and C-ras oncogene 1 (ROSI) genes in NSCLC has enabled the creation of small-molecule tyrosine kinase inhibitor drugs (TKIs) targeting these driver mutations. These targeted drugs are more effective against brain metastases than conventional chemotherapy drugs (39). Even so, in the retrospective study of Manguson et al., the worst results in patients with EGFR-TKI-naive NSCLC were seen in patients who received TKI alone. The best overall survival probability was achieved with upfront SRS followed by EGFR-TKI (40). New tyrosine kinase inhibitors, with few exceptions like crizotinib, can effectively treat brain metastases by crossing the blood-brain barrier more easily(41). Third-generation EGFR inhibitors such as osimertinib, second-generation ALK inhibitors such as alectinib (42) and brigatinib (43), and third-generation ALK inhibitors such as lorlatinib (44) are highly effective against brain metastases and provide cerebral control until drug resistance develops.

Immunotherapy drugs have recently been effective against brain metastases in lung cancers and paved the way for a new generation of systemic treatments. Immunotherapeutic drugs cross the blood-brain barrier better than systemic chemotherapy drugs. However, treating symptomatic brain metastases with immunotherapy alone is not appropriate. There is a systemic collaboration between SRS and immunotherapies called the abscopal effect, but the clinical significance of which still needs to be fully clarified (45). However, it has been reported that the risk of toxicity increases 2.4 times in concurrent applications compared with sequential chemoradiotherapy (46). On the other hand, although the expectation of lymphopenia (which has a role in the efficacy of immunotherapy) with WBRT is relatively higher, radiation necrosis is a side effect more specific to SRS. Therefore, only some reports focus on integrating WBRT and immunotherapies. For example, Gerber et al. performed a retrospective analysis of 13 patients with brain metastases treated with WBRT within 30 days of ipilimumab administration. Their findings revealed that the combination of ipilimumab and WBRT provided encouraged efficacy (47). None of the patients in our study received TKI or immunotherapy, which are now considered more effective than traditional chemotherapy on the CNS. Therefore, the systemic treatment received by the patients in this study did not have the power to affect the on-objective findings.

Our study obtained a response in 8 (67%) of 12 patients in 8-10 weeks of follow-up. The complete radiological response was obtained in 2 (50%) of 4 patients with solitary lesions. The pathological diagnosis of these two patients was adenocarcinoma, and the initial diameter of the metastasis was 1.1 cm and 0.6 cm. The pathological diagnosis of a patient with solitary brain metastases in whom a complete response could not be obtained was squamous cell carcinoma, whose initial diameter was 4.1 cm. In this patient, the diameter of metastasis measured in the brain MRI performed at 8 weeks following WBRT was 3.5 cm. The lesion diameter of the other patient with a solitary lesion re-

mained stable. Both patients were referred to SRS following the second imaging. One of the lesions disappeared in 2 (25%) of 8 patients with multiple metastases. The diameters of the disappeared lesions were 0.8 cm in both patients. In all patients who presented with symptoms, WBRT provided improvement.

In the neoadjuvant treatment of rectal cancer, it has been shown that the probability of a response increases when the waiting time is prolonged following RT (48). Similarly, a waiting period approach is applied in treating superior sulcus tumors (49). Our study obtained a partial or complete response in 58% of the patients due to an MRI examination performed 4-6 weeks after palliative RT for brain metastasis. However, further prolonging the waiting time increased the response rate in only one patient. 4-6 and 8-10. When the lesion volumes in the weeks were compared, a significant rate could not be reached (p=0.398). A significant response was observed when the diagnostic MRI was compared with the first MRI after treatment. (p=0.008). However, this response rate was not associated with survival or local progression.

In patients with multiple brain metastases, SRS can be reserved for applying to a few lesions that progress after WBRT. Most of the series that included the patients treated with SRS because of progressed/recurrent brain disease consisted of patients with recurred after WBRT or were far from the original SRS site (50). Thus, the risk of necrosis that may develop due to SRS application after WBRT does not increase. There haven't been any studies that compare SRS with residual disease in patients who receive WBRT versus those who get SRS if the disease progresses.

Although our study has produced valuable insights, it's important to acknowledge its limitations. One such limitation is the relatively small sample size of the study, which may impact the generalizability of our findings. Another limitation is the lack of association of WBRT with new agents in our study. On the other hand, this may provide insight for patients with brain metastases who are not treated with new agents. We believe that the findings from this study will constitute a small step toward future research in this field.

Conclusion

Lung cancer is the most important cause of brain metastasis. WBRT has historically been the only treatment for brain metastases. While this treatment alleviates the patient's symptoms by shrinking the radiologically visible lesions or preventing the growth of these lesions, it also contributes to disease control in the brain in the long term by affecting the microscopic disease foci that cannot be visualized in the brain tissue. Recently, it has been understood that long survival times can be achieved by adding SRS and/or surgery to the treatment in the appropriate patient group. The response rates we obtained with WBRT in our study are promising in making patients with a controlled primary disease and metastasis(s) in the brain not suitable for surgery or SRS at the beginning become ideal for these treatments, which have curative potential with WBRT. However, how permanent the response obtained with WBRT is also important. Therefore, studies with a larger number of patients and longer follow-up periods are needed to support the findings of our research.

Author Contributions

Study conception and design: **Mert Saynak, Funda Çukurçayır, Kamuran İbiş, Alaatin Özen, Zafer Koçak, Görkem Tükkan,** data collection: **Funda Çukurçayır, Canan Çelebi, Alaatin Özen, Görkem Türkkan, Kamuran İbiş,** analysis and interpretation of results: **Mert Saynak, Funda Çukurçayır, Canan Çelebi, Alaattin Özen, Görkem Türkkan, Kamuran İbiş, Mustafa Akın, Zafer Koçak,** draft manuscript preparation: **Mert Saynak, Funda Çukurçayır, Mustafa Akın, Alaatin Özen, Canan Çelebi.** The author(s) reviewed the results and approved the final version of the article.

Conflicts of Interest

There is no conflict of interest.

Ethical Approval

Eighteen patients with NSCLC diagnosed with brain metastases following the Trakya University Medical Faculty Clinical Researches Ethics each diagnosed with non-small cell lung cancer (NSCLC) and brain metastases, were thoughtfully included in this study, which garnered approval from the Clinical Research Ethics Committee of Trakya University Medical Faculty (Date: 18.08.2012, Decision No: 18/08).

REFERENCES

1. Smedby KE, Brandt L, Backlund ML, Blomqvist P. Brain metastases admissions in Sweden between 1987 and 2006. Br J Cancer. 2009;101(11):1919-1924. https://doi.org/10.1038/sj.bjc.6605373
2. Achrol AS, Rennert RC, Anders C, Soffietti R, Ahluwalia MS, Nayak L, et al. Brain metastases. Nat Rev Dis Primers. 2019;5(1):5. https://doi.org/10.1038/s41572-018-0055-y
3. Hubbs JL, Boyd JA, Hollis D, Chino JP, Saynak M, Kelsey CR. Factors associated with the development of brain metastases: analysis of 975 patients with early stage nonsmall cell lung cancer. Cancer. 2010;116(21):5038-5046. https://doi.org/10.1002/cncr.25254
4. Noordijk EM, Vecht CJ, Haaxma-Reiche H, Padberg GW, Voormolen JH, Hoekstra FH, et al. The choice of treatment of single brain metastasis should be based on extracranial tumor activity and age. Int J Radiat Oncol Biol Phys. 1994;29(4):711-717. https://doi.org/10.1016/0360-3016(94)90558-4
5. Patchell RA, Tibbs PA, Walsh JW, Dempsey RJ, Maruyama Y, Kryscio RJ, et al. A randomized trial of surgery in the treatment of single metastases to the brain. N Engl J Med. 1990;322(8):494-500. https://doi.org/10.1056/NEJM199002223220802

6. Andrews DW, Scott CB, Sperduto PW, Flanders AE, Gaspar LE, Schell MC, et al. Whole brain radiation therapy with or without stereotactic radiosurgery boost for patients with one to three brain metastases: phase III results of the RTOG 9508 randomised trial. Lancet. 2004;363(9422):1665-1672. https://doi.org/10.1016/S0140-6736(04)16250-8
7. Aoyama H, Shirato H, Tago M, Nakagawa K, Toyoda T, Hatano K, et al. Stereotactic radiosurgery plus whole-brain radiation therapy vs stereotactic radiosurgery alone for treatment of brain metastases: a randomized controlled trial. Jama. 2006;295(21):2483-2491. https://doi.org/10.1001/jama.295.21.2483
8. Chang EL, Wefel JS, Hess KR, Allen PK, Lang FF, Kornguth DG, et al. Neurocognition in patients with brain metastases treated with radiosurgery or radiosurgery plus whole-brain irradiation: a randomised controlled trial. Lancet Oncol. 2009;10(11):1037-1044. https://doi.org/10.1016/S1470-2045(09)70263-3
9. Tsao MN, Sneed PK, McDermott MW, Larson DA. Radiosurgery and radiotherapy for non-small-cell lung cancer metastatic to brain. Clin Lung Cancer. 2001;2(3):197-203. https://doi.org/10.1016/j.cllc.2013.12.008
10. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. 1982;5(6):649-655. https://doi.org/10.1097/00000421-198212000-00014
11. Cox JD, Stetz J, Pajak TF. Toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European Organization for Research and Treatment of Cancer (EORTC). Int J Radiat Oncol Biol Phys. 1995;31(5):1341-1346. https://doi.org/10.1016/0360-3016(95)00060-C
12. Hazard LJ, Jensen RL, Shrieve DC. Role of stereotactic radiosurgery in the treatment of brain metastases. Am J Clin Oncol. 2005;28(4):403-410. https://doi.org/10.1097/01.coc.0000158438.79665.bb
13. Tsao MN, Xu W, Wong RK, Lloyd N, Laperriere N, Sahgal A, et al. Whole brain radiotherapy for the treatment of newly diagnosed multiple brain metastases. Cochrane Database Syst Rev. 2018;1(1):CD003869. https://doi.org/10.1002/14651858.CD003869.pub4
14. Galicich JH, French LA, Melby JC. Use of dexamethasone in treatment of cerebral edema associated with brain tumors. J Lancet. 1961;81:46-53.
15. Bezjak A, Adam J, Barton R, Panzarella T, Laperriere N, Wong CS, et al. Symptom response after palliative radiotherapy for patients with brain metastases. Eur J Cancer. 2002;38(4):487-496. https://doi.org/10.1016/S0959-8049(01)00150-2
16. Balasubramanian SK, Sharma M, Venur VA, Schmitt P, Kotecha R, Chao ST, et al. Impact of EGFR mutation and ALK rearrangement on the outcomes of non-small cell lung cancer patients with brain metastasis. Neuro Oncol. 2020;22(2):267-277. https://doi.org/10.1093/neuonc/noz155
17. Kocher M, Soffietti R, Abacioglu U, Villa S, Fauchon F, Baumert BG, et al. Adjuvant whole-brain radiotherapy versus observation after radiosurgery or surgical resection of one to three cerebral metastases: results of the EORTC 22952-26001 study. J Clin Oncol. 2011;29(2):134-141. https://doi.org/10.1200/JCO.2010.30.1655
18. Patchell RA, Tibbs PA, Regine WF, Dempsey RJ, Mohiuddin M, Kryscio RJ, et al. Postoperative radiotherapy in the treatment of single metastases to the brain: a randomized trial. JAMA. 1998;280(17):1485-1489. https://doi.org/10.1001/jama.280.17.1485
19. Yamamoto M, Kawabe T, Sato Y, Higuchi Y, Nariai T, Watanabe S, et al. Stereotactic radiosurgery for patients with multiple brain metastases: a case-matched study comparing treatment results for patients with 2-9 versus 10 or more tumors. J Neurosurg. 2014;121 Suppl:16-25. https://doi.org/10.3171/2014.8.GKS141421
20. Vecht CJ, Hovestadt A, Verbiest HB, van Vliet JJ, van Putten WL. Dose-effect relationship of dexamethasone on Karnofsky performance in metastatic brain tumors: a randomized study of doses of 4, 8, and 16 mg per day. Neurology. 1994;44(4):675-680. https://doi.org/10.1212/WNL.44.4.675
21. Muldoon LL, Soussain C, Jahnke K, Johanson C, Siegal T, Smith QR, et al. Chemotherapy delivery issues in central nervous system malignancy: a reality check. J Clin Oncol. 2007;25(16):2295-2305. https://doi.org/10.1200/JCO.2006.09.9861
22. Dukic SF, Heurtaux T, Kaltenbach ML, Hoizey G, Lallemand A, Vistelle R. Influence of schedule of administration on methotrexate penetration in brain tumours. Eur J Cancer. 2000;36(12):1578-1584. https://doi.org/10.1016/S0959-8049(00)00142-8
23. Verger E, Gil M, Yaya R, Vinolas N, Villa S, Pujol T, et al. Temozolomide and concomitant whole brain radiotherapy in patients with brain metastases: a phase II randomized trial. Int J Radiat Oncol Biol Phys. 2005;61(1):185-191. https://doi.org/10.1016/j.ijrobp.2004.04.061
24. Guerrieri M, Wong K, Ryan G, Millward M, Quong G, Ball DL. A randomised phase III study of palliative radiation with concomitant carboplatin for brain metastases from non-small cell carcinoma of the lung. Lung Cancer. 2004;46(1):107-111. https://doi.org/10.1016/j.lungcan.2004.02.019
25. Ushio Y, Arita N, Hayakawa T, Mogami H, Hasegawa H, Bitoh S, et al. Chemotherapy of brain metastases from lung carcinoma: a controlled randomized study. Neurosurgery. 1991;28(2):201-205. https://doi.org/10.1227/00006123-199102000-00005
26. Rodrigus P, de Brouwer P, Raaymakers E. Brain metastases and non-small cell lung cancer. Prognostic factors and correlation with survival after irradiation. Lung Cancer. 2001;32(2):129-136. https://doi.org/10.1016/S0169-5002(00)00227-0
27. Sperduto PW, Berkey B, Gaspar LE, Mehta M, Curran W. A new prognostic index and comparison to three other indices for patients with brain metastases: an analysis of 1,960 patients in the RTOG database. Int J Radiat Oncol Biol Phys. 2008;70(2):510-514. https://doi.org/10.1016/j.ijrobp.2007.06.074
28. Mulvenna P, Nankivell M, Barton R, Faivre-Finn C, Wilson P, McColl E, et al. Dexamethasone and supportive care with or without whole brain radiotherapy in treating patients with non-small cell lung cancer with brain metastases unsuitable for resection or stereotactic radiotherapy (QUARTZ): results from a phase 3, non-inferiority, randomised trial. Lancet. 2016;388(10055):2004-2014. https://doi.org/10.1016/S0140-6736(16)30825-X
29. De Angelis LM, Delattre JY, Posner JB. Radiation-Induced Dementia in Patients Cured of Brain Metastases. Neurology. 1989;39(6):789-796. https://doi.org/10.1212/WNL.39.6.789
30. Brown PD, Gondi V, Pugh S, Tome WA, Wefel JS, Armstrong TS, et al. Hippocampal Avoidance During Whole-Brain Radiotherapy Plus Memantine for Patients With Brain Metastases: Phase III Trial NRG Oncology CC001. J Clin Oncol. 2020;38(10):1019-1029. https://doi.org/10.1200/JCO.19.02767

31. Brown PD, Pugh S, Laack NN, Wefel JS, Khuntia D, Meyers C, et al. Memantine for the prevention of cognitive dysfunction in patients receiving whole-brain radiotherapy: a randomized, double-blind, placebo-controlled trial. Neuro Oncol. 2013;15(10):1429-1437. https://doi.org/10.1093/neuonc/not114
32. Bernhardt D, König L, Grosu AL, Rieken S, Krieg SM, Wick W, et al. DEGRO practical guideline for central nervous system radiation necrosis part 2: treatment. Strahlenther Onkol. 2022;198(11):971-980. https://doi.org/10.1007/s00066-022-01973-8
33. Dejonckheere CS, Scafa D, Käsmann L, Zeyen T, Potthoff A-L, Schäfer N, et al. Boswellia serrata for the Management of Radiation-Induced Cerebral Edema and Necrosis: A Systematic Meta-Narrative Review of Clinical Evidence. Advances in Radiation Oncology. 2025;10(4):101732. https://doi.org/10.1016/j.adro.2025.101732
34. Elmas Ö, Keser Şahin HH, Keskin E, Apsar Başaran Ö. Radioprotective Effects of Annona Muricata Leaf Extract in the Spinal Cord. Medical Journal of Western Black Sea. 2024;8(2):168-176. https://doi.org/10.29058/mjwbs.1359185
35. Li J, Bentzen SM, Li J, Renschler M, Mehta MP. Relationship between neurocognitive function and quality of life after whole-brain radiotherapy in patients with brain metastasis. Int J Radiat Oncol Biol Phys. 2008;71(1):64-70. https://doi.org/10.1016/j.ijrobp.2007.09.059
36. Gaspar LE, Mehta MP, Patchell RA, Burri SH, Robinson PD, Morris RE, et al. The role of whole brain radiation therapy in the management of newly diagnosed brain metastases: a systematic review and evidence-based clinical practice guideline. J Neurooncol. 2010;96(1):17-32. https://doi.org/10.1007/s11060-009-0060-9
37. DeAngelis LM, Delattre JY, Posner JB. Radiation-induced dementia in patients cured of brain metastases. Neurology. 1989;39(6):789-796. https://doi.org/10.1212/WNL.39.6.789
38. van der Steen-Banasik E, Hermans J, Tjho-Heslinga R, Caspers R, Leer JW. The objective response of brain metastases on radiotherapy. A prospective study using computer tomography. Acta Oncol. 1992;31(7):777-780. https://doi.org/10.3109/02841869209083870
39. Li G, Dai WR, Shao FC. Effect of ALK-inhibitors in the treatment of non-small cell lung cancer: a systematic review and meta-analysis. Eur Rev Med Pharmacol Sci. 2017;21(15):3496-3503.
40. Magnuson WJ, Lester-Coll NH, Wu AJ, Yang TJ, Lockney NA, Gerber NK, et al. Management of Brain Metastases in Tyrosine Kinase Inhibitor-Naïve Epidermal Growth Factor Receptor-Mutant Non-Small-Cell Lung Cancer: A Retrospective Multi-Institutional Analysis. J Clin Oncol. 2017;35(10):1070-1077. https://doi.org/10.1200/JCO.2016.69.7144
41. Metro G, Lunardi G, Floridi P, Pascali JP, Marcomigni L, Chiari R, et al. CSF Concentration of Crizotinib in Two ALK-Positive Non-Small-Cell Lung Cancer Patients with CNS Metastases Deriving Clinical Benefit from Treatment. J Thorac Oncol. 2015;10(5):e26-e27. https://doi.org/10.1097/JTO.0000000000000468
42. Tomasini P, Egea J, Souquet-Bressand M, Greillier L, Barlesi F. Alectinib in the treatment of ALK-positive metastatic non-small cell lung cancer: clinical trial evidence and experience with a focus on brain metastases. Ther Adv Respir Dis. 2019;13:1753466619831906. https://doi.org/10.1177/1753466619831906
43. Camidge DR, Kim HR, Ahn MJ, Yang JC, Han JY, Lee JS, et al. Brigatinib versus Crizotinib in ALK-Positive Non-Small-Cell Lung Cancer. N Engl J Med. 2018;379(21):2027-2039. https://doi.org/10.1056/NEJMoa1810171
44. Shaw AT, Bauer TM, de Marinis F, Felip E, Goto Y, Liu G, et al. First-Line Lorlatinib or Crizotinib in Advanced ALK-Positive Lung Cancer. N Engl J Med. 2020;383(21):2018-2029. https://doi.org/10.1056/NEJMoa2027187
45. Fionda B, Massaccesi M, Tagliaferri L, Dinapoli N, Iezzi R, Boldrini L. Abscopal effect and interventional oncology: state of art and future perspectives. Eur Rev Med Pharmacol Sci. 2020;24(2):773-776.
46. Lehrer EJ, Peterson J, Brown PD, Sheehan JP, Quiñones-Hinojosa A, Zaorsky NG, et al. Treatment of brain metastases with stereotactic radiosurgery and immune checkpoint inhibitors: An international meta-analysis of individual patient data. Radiother Oncol. 2019;130:104-112. https://doi.org/10.1016/j.radonc.2018.08.025
47. Gerber NK, Young RJ, Barker CA, Wolchok JD, Chan TA, Yamada Y, et al. Ipilimumab and whole brain radiation therapy for melanoma brain metastases. J Neurooncol. 2015;121(1):159-165. https://doi.org/10.1007/s11060-014-1617-9
48. Francois Y, Nemoz CJ, Baulieux J, Vignal J, Grandjean JP, Partensky C, et al. Influence of the interval between preoperative radiation therapy and surgery on downstaging and on the rate of sphincter-sparing surgery for rectal cancer: the Lyon R90-01 randomized trial. J Clin Oncol. 1999;17(8):2396. https://doi.org/10.1200/JCO.1999.17.8.2396
49. Rusch VW, Giroux DJ, Kraut MJ, Crowley J, Hazuka M, Winton T, et al. Induction chemoradiation and surgical resection for superior sulcus non-small-cell lung carcinomas: long-term results of Southwest Oncology Group Trial 9416 (Intergroup Trial 0160). J Clin Oncol. 2007;25(3):313-318. https://doi.org/10.1200/JCO.2006.08.2826
50. Maranzano E, Terenzi S, Anselmo P, Casale M, Arcidiacono F, Loreti F, et al. A prospective phase II trial on reirradiation of brain metastases with radiosurgery. Clin Transl Radiat Oncol. 2019;17:1-6. https://doi.org/10.1016/j.ctro.2019.04.003

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):60-64
DOI: 10.29058/mjwbs.1704810

Evaluation of viral respiratory pathogens detected in a tertiary university hospital between 2021 and 2025

Üçüncü basamak bir üniversite hastanesinde 2021-2025 tarihleri arasında saptanan viral solunum yolu etkenlerinin değerlendirilmesi

Canberk Çınar, Sümeyye Özkaya, İlknur Bıyık, Yeliz Tanrıverdi Çaycı, Demet Gür Vural, Kemal Bilgin, Asuman Birinci

Department of Medical Microbiology, Ondokuz Mayıs University Faculty of Medicine, Samsun, Türkiye

Cite this article as: *Çınar C et al. Evaluation of viral respiratory pathogens detected in a tertiary university hospital between 2021 and 2025. Med J West Black Sea. 2026;10(1):60-64.*

ABSTRACT

Aim: Infections caused by respiratory tract pathogens cause high morbidity and mortality and pose a significant threat to public health. In this study, we aimed to determine the prevalence and seasonal distribution of common respiratory viruses in patients admitted to our university hospital between 2021 and 2025.

Material and Methods: Between January 1, 2021 and January 1, 2025, 3131 nasal swab samples from patients with a preliminary diagnosis of acute respiratory tract infection were sent to the Molecular Laboratory of Ondokuz Mayıs University Faculty of Medicine Hospital. 24 different pathogens were detected by multiplex real time polymerase chain reaction method (Qiagen, Germany).

Results: Among the 3,131 samples analyzed, one or more viral agents were detected in 1,792 cases (57.23%). The most commonly identified viral pathogens were rhinovirus/enterovirus, respiratory syncytial virus A/B, SARS-CoV-2, and influenza A. The most frequently observed dual infection was influenza A combined with influenza A (subtype H1N1/2009), found in 22% of the co-infected cases. Seasonal distribution analysis showed that rhinovirus/enterovirus infections were more common in autumn, influenza A in winter, and respiratory syncytial virus A/B and SARS-CoV-2 in both autumn and winter months.

Conclusion: Determination of the viral pathogen by molecular methods in respiratory tract infections will be useful in preventing unnecessary use of antibiotics with correct treatment. It will also contribute to rapid and accurate management decisions with infection control measures and shorten hospitalization periods.

Keywords: Coinfection, multiplex PCR, respiratory viral panel

ÖZ

Amaç: Solunum yolu patojenlerinin neden olduğu enfeksiyonlar, yüksek morbidite ve mortaliteye yol açmakta olup halk sağlığı açısından önemli bir tehdit oluşturmaktadır. Bu çalışmada, 2021-2025 yılları arasında üniversite hastanemize başvuran hastalarda yaygın solunum yolu virüslerinin prevalansının ve mevsimsel dağılımının belirlenmesi amaçlanmıştır.

Gereç ve Yöntemler: Ondokuz Mayıs Üniversitesi Tıp Fakültesi Hastanesi Moleküler Laboratuvarı'na 1 Ocak 2021-1 Ocak 2025 tarihleri arasında akut solunum yolu enfeksiyonu ön tanısı olan hastalardan alınan 3131 nazal sürüntü örneği gönderilmiştir. 24 farklı patojenin saptanabildiği multipleks real time polimeraz zincir reaksiyonu yöntemi (Qiagen, Almanya) ile viral etkenlerin varlığı araştırılmıştır.

Bulgular: Toplam 3131 örneğin 1792'sinde (%57,23) bir veya daha fazla etken tespit edilmiştir. Enfeksiyona neden olan en yaygın viral etkenler rinovirüs/enterovirüs, respiratuar sinsityal virüs A/B, SARS-CoV-2, influenza A'dır. En yaygın ikili etken ise %22 prevalans ile influenza A+influenza A (alt tip H1N1/2009) olmuştur. Mevsimsel olarak değerlendirildiğinde; sonbaharda rinovirüs/enterovirüs, kışın influenza A, sonbahar-kış aylarında ise respiratuar sinsityal virüs A/B ve SARS-CoV-2'nin daha sık enfeksiyona neden olduğu görülmüştür.

Sonuç: Solunum yolu enfeksiyonlarında moleküler yöntemlerle viral patojenin belirlenmesi, doğru tedavi ile antibiyotiklerin gereksiz kullanımını önlemede yararlı olacaktır. Aynı zamanda enfeksiyon kontrol önlemleriyle hızlı ve doğru yönetim kararlarının alınmasına ve hastanede yatış sürelerinin kısalmasına katkı sağlayacaktır.

Anahtar Kelimeler: Koenfeksiyon, multipleks PZR, solunum yolu viral panel

Corresponding Author: Canberk Çınar ✉ cinarcanberk57@gmail.com **Received:** 23.05.2025 **Revision:** 24.11.2025 **Accepted:** 09.12.2025

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- The comprehensive prevalence of respiratory viruses in the region was determined with 3,131 samples during the 2021–2025 period.
- The viral positivity rate was 57.23%, with rhinovirus/enterovirus identified as the most dominant agent.
- High rates of co-infection were detected among influenza A subtypes, revealing simultaneous circulation dynamics.
- Distinct seasonal peaks were observed for rhinovirus/enterovirus in autumn and for influenza A in winter.
- Multiplex PCR was shown to provide rapid and accurate diagnosis, reducing unnecessary antibiotic use and improving clinical management.

INTRODUCTION

Respiratory tract infections (RTIs) are more commonly observed in children but also cause morbidity and mortality in adults, representing a significant public health concern worldwide (1). While mortality due to respiratory viral infections is low among healthy individuals in developed countries, it is higher in less developed regions, particularly among children, with approximately 2 million children aged 0–4 years dying annually worldwide (2,3). RTIs can be caused by bacteria, viruses, fungi, and parasites; however, viruses are known to be the primary causative agents in most cases. The prevalence and seasonal distribution of these microorganisms can vary between countries and even between regions within the same country (4,5). The most commonly encountered viral pathogens include influenza viruses, rhinoviruses, respiratory syncytial virus (RSV), human coronaviruses, and parainfluenza viruses (6). Clinically and radiologically, RTIs often present with similar features, and the absence of pathogen-specific signs or symptoms makes it difficult to differentiate between causative agents. Definitive diagnosis largely relies on microbiological testing. High-sensitivity, high-performance tests capable of detecting multiple groups of microorganisms simultaneously and rapidly have become indispensable. Accurate and timely identification of RTI pathogens enables prompt initiation of antiviral therapy and prevents unnecessary antibiotic use, which is particularly valuable given the current challenges of bacterial resistance. Furthermore, it contributes to reducing hospital stay, the risk of nosocomial transmission, and treatment costs (7–10). The aim of this study was to investigate the prevalence and seasonal distribution of respiratory pathogens in patients presenting to Ondokuz Mayıs University Faculty of Medicine with a preliminary diagnosis of acute RTI between 2021 and 2025.

MATERIAL and METHODS

Prior to the study, approval was obtained from the Clinical Research Ethics Committee of Ondokuz Mayıs University Faculty of Medicine (Date: 2025, Decision No: 47). A total of 3,131 nasal swab samples collected between 2021 and 2025 at the Molecular Laboratory of the Department of Medical Microbiology, Ondokuz Mayıs University, were included in the study. For patients who provided multiple samples within a month, only the first sample was included. The samples, transported in viral transport medium (UTM, Copan Diagnostics, Italy) under cold chain conditions, were analyzed at the Medical Microbiology Laboratory using the respiratory syndromic panel kit (Qiagen, Germany) according to the manufacturer's instructions. The respiratory syndromic viral panel (Qiagen, Germany) was performed following the recommendations of the manufacturer. The panel was capable of detecting the following pathogens within approximately 70 minutes: Adenovirus, Bocavirus, *Bordetella pertussis, Chlamydophila pneumoniae*, Coronavirus 229E, HKU1, NL63, OC43, Human metapneumovirus A/B, Influenza A, Influenza A (subtype H1), Influenza A (subtype H1N1/2009), Influenza A (subtype H3), Influenza B, *Mycoplasma pneumoniae*, Parainfluenza types 1–4, Rhinovirus/Enterovirus, RSV A/B, SARS-CoV-2, and *Legionella pneumophila.*

RESULTS

Among the total 3,131 samples received in our laboratory, at least one viral pathogen was detected in 1,792 cases (57.23%). The detected pathogens were distributed as follows: Rhinovirus/Enterovirus in 39.8%, RSV in 14.5%, SARS-CoV-2 in 13.6%, and Influenza A in 11.3% (Table 1).

Dual infections were detected in 412 patients (23%), triple infections in 63 patients (3.5%), and quadruple infections in 10 patients (0.56%). Among dual infections, the most common combination was Influenza A + Influenza A (subtype H1N1/2009) at 22% (n=91), followed by Influenza A + Influenza A (subtype H3) at 16.5% (n=68), and Rhinovirus/Enterovirus + RSV A/B at 9.4% (n=39). Among triple infections, the most frequently observed combination was Influenza A + Influenza A (subtype H1N1/2009) + Rhinovirus/Enterovirus at 7.9% (n=5). The seasonal distribution of Rhinovirus/Enterovirus, SARS-CoV-2, RSV A/B, and influenza A infections is shown in Figure 1.

Table 1: Annual distribution of respiratory pathogens

Pathogen	2021		2022		2023		2024		Total	
	n	%	n	%	n	%	n	%	n	%
Rhinovirus/Enterovirus	117	40.3	176	37.2	213	43	209	39	715	39.8
RSV A/B	90	31	75	15.8	51	10.3	44	8.2	260	14.5
SARS-CoV-2	34	11.7	70	14.8	72	14.5	69	12.8	245	13.6
Influenza A	15	5.1	58	12.2	46	9.2	85	15.8	204	11.3
Adenovirus DNA	5	1.7	58	12.2	43	8.6	33	6.1	139	7.7
Parainfluenza 3	47	16.2	20	4.2	32	6.4	22	4.1	121	6.7
Influenza A (subtype H1N1/2009)	0	0	3	0.6	28	5.6	75	14	106	5.9
Human Metapneumovirus A/B	5	1.7	34	7.2	35	7	24	4.4	98	5.4
Influenza A (subtype H3)	15	5.1	54	11.4	16	3.2	4	0.7	89	4.9

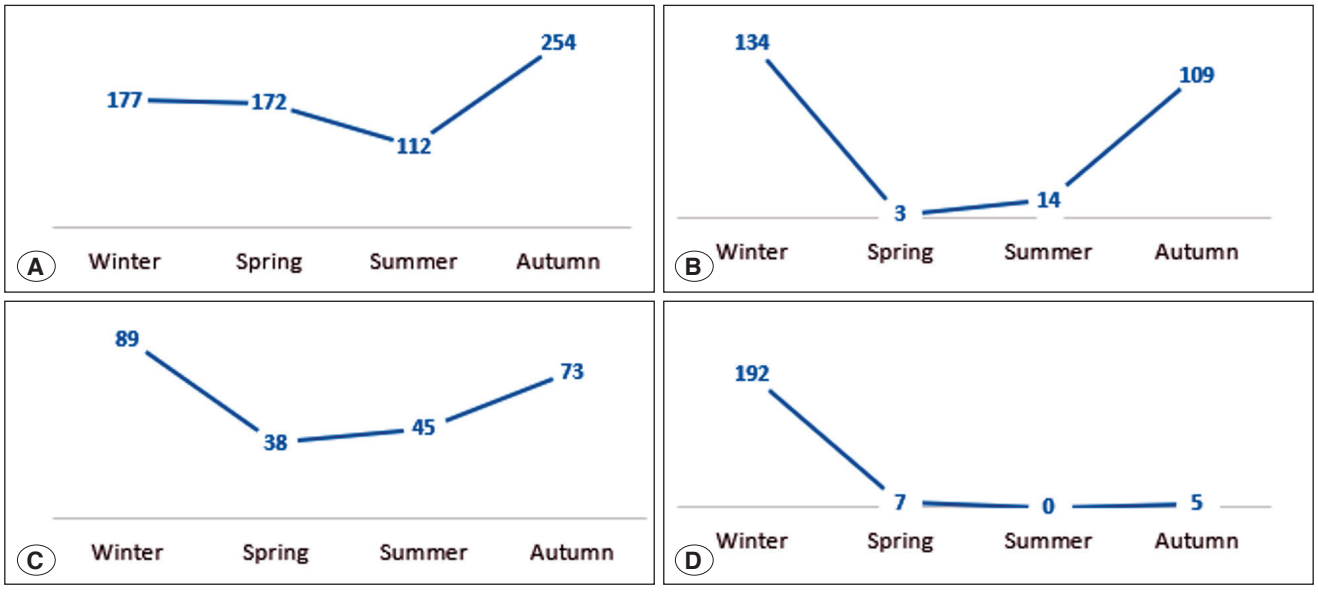


Figure 1: Analysis of the seasonal variations of detected pathogens. **A)** Rhinovirus/Enterovirus, **B)** SARS-CoV-2, **C)** RSV A/B, **D)** Influenza A.

(A) Rhinovirus/Enterovirus cases remained high during winter and spring, showed a marked decrease in summer, and increased sharply again in autumn. (B) SARS-CoV-2 cases were elevated in winter, dropped sharply in spring, remained low during summer, and rose again in autumn. (C) RSV A/B infections were highest in winter, reached their lowest level in spring, showed a slight increase in summer, and remained at moderate levels in autumn. (D) Influenza A cases peaked in winter, declined substantially in spring and especially in summer, and showed a mild increase in autumn. Overall, the data demonstrate distinct seasonal patterns among respiratory pathogens.

DISCUSSION

Respiratory tract infections are among the most common infectious diseases worldwide, increasing healthcare visits, hospitalizations, and work productivity losses, thereby imposing a significant economic burden (8). Acute RTIs account for a large proportion of all acute morbidities in developed countries. They are a leading cause of hospitalization among infants and young children in developed regions and a significant cause of mortality in developing countries (4). Studies using molecular diagnostic methods report detection rates of pathogens causing RTIs ranging from 30.9% to 96.1%, depending on the characteristics of the diagnostic kit, the demographic features of the patient population, and the timing of the study (11). While RTIs may be caused by bacteria, viruses, fungi, and parasites, viral agents are widely recognized as the predominant etiological factor (4). Accurate identification of viral pathogens cannot rely solely on clinical symptoms, as the manifestations are often non-specific and overlap among different viruses.

Therefore, laboratory-based tests are essential for etiological confirmation (7). The availability of specific antiviral treatments for influenza further emphasizes the importance of early and precise diagnosis. Recent advances in nucleic acid amplification techniques initially offered high sensitivity and rapidity, while modern multiplex systems now allow simultaneous detection of multiple pathogens, significantly improving diagnostic workflows (12, 13). Epidemiological data on respiratory pathogens show regional and seasonal variations. For example, RSV and influenza viruses are more common in winter, whereas rhinoviruses are prevalent from autumn through summer (5). Other studies have reported peaks in rhinovirus during October–November, and influenza virus, RSV, and human metapneumovirus during January–March (14). Influenza A is most frequent in winter, influenza B in spring, rhinovirus in autumn and spring, coronaviruses in summer, RSV in winter, and parainfluenza viruses in summer–autumn. Metapneumovirus and adenovirus infections also peak in winter (15, 16). In a study analyzing 5,102 clinical samples, approximately one-third of cases (33.4%) were positive for respiratory viruses, highlighting the dominant role of viral agents. Co-infections were observed in 18.6% of cases, with RSV + influenza A and RSV + parainfluenza being the most frequent combinations, indicating epidemiological co-circulation (5). In three-year data from Özdamar and Türkoğlu, at least one viral agent was detected in 83.5% of 283 samples, with the most common viruses being rhinovirus, adenovirus, influenza A/B, RSV A/B, human metapneumovirus A/B, and parainfluenza type 1 (17). Kuşkucu et al. analyzed 788 nasopharyngeal samples from patients diagnosed with acute RTIs between January 2010 and June 2018 using multiplex PCR. At least one viral agent was detected in 51.8% of samples, and multiple viruses in 7.2% of patients. The most frequently isolated agents were influenza A and B (11.4% and 5.5%), rhinovirus (14.9%), coronavirus (8.6%), RSV (7.1%), human metapneumovirus (4.1%), adenovirus (3.9%), and parainfluenza (3.6%) (15). Enterovirus and bocavirus were detected at rates below 1%. A single-center study at Süleyman Demirel University in 2019 examined 120 patients with suspected acute RTIs. At least one pathogen was detected in 59.2% (n=71) of cases, with RSV A/B and rhinovirus being the most frequent (16). In another study from March 2019 to December 2021, 8,825 respiratory samples were analyzed, and at least one viral agent was detected in 2,156 cases (24.4%). Single infections accounted for 85.5% of positive cases, predominantly rhinovirus, RSV A/B, human coronavirus, adenovirus, influenza virus, human metapneumovirus A/B, parainfluenza 1–4, influenza B, enterovirus, and bocavirus. Co-infections were observed in 14.5% of positive cases, with the most common combinations being adenovirus + rhinovirus, adenovirus + enterovirus, enterovirus + rhinovirus, rhinovirus + coronavirus, and RSV A/B + rhinovirus (18). In our study, rhinovirus/enterovirus was detected in 39.8% of patients, with a higher prevalence in autumn and winter, consistent with the literature. Clinical manifestations are usually mild in healthy individuals but may be severe with a higher risk of complications in children under five, the elderly, and immunocompromised patients (6, 19). The use of respiratory viral panels provides valuable information to guide timely, targeted antimicrobial therapy, reduce unnecessary antibiotic use, and initiate antiviral treatment when appropriate (20). During pandemics, these panels can also serve as a clinical decision-support tool for predicting short-term outcomes and guiding patient management (21).

Conclusion

Although our study is limited by its single-center design and the absence of comprehensive clinical correlation, continuous surveillance of circulating respiratory viruses remains critically important. Such monitoring not only contributes to the prevention of unnecessary or inappropriate antibiotic use, but also provides valuable guidance for clinicians in establishing timely and accurate treatment strategies, optimizing patient management, and supporting infection-control practices. Furthermore, population-level viral surveillance offers essential data for shaping public health policies, enabling early recognition of seasonal trends, outbreak preparedness, and resource allocation. In this context, multiplex real-time PCR assays for respiratory pathogens represent a highly effective diagnostic approach, as they allow for the rapid and sensitive identification of the etiologic agents of acute RTIs, thereby facilitating early initiation of targeted therapy, reducing hospital length of stay, and improving overall clinical outcomes.

Author Contributions

Study conception and design: **Canberk Çınar, Sümeyye Özkaya, İlknur Bıyık,** data collection: **Yeliz Tanrıverdi Çaycı, Kemal Bilgin,** analysis and interpretation of results: **Demet Gür Vural, Asuman Birinci,** draft manuscript preparation: **Canberk Çınar, Kemal Bilgin, Asuman Birinci.** The author(s) reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Ethical Approval

Our study received ethical committee approval from the Ondokuz Mayıs University Faculty of Medicine Clinical Research Department, with decision number 2025/47.

REFERENCES

1. Oliva J, Terrier O. Viral and bacterial co-infections in the lungs: dangerous liaisons. Viruses. 2021;13(9):1725-1725. https://doi.org/10.3390/v13091725

2. Alison M. Kesson Mini-symposium: microbiological diagnostic procedures in respiratory infection respiratory virus infection. Paediatr Respir Rev. 2007;8(3):240-248. https://doi.org/10.1016/j.prrv.2007.07.003
3. Henrickson KJ. Cost-effective use of rapid diagnostic techniques in the treatment and prevention of viral respiratory infections. Pediatr. Ann. 2005;34(1):24-31. https://doi.org/10.3928/0090-4481-20050101-08
4. Mahony JB, Petrich A, Smieja M. Molecular diagnosis of respiratory virus infections. Crit Rev Clin Lab Sci. 2011;48(5-6):217-249. https://doi.org/10.3109/10408363.2011.640976
5. Çiçek C, Arslan A, Karakuş HS, Yalaz M, Saz EU, Pullukçu H, et al. Prevalence and Seasonal Distribution of Respiratory Viruses in Patients with Acute Respiratory Tract Infections, 2002-2014. Bull Microbiol. 2015;49(2):188-200. https://doi.org/10.5578/mb.9024
6. Tregoning JS, Schwarze J. Respiratory viral infections in infants: causes, clinical symptoms, virology, and immunology. Clin Microbiol Rev. 2010;23(1):74-98. https://doi.org/10.1128/CMR.00032-09
7. Kesson AM. Respiratory virus infections. Pediatr Respir Rev. 2007;8(3):240-248. https://doi.org/10.1016/j.prrv.2007.07.003
8. Bayrakdar F, Altaş AB, Korukluoğlu G. Seasonal Distribution of Respiratory Viruses in Turkey Between 2009 and 2012. J Turk Microbiol Soc. 2013;43(1):56-66.
9. Uyar M, Kuyucu N, Tezcan S, Aslan G, Taşdelen B. Determination of the Frequency of Human Bocavirus and Other Respiratory Viruses Among 0-2 Years Age Group Children Diagnosed as Acute Bronchiolitis. Bull Microbiol. 2014;48(2):242-258. https://doi.org/10.5578/mb.7575
10. Sancaklı Ö, Yenigün A, Kırdar S. Results of Polymerase Chain Reaction in Nasopharyngeal Swab Specimens of Patients with Lower Respiratory Tract Infection. J Pediatr Inf. 2012;6(3):84-89. https://doi.org/10.5152/ced.2012.26
11. Goka EA, Vallely PJ, Mutton KJ, Klapper PE. Single and multiple respiratory virus infections and severity of respiratory disease: a systematic review. Pediatr Respir Rev. 2014;15(4):363-370. https://doi.org/10.1016/j.prrv.2013.11.001
12. Nijhuis R, Guerendiain D, Claas E, Templeton K. Comparison of ePlex respiratory pathogen panel with laboratory-developed real-time PCR assays for detection of respiratory pathogens. J Clin Microbiol. 2017;55(6):1938-1945. https://doi.org/10.1128/JCM.00221-17
13. Brotons P, Henares D, Latorre I, Cepillo A, Launes C, Muñoz-Almagro C. Respiratory Pathogen Panel and Anyplex II RV16 Tests for Multiplex Detection of Respiratory Pathogens in Hospitalized Children. J Clin Microbiol. 2016;54(12):2900-2904. https://doi.org/10.1128/JCM.01243-16
14. Appak Ö, Duman M, Belet N, Sayiner AA. Viral respiratory infections diagnosed by multiplex polymerase chain reaction in pediatric patients. J Med Virol. 2019;91(5):731-737. https://doi.org/10.1002/jmv.25379
15. Kuşkucu MA, Mete B, Tabak F, Midilli K. Prevalence and Seasonal Distribution of Respiratory Viruses in Adults Between 2010-2018. J Turk Microbiol Soc. 2020;50(1):21-26. https://doi.org/10.5222/TMCD.2020.021
16. Şirin MC. Investigation Of The Prevalence Of Viral And Bacterial Agents By Multiplex Pcr Method In Patients With A Pre-Diagnosis Of Acute Respiratory Tract Infection In Isparta Province. Med J SDU. 2022;29(2):171-178. https://doi.org/10.17343/sdutfd.990103
17. Özdamar M, Türkoğlu S. Detection of respiratory pathogens in lower respiratory tract infections by multiplex real time PCR in Kocaeli/Istanbul region in 2015-2017. Medeniyet Med J. 2018;33(3). https://doi.org/10.5222/MMJ.2018.71235
18. Aydoğan S, Kırca F, Gözalan A, Toyran A, Başyiğit T, Omay İ, et al. Viral Respiratory Tract Infection Agents Detected Between Years 2019-2021, COVID-19, and Co-Infections. Bull Microbiol. 2023;57(4):650-659. https://doi.org/10.5578/mb.20239952
19. Goktas S, Sirin MC. Prevalence and seasonal distribution of respiratory viruses during the 2014-2015 season in Istanbul. Jundishapur J Microbiol. 2016;9(9):e39132. https://doi.org/10.5812/jjm.39132
20. Covert K, Bashore E, Edds M, Lewis PO. Utility of the respiratory viral panel as an antimicrobial stewardship tool. J Clin Pharm Ther. 2021;46(2):277-285. https://doi.org/10.1111/jcpt.13326
21. Altunok İ, Özdemir S, Özkan A, Akça HŞ, İşat G, Kahraman HA, Gürlevik E. Frontal QRS-T angle in predicting short-term mortality in patients with SARS-CoV-2 infection in emergency service: an analytical study. Med J West Black Sea. 2023;7(2):167-173. https://doi.org/10.29058/mjwbs.1341878

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):65-69
DOI: 10.29058/mjwbs.1732952

Comparative analysis of pH measurements: Urine dipstick and pH meter

pH ölçümlerinin karşılaştırmalı analizi: İdrar dipstik ve pH metre

Murat Can, Berrak Güven, Osman Emre Ünlü

Department of Biochemistry, Zonguldak Bülent Ecevit University, Faculty of Medicine, Zonguldak, Türkiye

Cite this article as: *Can M et al. Comparative analysis of pH measurements: Urine dipstick and pH meter. Med J West Black Sea. 2026;10(1):65-69.*

ABSTRACT

Aim: Urine pH is a critical parameter in assessing an individual's metabolic and physiological status. The aim of this study was to compare dipstick measurement of urinary pH with gold standard pH meter readings.

Material and Methods: Each sample was tested by Jenway 3310 pH meter with a combined glass electrode and again using Meditape UC-9A urine dipsticks with Sysmex UC-3500 dipstick analyzer.

Results: A total of 3043 urine specimens were collected immediately after voiding from in-patients and out-patients at the Zonguldak Bulent Ecevit University Hospital. A strong correlation was found between urine dipstick and pH meter, with a rho value of 0.909 (95% CI 0,902 to 0.915), p<0.001. The bias values were between techniques was -0.18 (95% CI -0.1893 to -0.1695) with the two standard deviations establisheing the lower limit at –0.72 (95% CI -0.74 to -0.70) and the upper limit at 0.36 (95% CI 0.34 to 0.38), respectively.

Conclusion: Urine dipsticks offer expediency and practicality, making them suitable for routine checks and initial screenings. Conversely, pH meters deliver precision and reliability that is pivotal for comprehensive clinical assessments and research applications. Therefore, a pH meter with a glass electrode is mandatory when accurate measurement is necessary.

Keywords: Urine pH, urine dipstick, pH meter

ÖZ

Amaç: İdrar pH'ı, bir bireyin metabolik ve fizyolojik durumunu değerlendirmede kritik bir parametredir. Bu çalışmanın amacı, idrar pH'ının dipstik ölçümünü altın standart yöntem olan pH metre ile karşılaştırmaktır.

Gereç ve Yöntemler: Her numune, kombine cam elektrotlu Jenway 3310 pH metre ile ve ardından Sysmex UC-3500 idrar analizörü ile Meditape UC-9A idrar dipstikleri kullanılarak hemen test edildi.

Bulgular: Zonguldak Bülent Ecevit Üniversitesi Hastanesi'nde yatan ve ayaktan hastalardan toplam 3043 idrar örneği toplandı. İdrar dipstikleri ile pH metre arasında güçlü bir korelasyon bulundu, rho değeri 0,909 (95% CI 0,902 ila 0,915), p<0,001. Yöntemler arasındaki sapma değerleri -0,18 (95% CI -0,1893 ila -0,1695) idi, iki standart sapma alt sınırı -0,72 (95% CI -0,74 ila -0,70) ve üst sınırı 0,36 (95% CI 0,34 ila 0,38) olarak belirledi.

Sonuç: İdrar dipstikleri, rutin kontroller ve ilk taramalar için kolaylık ve pratiklik sunar. Buna karşılık, pH ölçerler kapsamlı klinik değerlendirmeler ve araştırma uygulamaları için çok önemli olan hassasiyet ve güvenilirlik sunar. Bu nedenle, doğru ölçüm gerektiğinde cam elektrotlu bir pH ölçer zorunludur.

Anahtar Kelimeler: İdrar pH, idrar dipstik, pH metre

Highlights

- This study compared urinary pH in 3,043 samples using a dipstick analyzer versus a glass-electrode pH meter.
- Although a strong correlation was observed (ρ = 0.909), this did not translate into acceptable clinical agreement.
- Bland–Altman analysis showed a negative bias and wide limits of agreement, indicating frequent measurement discrepancies.
- Clinically relevant errors occurred in approximately 7.5% of dipstick readings, often differing by ±1 pH unit.
- Accurate urine pH assessment—especially in pH-dependent therapies—requires a calibrated pH meter rather than dipstick testing.

Corresponding Author: Murat Can ✉ drcanmurat@yahoo.com
Received: 03.07.2025 **Revision:** 05.10.2025 **Accepted:** 22.12.2025

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

INTRODUCTION

Urine pH is a critical parameter in assessing an individual's metabolic and physiological status. It serves as a reflection of the body's acid-base balance and can provide valuable insights into various health conditions (1). The normal range for urine pH typically falls between 4.6 and 8.0, influenced by factors such as diet, hydration, and underlying health issues (2). For example, a diet rich in protein may lead to lower urine pH, reflecting increased acidity and conversely, a vegetarian diet often results in a more alkaline urine pH (2). Furthermore, hydration status significantly impacts urine concentration and consequently its pH (2).

Clinically, urine pH testing can be instrumental in diagnosing various disorders. For instance, a consistently high urine pH may suggest urinary tract infections (UTIs) caused by urease-producing bacteria, while a low pH can indicate conditions such as diabetic ketoacidosis or chronic respiratory acidosis (3-5). Moreover, urine pH can influence the solubility of certain substances, playing a critical role in the formation of kidney stones. Calcium oxalate stones, for example, tend to form in acidic urine, while struvite stones form more readily in alkaline urine (6).

Monitoring urine pH can also aid in managing certain medical treatments. Patients undergoing treatments for metabolic disorders or those on specific diets may need regular assessments of urine pH to evaluate the effectiveness of their management strategies (7,8).

Urine pH is a vital indicator of overall health, providing essential information regarding metabolic processes, dietary influences, and potential health complications. Its evaluation is an indispensable component of clinical diagnostics and management, thereby underscoring the necessity of further research into the diverse implications of urine pH variations in health and disease. The aim of this study was to compare dipstick measurement of urinary pH with gold standard pH meter readings.

MATERIAL and METHODS

A total of 3043 urine specimens were collected immediately after voiding from in-patients and out-patients at the Zonguldak Bulent Ecevit University Hospital. Samples were taken from both male and female patients. The protocol was approved by the Ethics Committee of Zonguldak Bulent Ecevit University (Date: January 13, 2025, Decision No: 2025/01).

Each sample was immediately tested by a fully trained operator via the two urinary pH testing methods. Once with the reference laboratory Jenway 3310 pH meter with a combined glass electrode (Jenway, Staffordshire, UK) and again using Meditape UC-9A urine dipsticks (Sysmex Corporation, Kobe, Japan). The pH meter was calibrated prior to use with pH solutions of pH 4, 7, and 10. The pH probe was placed in the urine sample and a pH reading was recorded for all samples tested with the previous methods. The Sysmex UC-3500 (Sysmex Corporation, Kobe, Japan) dipstick analyzer and an two levels of external daily quality control were performed.

Statistical analyses were conducted using SPSS Statistics Version 19.0 (IBM Corporation, Armonk, NY, USA) and MedCalc Statistical Software version 23.1.3 (MedCalc Software Ltd, Ostend, Belgium). Normality assumptions were controlled by the Shapiro–Wilk test. Descriptive analyzes were presented using mean±SD, median (min-max) or n (%), where appropriate. For pairwise comparisons between related samples, the paired samples t-test was used for normally distributed data. A two-tailed *p*-value of less than 0.05 was considered to be significant. To explore the linear relationships among the parameters, Spearman correlation coefficients and linear regression equations were calculated. The Bland-Altman test was applied to assess the agreement between pH results, with 95% confidence interval limits of agreement. We calculated the exact match concordance rate (%) and concordance rate (%) with ±1 grading difference rates (%) between Sysmex UC-3500 and Jenway 3310 pH meter. The sample size was calculated using G Power 3.1.9.4 software. The sample size calculation was based on a effect size 0.1, with a p-value of 0.05, a power of 95%, and a total sample size of at least 1073 subjects.

RESULTS

1450 of the specimens were from males, while 1593 were from females. The ages of the patients ranged from 1 to 95, with a mean age 52.0±18.9 median age of 55. A strong correlation between pH dipstick and pH meter was found, with a rho value of 0.909 (95% CI 0,902 to 0,915), p<0.001 (Figure 1). The linear function of pH dipstick for estimating pH meter was Y=-0.380 (95% CI -0,4650 to -0,2700) + 1.10 X (95% CI 1,0800 to 1,1150).

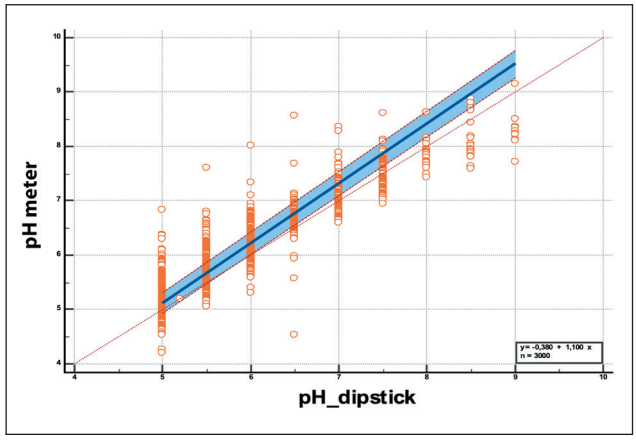


Figure 1: A comparative analysis of pH levels

Table 1. pH values of the patients with dipstick and pH meter

Dipstick pH Value	pHmeter				p-value
	Mean±SD	Difference±SD	Median	Min-Max	
5.0 (n=1036)	5.21±0.25	0.21±0.25	5.22	4.20-6.84	<0.001
5.5 (n=991)	5.69±0.24	0.19±0.24	5.61	5.05-7.61	<0.001
6.0 (n=424)	6.24±0.27	0.24±0.27	6.23	5.30-8.01	<0.001
6.5 (n=215)	6.63±0.28	0.13±0.28	6.62	5.54-8.57	<0.001
7.0 (n=172)	7.05±0.27	0.05±0.27	6.99	6.60-8.36	0.160
7.5 (n=107)	7.48±0.29	0.02±0.29	7.41	6.94-8.62	0.092
8.0 (n=47)	7.82±0.24	0.18±0.24	7.80	7.44-8.63	<0.001
8.5 (n=28)	8.14±0.39	0.36±0.39	8.02	7.59-8.88	<0.001
9.0 (n=22)	8.29±0.34	0.71±0.34	8.24	7.72-9.15	<0.001

Table 2. The distribution of the pH values

	pH	± 1 rank match concordance rate (%)	pHmeter (n)									
			<5.0	5-5.5	5.5-6.0	6.0-6.5	6.5-7.0	7.0-7.5	7.5-8.0	8.0-8.5	8.5-9.0	>9.0
Dipstick (n)	5.0 (n=1036)	90.3	**201**	**735**	93	6	1	0	0	0	0	0
	5.5 (n=991)	91.7	0	**207**	**702**	70	11	0	1	0	0	0
	6.0 (n=424)	83.9	0	3	**55**	**301**	63	1	0	1	0	0
	6.5 (n=215)	93.0	0	0	4	**47**	**153**	9	2	0	0	0
	7.0 (n=172)	94.7	0	0	0	0	**91**	**72**	7	2	0	0
	7.5 (n=107)	95.3	0	0	0	0	1	**63**	**39**	3	1	0
	8.0 (n=47)	87.3	0	0	0	0	0	2	**21**	**20**	4	0
	8.5 (n=28)	71.4	0	0	0	0	0	0	8	**9**	**11**	0
	9.0 (n=22)	40.1	0	0	0	0	0	0	1	8	**9**	4

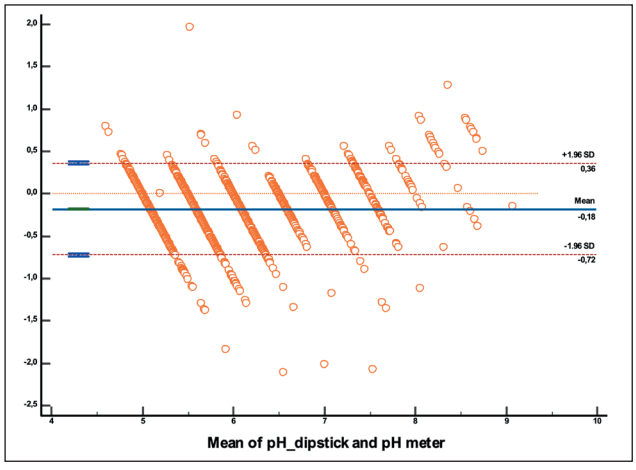


Figure 2: The means, mean differences, and agreement limits (bias ± 2 SD) of pH results according to Bland-Altman analyses

The Bland-Altman plots depict the differences between the means of pH dipstick versus pH meter (Figure 2). The bias values between techniques was -0.18 (95% CI -0.1893 to -0.1695) with the two standard deviations established the lower limit at –0.72 (95% CI -0.74 to -0.70) and the upper limit at 0.36 (95% CI 0.34 to 0.38), respectively.

The results of the comparison between the three methods when compared with the reference method are summarized in Tables 1 and 2.

DISCUSSION

The measurement of urine pH is a vital component of clinical diagnostics, offering critical insights into both metabolic and renal function. Fluctuations in urine pH can signal a variety of health issues, ranging from urinary tract infections to systemic metabolic disorders. Traditionally, urine pH has been evaluated using dipstick tests, which are popular due to their convenience and speed. However, the advent of pH meters presents a more accurate and dependable alternative. This essay will explore the advantages and limitations of both methods, highlighting their respective roles in clinical practice and research settings.

Dipstick tests for urine pH utilize a straightforward colorimetric method, wherein a reactive dipstick changes color in response to the acidity or alkalinity of a urine sample. In

a study conducted by Desai and Assimos, the pH levels of 100 fresh urine specimens from patients with urologic diseases were measured. The findings indicated that urinary dipsticks serve as reliable indicators of urinary pH, making them valuable tools for both physicians and patients to monitor pH manipulation therapy within the targeted range of 6 to 7 (9). Similarly, Johnson et al. analyzed 201 free-catch urine samples from 114 hospitalized dogs and concluded that urine reagent dipstick can provide estimates of pH for routine urinalysis (10). However, they cautioned against relying on these dipsticks when precise and consistent pH measurements are essential for the diagnosis, prevention, and management of diseases. This testing technique offers numerous advantages, particularly its ease of use and rapid results, making it an ideal choice for point-of-care testing in outpatient settings and laboratories with limited resources. Dipsticks are widely accessible, cost-effective, and require minimal training to operate, thereby facilitating routine screening. Nevertheless, while dipstick tests provide a quick overview of urine pH, they are inherently limited by their qualitative nature and the potential for subjective interpretation. Factors such as the freshness of the urine sample, laboratory conditions, and user variability can introduce inconsistencies in results, ultimately affecting diagnostic accuracy. Therefore, while dipstick tests are a valuable tool in clinical practice, they should be complemented with more precise methods when accuracy is paramount.

Mina et al. conducted a study involving 210 urine samples, utilizing the Mettler Toledo Seven Compact pH meter method, which demonstrated an 85.7% agreement with the dipstick method (11). In a separate evaluation, Coninck et al. assessed 77 urine aliquots from 20 healthy volunteers using reagent dipsticks, a portable pH meter, and a laboratory pH meter as the gold standard. Their findings indicated that the portable electronic pH meter provided more accurate readings compared to the reagent dipsticks according to the laboratory pH meter as the gold standard (12). Ilyas et al. collected urine samples from 200 patients attending stone clinics and emphasized the importance of using a pH meter for urinary pH readings in patients at risk for pH-dependent stone formation (13). In contrast to dipstick methods, pH meters offer a quantitative assessment of urine pH, delivering a higher level of precision and reproducibility. These meters operate on the principles of potentiometry, employing a glass electrode to accurately measure the hydrogen ion concentration in urine samples. This methodological rigor makes pH meters particularly well-suited for research environments and clinical investigations that demand high-quality data granularity. Moreover, pH meters provide definitive numerical values and facilitate the evaluation of urine samples under controlled conditions, thereby minimizing environmental influences and inter-user variability. However, the use of pH meters requires a certain level of technical proficiency and proper calibration, which can present challenges in settings where resources and trained personnel are limited. This study highlights that while the correlation between the two different modalities of urine pH measurement is statistically good, clinically significant discrepancies occur with an unacceptable frequency. Approximately 7.5% of dipstick measurements resulted in clinically relevant errors, with the majority deviating by one pH unit from the true urine pH value. Although a strong correlation coefficient (r=0.909) was observed between the two measurement methods, this does not necessarily imply clinical reliability. A correlation value only reflects the strength of a linear relationship, not the degree of agreement or interchangeability between methods. In clinical contexts such as urinary alkalinization therapy or stone prophylaxis, even a ±1 pH unit deviation can lead to inappropriate treatment decisions or missed therapeutic targets. Therefore, despite the statistically robust correlation, the observed discrepancies highlight that the dipstick method cannot substitute for pH meter measurements when accurate urine pH assessment is essential.

Wockenfus et al. conducted a study involving 116 urine samples from pediatric patients undergoing high-dose methotrexate therapy (14). They measured the pH of these samples using both the Clinitek Status dipstick and a pH meter. Their findings revealed a significant bias in pH readings among patients receiving high-dose methotrexate therapy, with an average discrepancy of 0.7 ± 0.4. Notably, this bias was most pronounced in urine samples with pH values exceeding 8. In our study, we observed a similar trend: at dipstick pH readings above 8.5, there was a tendency for the dipstick to undervalue the pH compared to the pH meter. A critical distinction between these two methods is their diagnostic utility. Urine pH levels can be influenced by various factors, including dietary habits, medications, and underlying health conditions. While dipstick tests may serve as a preliminary screening tool, they often lack the precision necessary for accurate clinical decision-making. In a cohort of 98 patients where pharmacological manipulation of urinary pH was indicated based on true pH measurements, it was found that 14 patients would not have received appropriate treatment, and 5 patients would have been unnecessarily medicated if only the dipstick pH values had been considered (15). In contrast, pH meters provide healthcare professionals with more reliable data, facilitating better patient management and therapeutic interventions. Our data indicate that while dipstick pH readings of 7 and 7.5 closely aligned with the mean pH meter readings, higher pH values were associated with a greater likelihood of overestimation by the dipstick method. Furthermore, when urine pH measurements are integrated with other diagnostic tests, the enhanced precision offered by pH meters significantly improves the interpretability of results, ultimately leading to more informed clinical judgments.

A limitation of our study is that dipstick pH results were treated as continuous variables for statistical analyses. Although this approach facilitates correlation and regression testing, it may not fully reflect the ordinal nature and potential rounding bias inherent in dipstick readings. Future studies employing categorical or interval-based analysis could provide additional insights into the agreement between the two methods.

Conclusion

Both dipstick tests and pH meters offer distinct advantages and limitations when it comes to measuring urine pH. Dipsticks are valued for their speed and convenience, making them ideal for routine checks and initial screenings. In contrast, pH meters provide the precision and reliability essential for thorough clinical assessments and research applications. The selection between these two methods should be tailored to the specific context, ideally combining the rapid results of dipsticks with the accuracy of pH meters to create a comprehensive diagnostic strategy. Enhancing our understanding and methodologies for measuring urine pH can significantly elevate its role in clinical practice, ultimately leading to improved patient outcomes and stimulating further research in metabolic health. Therefore, when precise measurements are critical, the use of a pH meter equipped with a glass electrode is indispensable.

Author Contributions

Study conception and design: **Murat Can, Berrak Güven;** data collection: **Murat Can, Berrak Güven, Osman Emre Ünlü;** analysis and interpretation of results: **Murat Can, Berrak Güven, Osman Emre Ünlü;** draft manuscript preparation: **Murat Can, Berrak Güven, Osman Emre Ünlü**. The author(s) reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Financial Support

The authors did not receive financial support for the study.

Ethical Approval

The research was approved by the Research Ethics Committee at Ethics Committee of Zonguldak Bulent Ecevit University (Date: January 13, 2025, Decision No: 2025/01).

REFERENCES

1. Eknoyan G. Looking at the urine: the renaissance of an unbroken tradition. Am J Kidney Dis. 2007;49(6):865-872. https://doi.org/10.1053/j.ajkd.2007.04.003
2. Welch AA, Mulligan A, Bingham SA, Khaw KT. Urine pH is an indicator of dietary acid-base load, fruit and vegetables and meat intakes: results from the European Prospective Investigation into Cancer and Nutrition (EPIC)-Norfolk population study. Br J Nutr. 2008;99(6):1335-1343. https://doi.org/10.1017/S0007114507862350
3. Lai HC, Chang SN, Lin HC, Hsu YL, Wei HM, Kuo C, et al. Association between urine pH and common uropathogens in children with urinary tract infections. J Microbiol Immunol Infect. 2021;54(2):290-298. https://doi.org/10.1016/j.jmii.2019.08.016
4. Ogawa S, Nako K, Okamura M, Ito S. Lower urinary pH is useful for predicting renovascular disorder onset in patients with diabetes. BMJ Open Diabetes Res Care. 2015;3(1):e000097. https://doi.org/10.1136/bmjdrc-2015-000097
5. Mittal K, Aggarwal HK, Rungta N, Patki V. Respiratory acidosis and alkalosis. J Pediatr Crit Care. 2021;8(3):161-164. https://doi.org/10.21304/2021.0803.007
6. Wang Z, Zhang Y, Zhang J, Deng Q, Liang H. Recent advances on the mechanisms of kidney stone formation. Int J Mol Med. 2021;48(2):149. https://doi.org/10.3892/ijmm.2021.4962
7. Galan-Llopis JA, Torrecilla-Ortiz C, Luque-Gálvez MP, Peris-Nieto X, Cuñé-Castellana J. Urinary pH as a target in the management of lithiasic patients in real-world practice: monitoring and nutraceutical intervention for a nonlithogenic pH range. Clin Med Insights Urol. 2019;12:1179561119877343. https://doi.org/10.1177/1179561119877343
8. Ordaz G, Dagà U, Budia A, Pérez-Lanzac A, Fernández JM, Jordán C. Urinary pH and antibiotics, choose carefully: a systematic review. Actas Urol Esp. 2023;47(7):408-415. https://doi.org/10.1016/j.acuro.2023.02.003
9. Desai RA, Assimos DG. Accuracy of urinary dipstick testing for pH manipulation therapy. J Endourol. 2008;22(6):1367-1370. https://doi.org/10.1089/end.2008.0063
10. Johnson KY, Lulich JP, Osborne CA. Evaluation of the reproducibility and accuracy of pH-determining devices used to measure urine pH in dogs. J Am Vet Med Assoc. 2007;230(3):364-369. https://doi.org/10.2460/javma.230.3.364
11. Mina A, Stathopoulos J, Sinanian T, McNeice L, Holmes D, Fletcher KL, et al. Comparison of different methods used in drugs of abuse for sample validity testing including pH methods, specific gravity methods, TECO™ drug adulteration test strip and oxidant assay. Adv Lab Med. 2021;2(4):550-557. https://doi.org/10.1515/almed-2021-0035
12. Coninck V, Keller EX, Monsalve MR, Doizi S, Audouin M, Haymann JP, et al. Evaluation of a portable urinary pH meter and reagent strips. J Endourol. 2018;32(7):647-652. https://doi.org/10.1089/end.2017.0855
13. Ilyas R, Chow K, Young JG. What is the best method to evaluate urine pH? A trial of three urinary pH measurement methods in a stone clinic. J Endourol. 2015;29(1):70-74. https://doi.org/10.1089/end.2014.0408
14. Wockenfus AM, Koch CD, Conlon P, Sorensen L, Cambern K, Chihak A, et al. Discordance between urine pH measured by dipstick and pH meter: implications for methotrexate administration protocols. Clin Biochem. 2013;46(1-2):152-154. https://doi.org/10.1016/j.clinbiochem.2012.10.009
15. Kwong T, Robinson C, Spencer D, Wiseman OJ, Frankl F. Accuracy of urine pH testing in a regional metabolic renal clinic: is the dipstick accurate enough? Urolithiasis. 2013;41(2):129-132. https://doi.org/10.1007/s00240-012-0526-3

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):70-80
DOI: 10.29058/mjwbs.1733561

Applications and research trends of large language models in healthcare: A bibliometric analysis

Sağlık hizmetlerinde büyük dil modellerinin uygulamaları ve araştırma eğilimleri: Bibliyometrik bir analiz

Hakan Öztürk[1], **Elvan Hayat**[2]

[1]Department of Biostatistics, Aydın Adnan Menderes University, Faculty of Medicine, Aydın, Türkiye
[2]Department of Econometrics, Aydın Adnan Menderes University, Faculty of Political Sciences, Aydın, Türkiye

Cite this article as: *Öztürk H and Hayat E. Applications and research trends of large language models in healthcare: a bibliometric analysis. Med J West Black Sea. 2026;10(1):70-80.*

ABSTRACT

Aim: This study aims to reveal the current scientific studies and research trends on the use of Large Language Models (LLM) in the field of health by bibliometric analysis.

Material and Methods: Data were sourced from open-access clinical research articles published between 2023 and 2025, extracted from the Web of Science database. A total of 173 articles meeting the inclusion criteria were analyzed using the “Bibliometrix” R package and the “Biblioshiny” interface, focusing on publication trends, co-authorship networks, keyword co-occurrences, and thematic mapping.

Results: Results revealed a substantial rise in academic interest in LLMs within healthcare, with significant contributions from researchers in the United States, China, and various European countries. Prominent keywords such as "Artificial Intelligence," "ChatGPT," and "Medical Education" indicate that LLMs are increasingly explored for educational and clinical support applications. Thematic analysis identified emerging research areas including "performance," "health disparities," and "ethical challenges." While LLMs offer considerable opportunities in healthcare, they also present notable risks like data privacy issues, misinformation, and ethical oversight gaps.

Conclusion: In conclusion, this study provides a comprehensive overview of LLM applications in healthcare, highlighting key research themes and identifying gaps within the current literature.

Keywords: Bibliometric analysis, ChatGPT, generative artificial intelligence, large language models, medicine

ÖZ

Amaç: Bu çalışma, sağlık alanında Büyük Dil Modelleri’nin (LLM) kullanımına ilişkin mevcut durumu ve araştırma eğilimlerini bibliyometrik analiz yöntemiyle ortaya koymayı amaçlamaktadır.

Gereç ve Yöntemler: Veriler, 2023-2025 yılları arasında yayımlanmış açık erişimli klinik araştırma makaleleri arasından seçilerek Web of Science veri tabanından elde edilmiştir. Dahil etme kriterlerini karşılayan 173 makale, R programında “Bibliometrix” paketi ve “Biblioshiny” arayüzü kullanılarak analiz edilmiştir. Yayın eğilimleri, ortak yazarlık ağları, anahtar kelime eş-oluşumları ve tematik haritalama gibi temel bibliyometrik göstergeler incelenmiştir.

Bulgular: Bulgular, sağlık alanında LLM’lere yönelik bilimsel ilginin belirgin şekilde arttığını ve ABD, Çin ve bazı Avrupa ülkelerinin bu alandaki araştırmalara öncülük ettiğini göstermektedir. “Yapay Zekâ”, “ChatGPT” ve “Tıp Eğitimi” gibi anahtar kelimelerin öne çıkması, LLM’lerin eğitim ve klinik destek sistemlerinde giderek daha fazla araştırıldığını ortaya koymaktadır. Tematik analizde, “performans”, “sağlık eşitsizlikleri” ve “etik zorluklar” gibi konuların yükselen araştırma alanları olduğu görülmüştür. LLM’ler sağlık hizmetlerinde önemli fırsatlar sunarken, veri gizliliği, yanlış bilgi üretimi ve etik denetim eksiklikleri gibi riskler de barındırmaktadır.

Sonuç: Sonuç olarak, bu çalışma sağlık alanında LLM’lerin kullanımına yönelik kapsamlı bir haritalama sunarak, öne çıkan araştırma temalarını ve literatürdeki boşlukları ortaya koymaktadır.

Anahtar Kelimeler: Bibliyometrik analiz, ChatGPT, üretken yapay zeka, büyük dil modelleri, tıp

Corresponding Author: Hakan Öztürk ✉ h.ozturk3509@gmail.com **Received:** 03.07.2025 **Revision:** 28.11.2025 **Accepted:** 24.12.2025

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- The study identifies influential sources and emerging hotspots in healthcare-related Large Language Models research.
- WoS-based mapping of 2023-early 2025 publications characterizes the post-ChatGPT landscape of LLM research in healthcare.
- Output and collaborations cluster in major hubs (notably the USA and China) and high-capacity institutions, while Turkey remains peripheral in co-authorship networks.
- Core themes focus on AI/ChatGPT and performance; acceptance, risk, and equity/governance remain less mature, indicating key gaps and priorities for future work.

INTRODUCTION

Large Language Models (LLMs), in parallel with recent advances in natural language processing, offer substantial potential for health sciences and their applications (1). As healthcare increasingly depends on digital tools and data-driven decision-making, understanding how LLMs contribute to clinical, educational, and administrative settings becomes essential. The ability of these models to analyze and interpret complex data has the potential to transform the way healthcare is delivered, from diagnosis to treatment, from drug discovery to personalized medicine. Their superior performance in natural language processing tasks is particularly advantageous in analyzing text-based health data (1). LLMs can improve patient compliance by making medical instructions more understandable and reducing communication errors (2). By analyzing data from patient records, LLMs can improve adherence to medical prescriptions and minimize misunderstandings between healthcare professionals and patients. These models can also improve health literacy by rephrasing medical information in a language that non-experts can understand, helping patients make more informed decisions about their health (3). LLM applications in health sciences are not only limited to patient care and communication, but also play an important role in developing clinical decision support systems, predicting disease progression and accelerating medical research (4).

Since large language models, such as ChatGPT and other generative artificial intelligence (AI) models, are trained on massive amounts of health and medical data, they can develop specialized knowledge specific to different medical disciplines. In this way, they can make significant contributions to healthcare professionals in areas such as clinical decision support, early diagnosis of diseases and drug discovery. By analyzing the information in the medical literature, LLMs can specialize in different medical fields such as radiology, pathology and oncology and follow the latest developments in these fields (4,5). However, the potential benefits of LLMs in the healthcare field also come with some risks and limitations that need to be carefully considered. Although LLMs have human-like abilities to analyze and interpret medical information, they may not always produce accurate and reliable results (6). It should be kept in mind that the texts produced by these models may be incompatible with clinical and societal values and may contain misleading medical information (7). The use of LLMs, especially in clinical tasks, requires extensive expert training and rigorous validation processes. The use of models such as ChatGPT by the general public as a substitute for professional medical advice carries potential risks. Developers, healthcare professionals, ethics experts and regulatory authorities should act in collaboration (5). Transparency, accountability and ethical principles should be observed in the application of LLMs in health (8,9).

Bibliometric analysis is a method that aims to examine research trends, collaboration networks and academic impacts based on quantitative data of scientific publications. It maps knowledge production processes through criteria such as publication numbers, citation data, inter-author relationships and keyword analysis. This method helps to understand the body of knowledge in a particular field and identify prominent themes and gaps. Especially in rapidly changing disciplines such as health sciences, it provides researchers with an important tool for both systematically reviewing existing literature and providing strategic directions for future research (10,11).

Despite the rapidly increasing number of studies on LLM research in healthcare, comprehensive evaluations that systematically map the intellectual structure, thematic evolution, collaboration patterns, and emerging hotspots of the field remain limited. Existing publications primarily examine technical capabilities, ethical considerations, or specific application areas, but they do not offer an overarching, data-driven picture of how the field is evolving over time. In particular, research focusing on the post-ChatGPT era (2023 onwards), when the volume and diversity of publications noticeably increased, is still limited (12,13). Although

the field has expanded rapidly, the extent and characteristics of this acceleration have yet to be systematically documented. Therefore, this study aims to address this gap by conducting a comprehensive bibliometric analysis of LLM-related publications in the health sciences, identifying major research themes, collaboration networks, influential authors, and emerging trends.

MATERIAL and METHODS

Data Sources, Data Collection and Search Strategies

This research is a descriptive study in which studies published in the Web of Science (WoS) database between 2023 and 2025 on the use of LLM models in the field of health sciences are examined by bibliometric analysis method. The year 2023 was chosen as the starting point because it coincides with the period following the public release of widely accessible generative AI tools such as ChatGPT, when applications of LLMs in healthcare began to expand rapidly. The search query was constructed using key terms reflecting the scope of the study, including "Large Language Model*", "LLM*", "ChatGPT", "Generative AI", "Health", "Healthcare", "Medical", "Medical Practice", "Clinical Practice" and "Medicine". The search was conducted on the WoS database in a single session on 22.02.2025.

In the first screening, 1458 articles were found. The inclusion criteria were that they were published between 2023 and 2025, open access and full-text available in English. Exclusion criteria were articles without access to the full text, review articles and irrelevant studies based on manual evaluation of titles and abstracts. As a result, publications from 2025 that had been accepted or were in press but had not yet been indexed in WoS by 22.02.2025 could not be captured, which may lead to underestimation of the total number of publications for that year.

The flowchart created by taking into account the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines is presented in Figure 1. In this diagram, the boxes on the main vertical axis show the number of records retained at each PRISMA phase (identification, screening, eligibility and inclusion), while the boxes on the right-hand side indicate the reasons for exclusion and how many records were removed at each step. In the first search, 818 articles were accessed by selecting open access articles from the 1458 articles. Then, review articles (n=81) and enriched cited references (n=198) were removed and the years 2023-2024 and 2025 were selected. Thus, 539 articles were accessed. When the document type "article" was selected among these, 173 articles were identified.

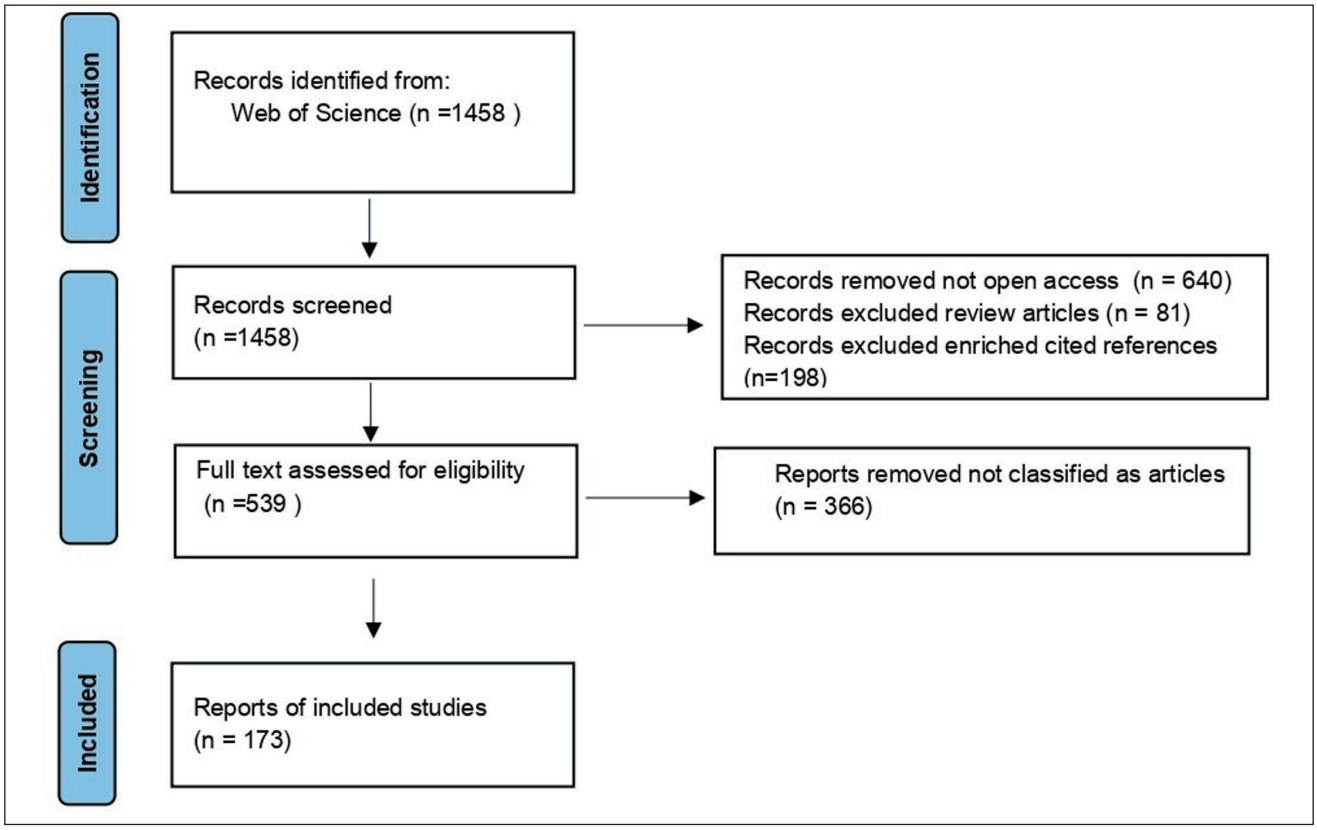


Figure 1: PRISMA flow diagram of the study selection process for LLM-related publications in healthcare

When these articles were analyzed according to Web of Science indexes, it was determined that 103 articles were published in SCI indexed journals, 6 articles in SSCI indexed journals and 64 articles in ESCI indexed journals.

The bibliometric analysis was conducted using the Biblioshiny web interface of the Bibliometrix R package. Analyses were conducted in R using the Bibliometrix package and its Biblioshiny graphical interface to perform both descriptive bibliometrics (e.g., annual scientific production, most productive authors, sources and countries) and science mapping (e.g., co-authorship, co-citation and keyword co-occurrence networks). Bibliometrix provides a comprehensive framework for importing bibliographic data from various academic databases, including WoS, and performing bibliometric analyses (11).

In this study, various bibliometric indicators were utilized, including document analysis to identify the most frequently used terms in the literature, keyword cloud visualization to highlight prominent topics, and keyword co-occurrence analysis to detect emerging research trends. Co-authorship (authors, institutions, countries) and keyword co-occurrence networks were generated using full counting, and the association strength normalization method was applied to construct the similarity matrices. For network visualization and community detection, the clustering procedure implemented in Bibliometrix/Biblioshiny (Louvain algorithm) was used to identify groups of closely related authors, countries and topics. To enhance interpretability and reduce visual noise, only items with at least three occurrences (for keywords) or at least two documents (for authors and countries) were included in the network maps, and isolated nodes were excluded. This approach allowed for a detailed examination of the evolution and significance of the topics covered in the academic literature.

RESULTS

As a result of the bibliometric analysis of 173 articles that met the inclusion criteria in the study, it was determined that these articles were published in 93 different sources, the number of authors working on this subject was 961, and the number of single-author articles was 13.

Most Relevant Sources

"Most Relevant Sources", which shows the journals with the highest number of publications and citations in the relevant academic field, is given in Table 1. In this study, the term "most relevant sources" refers to journals with the highest number of LLM-related articles within the analyzed WoS corpus. Table 1 shows the most important academic journals ranked according to the number of publications and the number of articles in these journals. Specifically, Table 1 presents the top 10 journals that published LLM-related articles in health sciences, ranked in descending order according to the number of articles identified in our dataset. The column "Sources" lists the names of the journals, while the column "Articles" indicates the number of publications from each journal. Accordingly, Journal of Medical Internet Research (JMIR) stands out as the journal with the highest number of articles published with 18 publications. Cureus Journal of Medical Science (14 publications) and JMIR Medical Education (13 publications) rank second and third, respectively. IEEE Access (5 publications), JMIR Mental Health (5 publications) and PLOS ONE (5 publications) also make notable contributions.

Most Relevant Affiliations

The institutions (affiliations) that produce the highest number of scientific publications in the relevant academic field are presented in Table 2. Table 2 presents universities and research institutions, along with the number of LLM-relat-

Table 1: Most Relevant Sources: Distribution of LLM-related Scientific Publications by Journal

Sources	Articles
Journal of Medical Internet Research	18
Cureus Journal of Medical Science	14
JMIR Medical Education	13
IEEE Access	5
JMIR Mental Health	5
Plos One	5
JMIR Medical Informatics	4
Proceedings of the ACM on Interactive Mobile Wearable and Ubiquitous Technologies-IMWUT	4
AI	3
JAMA Network Open	3

JMIR: Journal of Medical Internet Research, **IEEE:** Institute of Electrical and Electronics Engineers, **JAMA:** Journal of the American Medical Association

Table 2: Most Relevant Institutions

Affiliation	Articles
Harvard University	16
University of California System	15
Feinberg School of Medicine	12
Northwestern University	12
University of Ulsan	12
Vanderbilt University	12
Sichuan University	10
Chongqing Medical University	9
Harvard University Medical Affiliates	9
Harvard Medical School	8

ed publications produced by each within the analyzed WoS dataset. More specifically, it shows the top 10 affiliations, ordered in descending number of articles. The column "Affiliation" displays the institution name as indexed in WoS, while the column "Articles" reports the number of publications from our final corpus associated with that institution. Table 2 reveals which academic institutions conduct more research in this field. Harvard University stands out as the academic institution that produces the most publications with 16 publications. The University of California System ranks second with 15 publications. The leadership of these two institutions reflects the leading role of the US in scientific productivity. Other leading universities, Feinberg School of Medicine, Northwestern University, University of Ulsan and Vanderbilt University, have a remarkable scientific production with 12 publications. These institutions are particularly active in the fields of medicine, health informatics and artificial intelligence applications. China-based universities such as Sichuan University (10 publications) and Chongqing Medical University (9 publications) demonstrate the strength of academic research in Asia. This finding reflects China's growing scientific activity in the fields of medical informatics and artificial intelligence. Harvard University Medical Affiliates (9 publications) and Harvard Medical School (8 publications) show the scope of Harvard University's research in medicine and health informatics. These results are important to understand which institutions are driving scientific research and to identify academic centers for future collaboration.

Most Relevant Words

Findings on the most relevant words that are prominent in academic publications are given in Table 3. In Table 3, the "Words" column lists the author keywords, while the "Occurrences" column shows how many times each keyword appears in the dataset. Table 3 provides an understanding of the trends in a particular field by revealing which topics are emphasized more in academic studies. "Artificial Intelligence" stands out as the most frequently used keyword with 16 repetitions. "ChatGPT" ranks second with 10 occurrences. This finding shows that AI and ChatGPT are increasingly being examined in academic research and the use of these technologies in various fields is being investigated. The word "Education" was used 8 times and the word "Performance" was used 7 times. This shows the growing academic interest in how AI is being used, especially in the context of education.

Table 3: Most Relevant Keywords

Words	Occurrences
artificial-intelligence	16
chatgpt	10
education	8
performance	7
challenges	5
accuracy	4
AI	4
depression	4
health	4
impact	4

AI: Artificial intelligence

Tree Map

The tree map visualizing the most frequently used keywords in scientific studies and their importance in the academic literature is given in Figure 2. The size of each box reflects the number of repetitions of the relevant word and shows which themes are at the forefront of scientific research. Each rectangle represents one author keyword, and the different colors are used only to visually distinguish the keywords, while the numbers and percentages inside each box indicate the absolute frequency and relative share of that keyword in the dataset. According to Figure 2, "Artificial Intelligence" is the keyword that ranks first with 82 repetitions (19%). "ChatGPT" ranked second with 63 occurrences (14%). These findings show that artificial intelligence and large language models are becoming an increasingly central topic in scientific research. "Large Language Models" has an important place with 48 occurrences (11%). The terms "LLM" and "Large Language Model" were used 33 times in total. The term "Generative AI" was used 11 times. The term "medical education" was used 32 times (7%). "Chatbots" is included in the literature as one of the applications of large language models with 8 repetitions.

Thematic Map

The thematic map presented in Figure 3 visualizes the most common themes and their importance levels in the scientific literature through bibliometric analysis. Figure 3 is divided into four sections where keywords are classified in terms of their density and centrality. On the horizontal axis, the "Relevance degree (Centrality)" indicates how strongly a theme is connected to other themes in the overall network, while on the vertical axis the "Development degree (Density)" reflects the internal cohesion and maturity of each theme. The size of the circles represents the frequency of the keywords in the dataset, and the different colors indicate distinct thematic clusters. These are Motor Themes (Upper Right Quadrant), Core Themes (Lower Right Quadrant), Niche Themes (Upper Left Quadrant) and Emerging or Weakening Themes (Lower Left Quadrant).

In the motor themes region, the most powerful and guiding themes in the scientific field with high centrality and

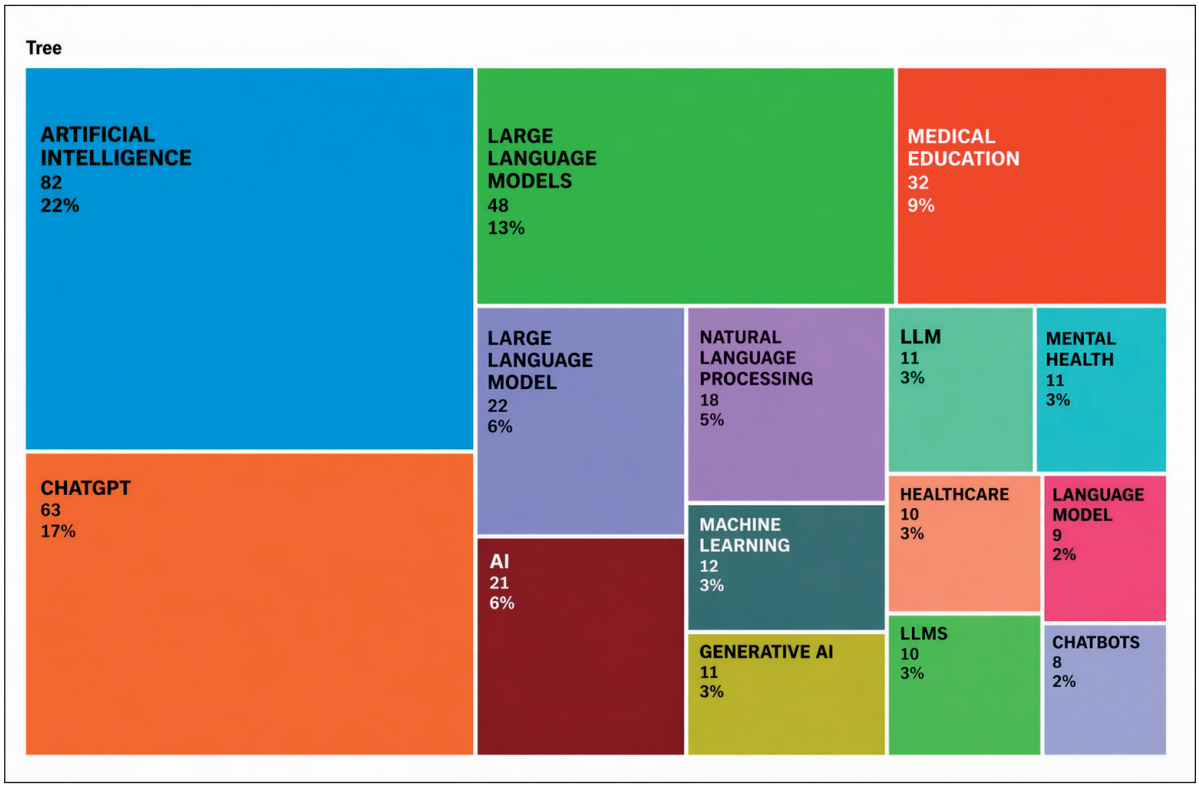


Figure 2: Treemap of the most frequently used author keywords in LLM-related healthcare publications

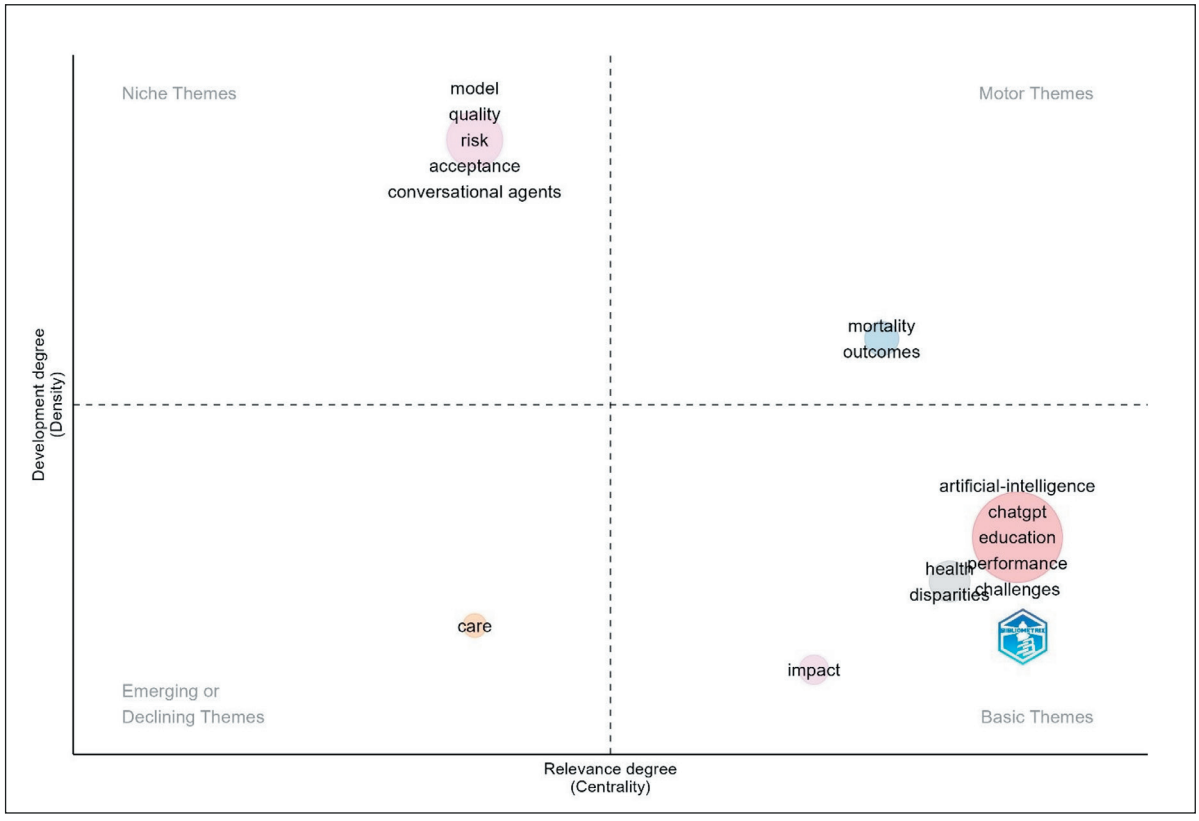


Figure 3: Thematic Map

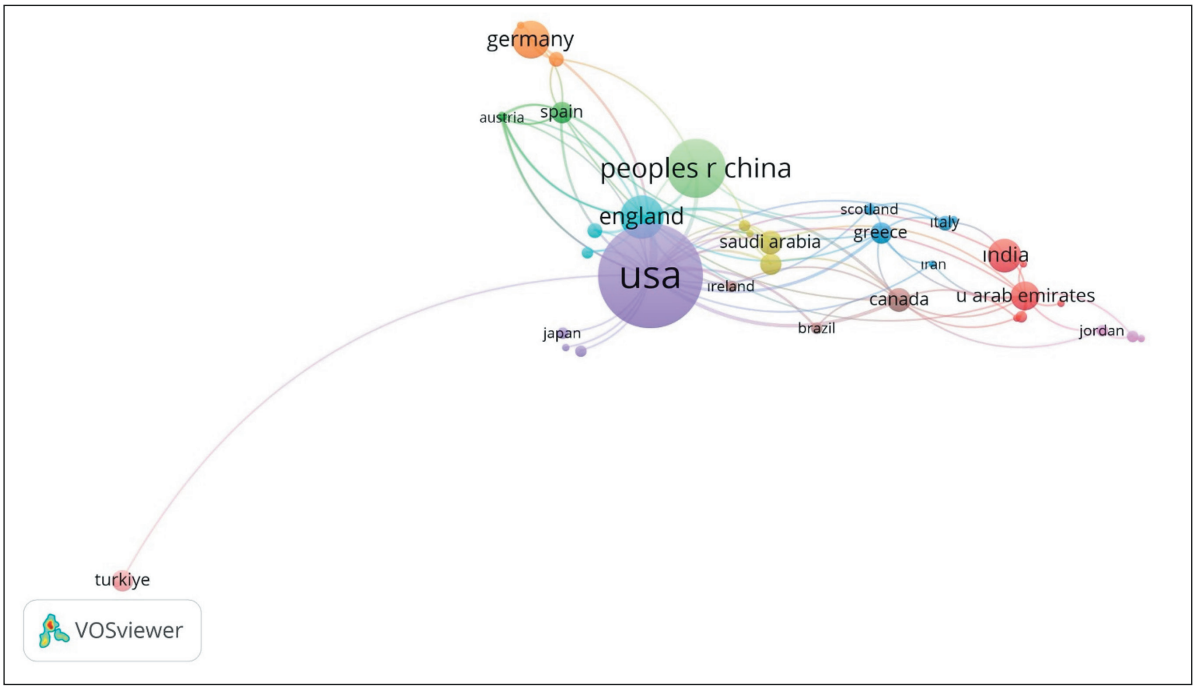


Figure 4: Academic Collaboration Network

sophistication in the literature are located. "mortality" and "outcomes" are located in this region, but do not form a large cluster. This shows that mortality and clinical outcomes remain an important area in health and medical research, but topics such as AI and ChatGPT are becoming more central.

Core Themes include topics that have been widely studied and are central to the literature but are not yet fully developed. "artificial intelligence", "chatgpt", "education", "performance", "health disparities", "challenges" and "impact" are included in this section. This finding shows that artificial intelligence and chatgpt have become a core academic topic, especially in the context of education and health. Issues such as health inequalities and performance are indicative of the growing interest in how AI-enabled systems are being used in health.

Niche Themes include topics that are more specific and intensively studied but not central to the literature. "model", "quality", "risk", "acceptance" and "conversational agents" fall into this category. This shows that academic work on conversational agents (e.g. chatGPT) and the reliability, risks and acceptance of language models is intensifying. However, these topics have not yet become a large-scale central theme in the literature.

Emerging or weakening themes include the concept of "care". This suggests that work on care services of AI and ChatGPT is still developing or has not made a strong impact in the literature.

Academic Collaboration Network

In order to analyze scientific collaborations between countries, a bibliometric map derived from co-authorship data is presented in Figure 4. This map was created using VOSviewer software, collaborative relationships between countries are represented by nodes and edges. Node sizes vary depending on the number of co-authorships of the respective country, with larger nodes indicating more collaboration. In this visualization, the thickness of the connecting lines ("edges") reflects the strength of collaboration between two countries (number of co-authored documents), and different colors indicate distinct clusters of closely collaborating countries. Figure 4 shows that the node with the largest and most densely connected network is USA. Accordingly, it can be said that the USA has a central position in international scientific collaborations and collaborates with a large number of countries. Countries such as the UK, China and Germany also stand out as important nodes in the network. These countries have a wide network of cooperation and are connected to a large number of countries. In Europe, there are strong links between the UK, Germany and Spain. Countries such as India and the United Arab Emirates have developed closer ties with specific regions. Turkey is positioned far from the network and has relatively few connections in global collaborations. As noted above, the different colors correspond to clusters of countries that are more closely connected to each other than to countries

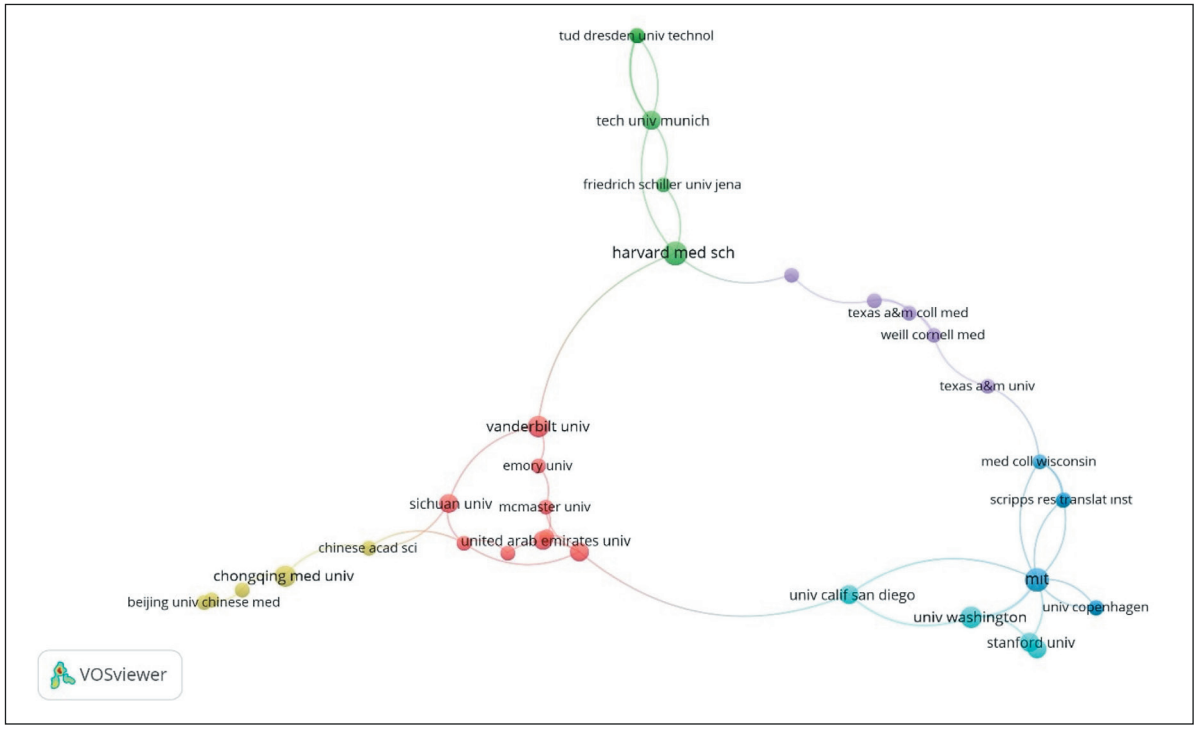


Figure 5: Cooperation Network between Academic Institutions

in other clusters. The blue-green network, which includes the US and the UK, shows strong academic collaborations between developed countries, while the red-pink network, which includes China and India, presents a more regional model of collaboration. This analysis reveals that the US has a central position in scientific research, while the UK, Germany and China stand out as important cooperation actors.

Co-authorship Analysis

As a result of the co-authorship analysis conducted in the study, Figure 5 presents a map of the collaborations between universities and research centers. Each circle (node) represents a university or research institution, the size of the node is proportional to the number of co-authored documents, the connecting lines (edges) indicate co-authorship links, and different colors denote clusters of institutions that collaborate more frequently with each other. Accordingly, one of the most important nodes is Harvard Medical School, shown in green, which is connected to various academic institutions from different geographies. This shows that Harvard has a wide network of academic collaborations at the international level. In particular, it is directly linked to three major universities in Germany (TU Dresden, Technical University of Munich and Friedrich Schiller University Jena), indicating that it also attaches importance to collaborations in Europe. Leading US universities such as the Massachusetts Institute of Technology (MIT), the University of Washington and Stanford University are notable for their strong collaboration. When regional clusters and academic collaboration are examined, China-based institutions (Sichuan University, Chongqing Medical University, Beijing University of Chinese Medicine and Chinese Academy of Sciences), which can also be called the Asian cluster and shown in yellow in the figure, form a strong collaboration network within themselves. This may reflect the influence of regional science policies and local research funding. The Middle East and Canada Collaboration is also shown in red in the figure, with a notable academic partnership between the United Arab Emirates University, McMaster University and Emory University. This cluster represents the academic ties between Canada, the US and the Middle East. Another noteworthy element in the figure is the relatively low number of direct links between China-based academic institutions and major research universities in the US. This suggests that regional trends in scientific collaborations are strong and that some academic barriers may exist globally.

DISCUSSION

This study reveals that academic productivity towards Large Language Models in healthcare has increased significantly in recent years and research trends are concentrated around specific themes. By focusing on WoS-indexed publications between 2023 and early 2025, our analysis quantitatively characterizes the “post-ChatGPT” period and shows how scientific output and collaboration are distributed across

countries, institutions and research topics. The results of the bibliometric analysis show that the United States, China and some European countries are leading the way in this field and that scientific collaboration networks are shaped around these countries. The leadership of institutions such as Harvard University, University of California System and Northwestern University at both the publication and collaboration level reveals that the central actors of AI research in healthcare are largely institutions with high research capacity (14,15). Compared with previous narrative and scoping reviews of LLMs in healthcare, the present study adds a data-driven mapping of where this research is produced and which institutions occupy structurally central positions in the global collaboration network.

The fact that the most frequently encountered keywords include “Artificial Intelligence”, ‘ChatGPT’, “Medical Education” and “Performance” shows that LLM technologies are widely used in health systems not only in decision support processes but also in areas such as health education, communication and service quality (1,16). Our treemap results further quantify this pattern by showing that “artificial intelligence” and “ChatGPT” together account for a substantial proportion of all keywords, confirming that LLMs have rapidly become a focal topic in health-related AI research. Tree map and thematic map analyses show that AI-based approaches also touch social dimensions such as “health disparities” and “challenges”. These findings indicate that LLM applications should be evaluated not only in technological but also in socio-ethical contexts (8,17). In this respect, our thematic structure, in which terms related to equity and challenges appear alongside performance-oriented keywords, empirically supports the growing emphasis on ethical and societal implications reported in recent literature. Similarly, a recent national bibliometric study from Turkey has highlighted ethical and governance-related themes in AI applications in healthcare, underscoring the importance of these issues in the local context as well (18).

According to the Thematic Map analysis, concepts such as “artificial intelligence” and “ChatGPT” appear as core themes, while terms such as ‘acceptance’, “conversational agents” and “risk” are more niche or emerging areas. This suggests that dimensions such as user perception, trust, and ethical awareness are still immature research areas in the adoption process of LLM technologies (19,20). Our finding that these topics cluster as niche or emerging themes indicates that, despite intense conceptual and opinion-based discussion in the literature, empirical and large-scale studies on acceptance, risk perception and real-world deployment of conversational agents in healthcare remain relatively limited. Many studies emphasize that the use of LLM, especially in clinical settings, should be evaluated within the framework of security, accuracy and privacy (21,22), and our keyword and thematic analyses complement these concerns by demonstrating that such issues are beginning to gain visibility but have not yet matured into firmly established core themes.

In the analysis of cooperation between countries, it is seen that countries such as the USA, the UK, China and Germany are pioneers in scientific productivity and have extensive cooperation networks. This shows that LLM-based health research is largely concentrated in specific scientific centers (10). It is important for developing countries such as Turkey to develop more international collaboration in order to utilize these technologies effectively and reliably in the local context. In our co-authorship map, Turkey appears as a small, weakly connected node located at the periphery of the global network, indicating limited participation in large, multicenter LLM projects. This peripheral position may reduce the visibility and impact of national research outputs and slow down the transfer of methodological innovations and best practices into the local health system. Strengthening collaboration with leading centers in North America, Europe and Asia would therefore be strategically important for capacity building, access to high-quality datasets and the development of context-appropriate LLM applications for the Turkish healthcare system.

The results of this study are in line with the findings in the literature on the role of LLMs in healthcare. For example, Liu et al. (1) state that LLMs are effective in diagnostic support, medical summarization, clinical question answering and patient interaction; however, more research is needed on issues such as ethics, transparency and model interpretability. Similarly, Bi et al. (14) state that LLMs can contribute to multimodal data integration in healthcare, but data standardization and reliability issues are still important challenges to be solved. Our mapping refines these observations by showing that performance-related terms (e.g. “performance”, “accuracy”, “outcomes”) are already central in the keyword structure, whereas issues such as data quality, standardization and explainability are present but less prominent, suggesting a gap between the technical potential of LLMs and the maturity of research on their safe and robust implementation.

It is also emphasized that generative models such as ChatGPT offer potential in curriculum development, simulation and assessment processes in medical education, but cannot replace direct human interaction with faculty members (19,20). It is suggested that such technologies should be considered as supportive and complementary tools in educational practices. Consistent with this, our results show that “medical education” and related terms form an important thematic area, indicating that educational applications are one of the most actively explored domains for LLMs in health sciences; however, other application areas such as long-term outcome evaluation and health service delivery

appear less developed in the thematic map, highlighting opportunities for future empirical research.

However, it is also stated that LLM-based systems have serious limitations such as "hallucination" (generating false information), misjudging the context, creating user bias and unethical guidance (8,22). It is clear that such systems should only be used in a supervised and transparent environment, especially in clinical applications. As a matter of fact, institutions such as the World Health Organization (WHO) also emphasize that artificial intelligence applications in health should be supervised within an ethical framework (23). The presence of keywords such as "challenges", "health disparities" and "risk" in our analysis is consistent with these concerns and suggests that the research community has started to incorporate ethical and governance issues into the LLM agenda; nevertheless, their positioning as non-motor themes indicates that systematic, outcome-oriented evaluations of ethical frameworks and governance models are still relatively scarce.

Conclusion

This study provides a comprehensive mapping of the use of LLMs in healthcare, revealing the direction of academic interest, strong themes and existing gaps. Our findings highlight three main contributions: (i) a detailed description of the global and institutional landscape of LLM-related health research in the immediate post-ChatGPT period, (ii) an identification of core and emerging themes that structure the current literature, and (iii) an evidence-based demonstration of collaboration asymmetries, particularly the peripheral position of some developing countries such as Turkey in international co-authorship networks. Multifaceted, interdisciplinary and ethics-based research approaches are needed to integrate LLM technologies into health systems. Future studies should therefore prioritize rigorous, data-driven evaluations of clinical effectiveness, safety and equity impacts of LLM applications, as well as strategies to broaden international collaboration and reduce geographic and thematic imbalances identified in this bibliometric analysis.

Limitations

This study has several limitations that should be considered when interpreting the findings. First, the analysis includes only Web of Science-indexed, open-access, English-language articles, which may underrepresent studies from other databases, subscription journals or non-English sources. Second, the time window is restricted to 2023-22 February 2025, so publication counts for 2025 are provisional and later-indexed articles are not captured. Third, the search strategy relied on a predefined set of LLM-related keywords (e.g. "Large Language Model", "LLM", "ChatGPT", "Generative "), meaning that studies using different terminology may have been missed.

Author Contributions

Study conception and design: **Hakan Öztürk, Elvan Hayat;** data collection: **Hakan Öztürk, Elvan Hayat;** analysis and interpretation of results: **Hakan Öztürk, Elvan Hayat;** draft manuscript preparation: **Hakan Öztürk, Elvan Hayat.** The authors reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Financial Support

The authors declare that the study received no funding.

Ethical Approval

Ethical approval was not required for this study as it is a bibliometric analysis based on publicly available data.

REFERENCES

1. Liu J, Yang M, Yu Y, Huang X, Zhang Q. Large language models in bioinformatics: applications and perspectives. arXiv. 2024;2401.04155.
2. Adedeji A, Sanni M, Ayodele E, Okoro M, Bello A. The multicultural medical assistant: can LLMs improve medical ASR errors across borders? arXiv. 2025;2501.15310. https://doi.org/10.48550/arXiv.2501.15310
3. Wen-Ching Wu MSN. Effects of a Steno Diabetes Dialogue Card-based intervention on self-perceived health, health literacy, and glycemic control in older adults with type 2 diabetes mellitus living in the community. Hu Li Za Zhi. 2025;72(1):64-75. https://doi.org/10.6224/JN.202502_72(1).09
4. Karabacak M, Margetis K. Embracing large language models for medical applications: opportunities and challenges. Cureus. 2023;15(12):e39305. https://doi.org/10.7759/cureus.39305
5. Vrdoljak J, Boban Z, Vilović M, Kraljević T, Petrović L. A review of large language models in medical education, clinical decision support, and healthcare administration. Healthc (Basel). 2025;13(6):603. https://doi.org/10.3390/healthcare13060603
6. Alber DA, Yang Z, Alyakin A, Brown P, Chen L. Medical large language models are vulnerable to data-poisoning attacks. Nat Med. 2025;31:1-9. https://doi.org/10.1038/s41591-024-03445-1
7. Tangsrivimol JA, Darzidehkalani E, Virk HU, Smith J, Lee V. Benefits, limits, and risks of ChatGPT in medicine. Front Artif Intell. 2025;8:1518049. https://doi.org/10.3389/frai.2025.1518049
8. Clusmann J, Kolbinger FR, Muti HS, O'Connor C, Peters G. The future landscape of large language models in medicine. Commun Med. 2023;3(1):141. https://doi.org/10.1038/s43856-023-00370-1
9. Kumar K, Ashraf T, Thawakar O, Anwer RM, Singh P. LLM post-training: a deep dive into reasoning large language models. arXiv. 2025;2502.21321. https://doi.org/10.48550/arXiv.2502.21321
10. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res. 2021;133:285-296. https://doi.org/10.1016/j.jbusres.2021.04.070

11. Aria M, Cuccurullo C. Bibliometrix: an R-tool for comprehensive science mapping analysis. J Informetr. 2017;11(4):959-75. https://doi.org/10.1016/j.joi.2017.08.007
12. Barrington NM, Gupta N, Musmar B, Doyle D, Panico N, Godbole N, et al. A bibliometric analysis of the rise of ChatGPT in medical research. Med Sci (Basel). 2023;11(3):61. https://doi.org/10.3390/medsci11030061
13. Koo M. ChatGPT research: a bibliometric analysis based on the Web of Science from 2023 to June 2024. Knowledge. 2025;5(1):4. https://doi.org/10.3390/knowledge5010004
14. Bi Z, Dip SA, Hajialigol D, Kommu S, Liu H, Meng L, Patel R. AI for biomedicine in the era of large language models. arXiv. 2024;2403.15673. https://doi.org/10.48550/arXiv.2403.15673
15. Gencer G, Gencer K. Large language models in healthcare: a bibliometric analysis and examination of research trends. J Multidiscip Healthc. 2025;18:223-238. https://doi.org/10.2147/JMDH.S502351
16. Cascella M, Semeraro F, Montomoli J, Rossi A, Bianchi F. The breakthrough of large language models release for medical applications: 1-year timeline and perspectives. J Med Syst. 2024;48(1):22. https://doi.org/10.1007/s10916-024-02045-3
17. Lawrence HR, Schneider RA, Rubin SB, Davis K, Williams L. The opportunities and risks of large language models in mental health. JMIR Ment Health. 2024;11(1):e59479. https://doi.org/10.2196/59479
18. Çelik Ö, Kaya E. Artificial intelligence and ethics in healthcare: a bibliometric analysis. Süleyman Demirel University Visionary Journal. 2024;15(43):1046-1062. https://doi.org/10.21076/vizyoner.1455659
19. Abd-Alrazaq A, AlSaad R, Alhuwail D, Househ M, Shah Z. Large language models in medical education: opportunities, challenges, and future directions. JMIR Med Educ. 2023;9(1):e48291. https://doi.org/10.2196/48291
20. Khan AA, Akbar MA, Fahmideh M, Rahman N, Li C. AI ethics: an empirical study on the views of practitioners and lawmakers. IEEE Trans Comput Soc Syst. 2023;10(6):2971-2984. https://doi.org/10.1109/TCSS.2023.3251729
21. Yu D, Bondi M, Hyland K. Can GPT-4 learn to analyse moves in research article abstracts? Appl Linguist. 2024;amae071. https://doi.org/10.1093/applin/amae071
22. Zhou Z, Li J, Zhang Z, Chen F, Wu X. Examining how the large language models impact the conceptual design with human designers: a comparative case study. Int J Hum-Comput Interact. 2024;1-17. https://doi.org/10.1080/10447318.2024.2370635
23. World Health Organization. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. Geneva: WHO; 2024. ISBN: 978-92-4-008475-9

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):81-88
DOI: 10.29058/mjwbs.1783501

Clinical characteristics and hematological inflammatory markers in pediatric peripheral facial paralysis: A retrospective study

Pediatrik periferik fasiyal paralizide klinik özellikler ve hematolojik inflamatuvar belirteçler: Retrospektif bir çalışma

Meral Karadağ[1], **Melek Yaman Ortaköylü**[2]

[1]Department of Pediatric Neurology, Malatya Training and Research Hospital, Malatya, Türkiye
[2]Department of Pediatric Hematology and Oncology, Malatya Training and Research Hospital, Malatya, Türkiye

Cite this article as: *Karadağ M and Yaman Ortaköylü M. Clinical characteristics and hematological inflammatory markers in pediatric peripheral facial paralysis: a retrospective study. Med J West Black Sea. 2026;10(1):81-88.*

The study was presented as an oral presentation at the 5th International Eastern Pediatrics Congress.

ABSTRACT

Aim: This study aimed to determine whether inflammatory indices derived from complete blood count indicate systemic inflammation at presentation in pediatric idiopathic Bell's palsy.

Material and Methods: We retrospectively analyzed 23 children with Bell's palsy and 69 age- and sex-matched healthy controls (March 2022–June 2025). Peripheral blood was drawn at presentation and before corticosteroid therapy. Indices included neutrophil/lymphocyte ratio, platelet/lymphocyte ratio, systemic immune-inflammation index, systemic inflammation response index, lymphocyte/monocyte ratio, eosinophil/lymphocyte ratio, derived neutrophil/lymphocyte ratio, and red cell distribution width/platelet ratio, mean platelet volume, platelet distribution width, plateletcrit were analyzed comparatively.

Results: Of the patients, 12 (52.2%) were female and 11 (47.8%) were male, with a mean age of 12.1 ± 3.8 years (range: 2–17). Based on the House–Brackmann classification, 17 (73.9%) patients were grade 2, 5 (21.7%) were grade 3, and 1 (4.4%) was grade 4. No statistically significant differences were found between the patient and control groups in terms of neutrophil/lymphocyte ratio, platelet/lymphocyte ratio, systemic immune-inflammation index, systemic inflammation response index, lymphocyte/monocyte ratio, eosinophil/lymphocyte ratio, derived neutrophil/lymphocyte ratio, plateletcrit, and red cell distribution width/platelet ratio values (for all variables $p > 0.05$). Mean platelet volume and platelet distribution width were lower in patients (Cliff's $\delta = -0.34$ for both; 95% CI –0.58 to –0.10, $p = 0.015$ for each). House–Brackmann grade correlated with plateletcrit (Spearman's $\rho = 0.58$; 95% CI 0.24 to 0.79; $p = 0.004$).

Conclusion: Modest reductions in mean platelet volume and platelet distribution width, along with a positive correlation between plateletcrit and clinical severity, may indicate platelet-axis involvement. Given the limited sample size and statistical power, these findings require confirmation in larger, standardized cohorts.

Keywords: Hematologic markers, inflammatory markers, pediatric bell's palsy

ÖZ

Amaç: Çocukluk çağı idiyopatik Bell paralizisinde başvuru anındaki tam kan sayımı kaynaklı inflamatuvar indekslerin sistemik inflamasyonu gösterip göstermediğini belirlemek amaçlandı.

Gereç ve Yöntemler: Mart 2022–Haziran 2025 tarihleri arasında üçüncü basamak pediatrik nöroloji kliniğinde izlenen Bell paralizili 23 çocuk ile yaş ve cinsiyete göre eşleştirilmiş 69 sağlıklı kontrol geriye dönük olarak analiz edildi. Periferik kan örnekleri başvuruda ve kortikosteroid tedavisi öncesinde alındı. İncelenen indeksler: nötrofil/lenfosit oranı, trombosit/lenfosit oranı, sistemik immün-inflamasyon indeksi, sistemik inflamasyon yanıt indeksi, lenfosit/monosit oranı, eozinofil/lenfosit oranı, türetilmiş nötrofil/lenfosit oranı, eritrosit dağılım genişliği/trombosit oranı, ortalama trombosit hacmi, trombosit dağılım genişliği ve plateletkrit; karşılaştırmalı analiz yapıldı. İstatistiksel anlamlılık için $p<0,05$ kabul edildi.

Bulgular: Hastaların 12'si (%52,2) kız, 11'i (%47,8) erkek; ortalama yaş 12,1±3,8 yıl (2–17). House–Brackmann sınıflamasına göre 17'si (%73,9) derece 2, 5'i (%21,7) derece 3, 1'i (%4,4) derece 4'tü. Nötrofil/lenfosit oranı, trombosit/lenfosit oranı, sistemik immün-inflamasyon indeksi, sistemik inflamasyon yanıt indeksi, lenfosit/monosit oranı, eozinofil/lenfosit oranı, türetilmiş nötrofil/lenfosit oranı, plateletkrit ve eritrosit dağılım genişliği/trombosit oranı açısından hasta ve kontrol grupları arasında anlamlı fark saptanmadı (tüm değişkenler için $p > 0,05$). Ortalama trombosit hacmi ve trombosit dağılım genişliği hasta grubunda daha düşüktü (Cliff's $\delta = -0.34$; %95 güven aralığı –0.58 ile –0.10; her iki değişken için $p = 0,015$). House–Brackmann derecesi ile plateletkrit arasında pozitif korelasyon saptandı (Spearman's $\rho = 0,58$; 95% güven aralığı 0.24–0.79; $p = 0,004$).

Sonuç: Ortalama trombosit hacmi ve trombosit dağılım genişliğindeki ılımlı düşüşler ile plateletkrit ve House–Brackmann derecesi arasındaki pozitif korelasyon, trombosit ekseninin olası etkilenimini düşündürmektedir. Ancak sınırlı örneklem nedeniyle sonuçlar, daha geniş ve metodolojik olarak standart kohortlarda teyit edilmelidir.

Anahtar Kelimeler: Hematolojik belirteçler, inflamatuvar belirteçler, pediatrik bell paralizisi

Corresponding Author: Melek Yaman Ortaköylü ✉ yaman.melek@gmail.com **Received:** 13.09.2025 **Revision:** 25.10.2025 **Accepted:** 24.12.2025

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- This retrospective study evaluated a broad panel of complete blood count–derived inflammatory indices in children with idiopathic peripheral facial paralysis compared with healthy controls.
- Modest reductions in MPV (Mean Platelet Volume) and PDW (Platelet Distribution Width), together with a positive correlation between PCT (Plateletcrit) and disease severity, suggest platelet-axis involvement in pediatric BP (Bell's palsy).
- As routinely available parameters, these indices may help support early clinical assessment in pediatric BP across diverse clinical settings.

INTRODUCTION

Peripheral facial paralysis is one of the most common motor cranial neuropathies. A lesion occurring at any point along the course of the facial nerve—the seventh cranial nerve—extending from its motor nucleus in the brainstem to the muscles of facial expression, may result in peripheral facial paralysis (1). Idiopathic peripheral facial paralysis, also known as Bell's palsy (BP), is characterized by sudden-onset unilateral facial muscle weakness in the absence of identifiable causes such as other cranial neuropathies, neoplasms, or infections. Clinical manifestations typically include the inability to close the eyelid on the affected side, drooping of the corner of the mouth, alterations in taste sensation, and hypersensitivity to sound (2).

Although not fully elucidated, BP is most attributed to edema that confines the facial nerve in the fixed bony canal, resulting in compression. Theories involving infection/inflammation and vascular ischemia support this mechanism (3). While the condition is most often idiopathic, viral infections—particularly herpes simplex virus—along with immune and inflammatory mechanisms have been implicated in its pathogenesis (4). The annual incidence in children has been reported as 2–4 per 100,000 (5). A favorable prognosis is expected, with approximately 70–90% of patients with BP recovering completely without treatment (6).

Hematological inflammatory parameters are biomarkers derived from complete blood counts (CBC) and serve as indirect indicators of systemic inflammation. In particular, the neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and lymphocyte-to- monocyte ratio (LMR) is widely used to assess inflammatory responses in various chronic diseases, infections, and malignancies (7). More recently, newly derived indices such as the Systemic Immune-Inflammation Index (SII) and the Systemic Inflammation Response Index (SIRI) have gained importance as prognostic markers (8). These parameters are commonly preferred in clinical practice due to their accessibility, low cost, and reliability (9).

The favorable response of BP to corticosteroids, along with contrast enhancement on MRI (magnetic resonance imaging), supports an inflammatory component in its pathogenesis. As a reliable inflammatory marker, NLR has been reported to increase during inflammatory processes in adult patients with BP. Mean platelet volume (MPV) and PLR are additional parameters associated with vascular diseases and inflammatory conditions. However, evidence in children is sparse, with only a few studies examining these hematologic markers in BP (3, 5, 10-12).

This study aimed to describe the clinical characteristics of pediatric BP and to examine potential associations between pathophysiology and selected CBC-derived inflammatory indices. By comparing these markers between patients and healthy controls, the study sought to assess whether such indices—derived from routine blood tests—might offer supportive evidence of systemic inflammation during the acute phase of the condition.

MATERIAL and METHODS

This retrospective cross-sectional study included 23 pediatric patients aged ≤ 18 years with a first episode of idiopathic BP, diagnosed at a tertiary pediatric neurology clinic between March 2022 and June 2025. Eligible patients were required to have complete clinical records and CBC parameters obtained at the time of presentation. Exclusion criteria included recurrent facial palsy, secondary causes of facial paralysis, acute infection, chronic systemic disease, use of immunomodulatory or anti-inflammatory medication prior to blood sampling, and incomplete clinical or laboratory data. Of all children presenting with facial paralysis, three were diagnosed with congenital facial palsy, two had facial nerve–involving tumors, and one had traumatic facial nerve injury; these cases were excluded from the analysis (Figure 1). The control group consisted of 69 healthy children, selected consecutively from routine outpatient visits at a 3:1 control-to-case ratio to improve statistical power due to the relative rarity of pediatric BP. Controls were matched to cases at the group level for age and sex ($p > 0.05$ for both) and were required to have no history of neurologic, infectious, inflammatory, or chronic disease, no medication use or infection within the preceding two weeks, and available CBC results. Demographic and clinical data, including age, sex, side of facial involvement, comorbidities, time to presentation, and

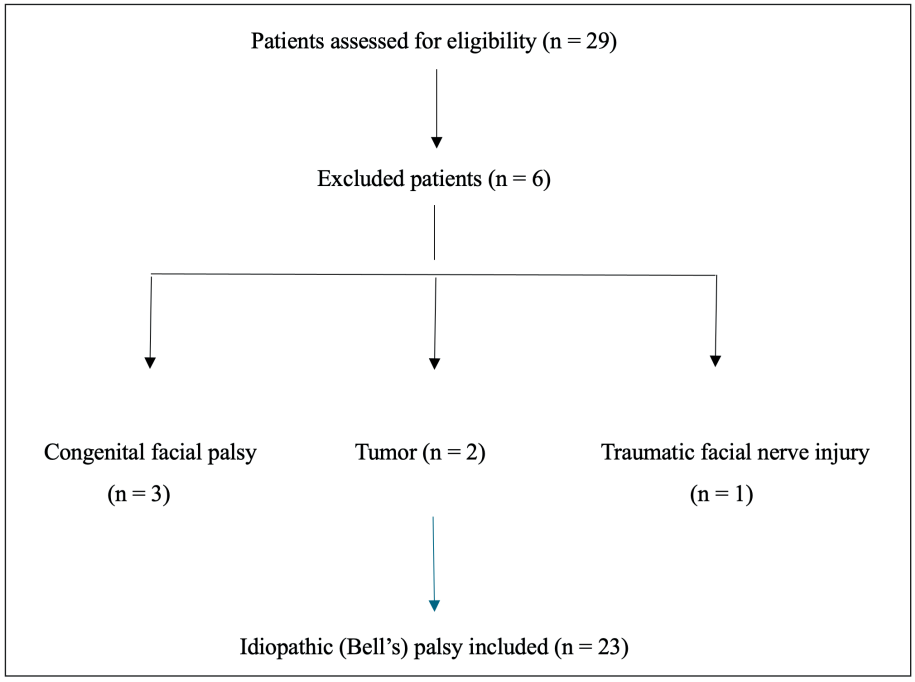


Figure 1: Flow diagram of patient selection

treatments administered, were retrieved retrospectively from medical records. Patients were classified according to the House–Brackmann (HB) grading system. The HB grading system classifies facial function into six grades, ranging from normal function (Grade I) to complete paralysis (Grade VI), based on standardized evaluation of resting symmetry, voluntary movement, and synkinesis. Grade I reflects normal facial function; Grade II indicates mild dysfunction with slight asymmetry; Grade III represents moderate dysfunction with obvious but non-disfiguring weakness; Grade IV corresponds to moderately severe dysfunction with incomplete eye closure; Grade V reflects severe dysfunction with barely perceptible movement; and Grade VI denotes total facial paralysis (13).

All blood samples were obtained at the initial presentation for facial palsy and before initiation of corticosteroid therapy. All CBC analyses were performed using the same automated hematology analyzer, Sysmex XN-1000, in the hospital's central laboratory to ensure consistency. From the CBC results, the following laboratory parameters were obtained: total leukocyte (×10³/μL), neutrophil (×10³/μL), lymphocyte (×10³/μL), monocyte (×10³/μL), eosinophil (×10³/μL), platelet (×10³/μL), MPV (fL), platelet distribution width (PDW) (%), red cell distribution width (RDW) (%), and plateletcrit (PCT) (%). Several derived inflammatory indices were calculated as follows: SII was calculated by multiplying the neutrophil and platelet counts and dividing the result by the lymphocyte count (neutrophil × platelet) / lymphocyte; SIRI was obtained by multiplying the neutrophil and monocyte counts and dividing by the lymphocyte count (neutrophil × monocyte)/ lymphocyte; LMR was calculated as lymphocyte count divided by monocyte count; eosinophil-to-lymphocyte ratio (ELR) as eosinophil count divided by lymphocyte count; derived NLR (d-NLR) as neutrophil count divided by the difference between total white blood cell count and neutrophil count; and RDW-to-platelet ratio (RPR) as RDW (%) divided by platelet count (10³/μL).

Data were evaluated using IBM SPSS Statistics, Version 25. Normality checks used the Kolmogorov–Smirnov and Shapiro–Wilk procedures. Skewed variables appear as median (min–max); normally distributed variables as mean ± SD. To interpret non-significant findings more accurately, we conducted a sensitivity analysis to estimate the minimal detectable effect for our sample size (n=23 vs. 69; α=0.05; 80% power). For two-group analyses, independent-samples t-tests were applied when assumptions held; otherwise, Mann–Whitney U tests were used. Categorical outcomes were analyzed with Pearson's chi-square or Fisher's exact tests. Significance threshold: $p<0.05$. Between-group contrasts were summarized with standardized effect sizes and 95% confidence intervals, reported as Cliff's delta (with a probability-of-superiority interpretation). All effect sizes are reported as Cliff's delta. For outcomes analyzed with independent-samples *t* tests, δ was derived from Hedges' *g* using $\delta = 2 \cdot \Phi(g/\sqrt{2}) - 1$, with confidence intervals transformed accordingly. Associations between HB grade (ordinal) and hematologic indices were summarized with Spearman's ρ and 95% CIs (Fisher's z method).

The study protocol was reviewed and approved by the Malatya Turgut Özal University Health Sciences Scientific Research Ethics Committee (Approval No: 2025/250, Date: August 5, 2025), and all procedures were conducted in accordance with the ethical principles outlined in the Declaration of Helsinki.

RESULTS

A total of 23 children were included in the final analysis. Among these, 12 were female (52.2%) and 11 were male (47.8%). Facial involvement was observed on the left side in 12 patients (52.2%) and on the right side in 11 patients (47.8%). The age range of the patients was between 2 and 17 years, with a mean age of 12.1 ± 3.8 years. Two patients (8.7%) had comorbid conditions, epilepsy (n=1) and scaphocephaly (n=1). The control group consisted of 69 healthy children (33 males, 36 females) with similar age and sex characteristics (Table 1).

When diagnoses were evaluated by year, five patients were identified in 2022, three in 2023, ten in 2024, and five in 2025. The highest number of cases by month was observed in February (4 patients) and September (3 patients). There was no statistically significant difference in the monthly (p=0.947) or seasonal (p=0.924) distribution of cases.

According to the HB classification, 17 patients (73.9%) were classified as grade 2, 5 patients (21.7%) as grade 3, and 1 patient (4.4%) as grade 4. The mean duration from symptom onset to clinical presentation was 3.6 days, ranging from 1 to 10 days. Steroid therapy was administered to 16 patients (69.5%). Central nervous system imaging was performed in 13 patients (56.5%); results were normal in 11, and incidental arachnoid cysts were detected in 2.

No statistically significant differences were found between the patient and control groups in terms of NLR, PLR, SII, SIRI, LMR, eosinophil count, ELR, RDW, derived NLR, PCT, and RDW/PLT values ($p > 0.05$). MPV and PDW were lower in patients (Cliff's $\delta = -0.34$ for both; 95% CI −0.58 to −0.10; $p = 0.015$) (Table 2). Values of hematologic indices across HB grades and their correlations with HB grade are shown in detail (Table 3). Among all markers, PCT showed a moderate, statistically significant correlation with clinical severity (Spearman's $\rho = 0.58$; 95% CI 0.24–0.79; p =

Table 1: Demographic characteristics of the patient and control groups

	Patient Group (n=23)	Control Group (n=69)	p-value
Sex (Female/Male)	12/11	36/33	1.000[a]
Mean age ±SD (years)	12.1 ± 3.8	12.1 ± 4.01	0.935[b]
Median age (min-max) (years)	12.5 (2.0-17.0)	12.5 (1.8-17.5)	

[a]Pearson Chi-square, [b]Mann-Whitney U
(n: number, max: maximum, min: minimum, SD: standard deviation)

Table 2: Comparison of Hematological Inflammatory Markers Between the Patient and Control Groups

	Patient Group	Control Group	p-value	Effect size (Cliff's δ; 95% Confidence Interval)
NLR	1.39 (0.37-5.91)	1.38 (0.26-6.03)	0.531[a]	δ = 0.09 (95% CI −0.19 to 0.36)
PLR	0.113 ± 0.042	0.110 (0.048-0.410)	0.763[b]	δ = −0.04 (95% CI −0.31 to 0.23)
ELR	0.034 (0.00-0.20)	0.053 (0.01-0.36)	0.104[a]	δ = −0.23 (95% CI −0.48 to 0.03)
LMR	5.59 ± 2.79	4.76 (1.14-12.01)	0.675[b]	δ = 0.06 (95% CI −0.22 to 0.34)
SII	465.9 (61.8-2224)	405 (84.1-2044.1)	0.384[a]	δ = 0.12 (95% CI −0.16 to 0.40)
SIRI	850 (214.8-5275.5)	780 (162.4-7568.8)	0.567[a]	δ = 0.08 (95% CI −0.20 to 0.36)
MPV	9.53 ± 1.00	10 (8.2-12.4)	0.015[a]*	δ = −0.34 (95% CI −0.58 to −0.10)
PCT	0.291 ± 0.085	0.315 ± 0.062	0.146[b]	δ = −0.20 (95% CI −0.44 to 0.07)
PDW	9.8 (8.3-14.4)	11.4 (8.0-16.9)	0.015[a]*	δ = −0.34 (95% CI −0.58 to −0.10)
d-NLR	1.034 (0.33-4.06)	1.101 (0.23-3.42)	0.404[a]	δ = 0.12 (95% CI −0.16 to 0.39)
RDW/PLT	0.043 (0.02-0.10)	0.041 (0.03-0.09)	0.925[a]	δ = −0.01 (95% CI −0.29 to 0.26)

[a]Mann Whitney U, [b]Independent Samples t-test

d-NLR: derived Neutrophil-to-Lymphocyte Ratio, **ELR:** Eosinophil-to-Lymphocyte Ratio, **LMR:** Lymphocyte-to-Monocyte Ratio, **MPV:** Mean Platelet Volume, **NLR:** Neutrophil-to-Lymphocyte Ratio, **PCT:** Plateletcrit, **PDW:** Platelet Distribution Width, **PLR:** Platelet-to-Lymphocyte Ratio, **PLT:** Platelet, **RDW:** Red Cell Distribution Width, **SII:** Systemic Immune-Inflammation Index, **SIRI:** Systemic Inflammatory Response Index (Data are presented as mean ± standard deviation or median (min–max) according to the distribution)

(Study sensitivity: with n=23 vs 69, minimal detectable effect at 80% power ≈ d=0.67 (Hedges' g) ≈ δ=0.37 (Cliff's delta), i.e., AUC≈0.685.)

Table 3: Correlations between House–Brackmann Grade and Hematologic Indices

	HB Grade 2 (n=17)	HB Grade 3 (n=5)	HB Grade 4 (n=1)	p-value	Effect size vs HB grade (Spearman ρ; 95% CI)
NLR	1.52 (0.37-3.95)	1.31 (0.81-5.91)	5.17	0.665[a]	ρ = 0.10 (95% CI –0.42 to 0.56)
PLR	0.105 ± 0.035	0.124 ± 0.051	0.200	0.247[a]	ρ = 0.25 (95% CI –0.20 to 0.62)
ELR	0.034 (0.010-0.201)	0.070 ± 0.074	0.005	0.477[a]	ρ = –0.16 (95% CI –0.53 to 0.43)
LMR	5.80 ± 2.89	5.56 ± 2.48	2.11	0.574[a]	ρ = –0.12 (95% CI –0.53 to 0.31)
SII	488 ± 264	400 (241-2124)	2224.0	0.364[a]	ρ = 0.20 (95% CI –0.34 to 0.64)
SIRI	825 (214-3368)	1043 (252-4557)	5275.5	0.364[a]	ρ = 0.20 (95% CI –0.34 to 0.64)
MPV	9.30 ± 0.92	10.14 ± 1.15	10.4	0.089[a]	ρ = 0.36 (95% CI –0.14 to 0.69)
PCT	0.26 ± 0.05	0.36 ± 0.09	0.45	0.004[a]*	ρ = 0.58 (95% CI 0.24 to 0.79)
PDW	9.7 (8.3-14.4)	10.98 ± 2.29	12.0	0.341[a]	ρ = 0.21 (95% CI –0.54 to 0.63)
d-NLR	1.29 ± 0.68	0.97 (0.68-4.06)	3.46	0.712[a]	ρ = 0.08 (95% CI –0.45 to 0.57)
RDW/PLT	0.044 (0.03-0.10)	0.037 ± 0.010	0.03	0.109[a]	ρ = –0.34 (95% CI –0.64 to 0.03)

[a]Spearman correlation

d-NLR: derived Neutrophil-to-Lymphocyte Ratio, **ELR:** Eosinophil-to-Lymphocyte Ratio, **LMR:** Lymphocyte-to-Monocyte Ratio, **MPV:** Mean Platelet Volume, **NLR:** Neutrophil-to-Lymphocyte Ratio, **PCT:** Plateletcrit, **PDW:** Platelet Distribution Width, **PLR:** Platelet-to-Lymphocyte Ratio, **PLT:** Platelet, **RDW:** Red Cell Distribution Width, **SII:** Systemic Immune-Inflammation Index, **SIRI:** Systemic Inflammatory Response Index (Data are presented as mean ± standard deviation or median (min–max) according to the distribution)

0.004), whereas correlations for other indices were weak and non-significant.

In the 1:1 age- and sex-matched subset (n = 23 vs. 23), both MPV and PDW values remained lower in patients compared to controls, although the differences were no longer statistically significant (p > 0.05). The direction of change was consistent with the main 3:1 analysis, suggesting that the primary findings were not driven by unequal group sizes but rather by the limited power of the smaller matched sample. The positive correlation between PCT and HB grade (ρ = 0.58, p = 0.004) persisted, supporting its robustness across analytic approaches.

DISCUSSION

In this study, we evaluated the hematological inflammatory markers in pediatric patients diagnosed with BP at presentation and compared the findings with those of a healthy control group. The primary objective was to investigate whether routinely available blood-based biomarkers could reflect an underlying systemic inflammatory process in these patients. The important role of inflammation in the pathogenesis of BP has been well documented in the literature. Öztürk et al. reported that SII and SIRI demonstrated high predictive value in the diagnosis of pediatric BP cases (AUC = 0.77 and 0.68, respectively) (10). Similarly, Liu et al. reported that NLR, PLR, and SIRI values were significantly increased in children with BP compared to the control group (4). Kim et al. suggested that NLR and PLR could serve as prognostic indicators in pediatric BP (14). In a recent pediatric cohort, NLR, PLR, and monocyte-to-lymphocyte ratio did not significantly differ between patients and controls but showed a positive correlation with recovery time, suggesting that leukocyte-derived ratios may serve as prognostic indicators in children with BP (11). Despite several prior reports in pediatric and adult BP showing significant shifts in inflammatory markers (15-18), we observed no statistically significant patient–control differences for most indices, including NLR, PLR, SII, SIRI, LMR, eosinophil count, ELR, RDW, derived NLR, PCT, and RDW/PLT. Several factors may account for this discrepancy. First, our study population consisted predominantly of patients with mild clinical presentations, as evidenced by the high proportion of HB grade 2 cases. It is possible that systemic inflammatory responses are more pronounced in moderate to severe cases, and thus, our cohort may not have exhibited sufficient variation in inflammatory marker levels to yield statistically significant results. Çelik et al. reported that commonly used hematologic marker did not demonstrate meaningful prognostic value in pediatric BP. In their cohort, early clinical severity was the only parameter associated with recovery, while CBC-derived inflammatory indices failed to predict short-term outcomes (19). Our findings are also in line with a recent single-center pediatric cohort by Di Sarno et al., who systematically evaluated prognostic factors, laboratory markers, and treatment strategies in 88 children with BP. In that study, no significant associations were observed between NLR, PLR, and HB scores at 2-month follow-up, despite an overall complete recovery rate of 81.8%. Taken together with our negative results for a broad panel of CBC-derived indices, these data suggest that readily available systemic inflammatory markers may have limited and inconsistent prognostic utility in

pediatric BP. Moreover, Di Sarno et al. reported that corticosteroid use—and particularly higher prednisone doses—was associated with a greater proportion of incomplete recovery, while vitamin supplementation had no measurable impact on outcomes. This pattern supports the notion that pediatric BP may be characterized by relatively subtle or heterogeneous systemic inflammatory changes that are not reliably captured by routine blood-based biomarkers and may not be readily modifiable by standard anti-inflammatory or adjunctive therapies (12). However, given the relatively small samples and predominance of mild cases in both cohorts, modest prognostic effects of these markers cannot be definitively excluded.

The patient group showed modestly lower MPV and PDW values compared with controls (Cliff's δ = –0.34 for both; 95% CI –0.58 to –0.10), suggesting that in approximately two-thirds of randomly paired observations, the corresponding control value was higher. Although statistically significant, the standardized effects are small-to-moderate, suggesting that any platelet-related change is subtle at the systemic level. Platelet size heterogeneity and activation are reflected by MPV and PDW. In classical systemic inflammation, they often increase, reflecting the release of larger, metabolically active platelets from the marrow. Conversely, several pediatric acute inflammatory states have reported lower MPV in the early phase, plausibly due to preferential consumption/margination of larger platelets at inflamed microvascular beds or a transient shift toward producing smaller platelets under cytokine pressure. In some settings, MPV has even been described as a negative acute-phase marker. Our samples were obtained at presentation and before corticosteroid therapy, consistent with an early-phase profile. Together with the positive correlation between PCT and clinical severity (ρ = 0.58), these findings may suggest a platelet-axis involvement in pediatric BP—i.e., altered platelet mass/turnover with relatively smaller and more uniform circulating platelets rather than a broad leukocyte-driven response. The relatively decreased MPV and PDW in affected patients could reflect a blunted or insufficient bone marrow response, or a distinct inflammatory pattern differing from classical systemic inflammation (20, 21). Sen et al. reported significantly elevated WBC, neutrophil, and platelet counts in pediatric patients with BP; however, the findings regarding MPV and PDW were inconsistent (5). In one study, a selective reduction in MPV was reported in pediatric BP, occurring exclusively in affected children (6/17 vs 0/17 controls, p = 0.007) (11). Further research is required to better understand how MPV and PDW are associated with pediatric BP.

Apart from PCT, the correlations of HB grade with CBC-based markers (NLR, PLR, SII, SIRI, LMR, ELR, d-NLR, RDW/PLT, MPV, PDW) were weak and did not reach statistical significance. A moderate positive and statistically significant correlation was observed between HB grade and PCT (Spearman's ρ=0.58, p = 0.004), indicating that greater clinical severity may be associated with increased total platelet mass. As it incorporates both platelet count and volume, PCT may provide a composite measure of platelet activation and systemic inflammatory burden. A previous study reported that PCT levels were significantly higher in patients with BP compared to healthy controls (p = 0.008); however, no significant association was observed between PCT values and recovery time (22). Its potential role as a marker of disease severity in pediatric BP deserves further investigation in larger, prospective studies.

The etiology of BP remains largely idiopathic; however, immune-mediated, viral, and ischemic mechanisms are widely believed to contribute to its pathogenesis. Neurotropic viral exposure is thought to activate inflammatory pathways in the facial nerve, leading to focal edema, myelin loss, and ischemia. In addition to infectious triggers, various systemic and environmental factors—such as hyperglycemia, uncontrolled hypertension, severe pre-eclampsia, migraine, and prior exposure to radiation—have been reported to predispose individuals to facial nerve palsy by promoting or exacerbating underlying inflammatory and ischemic processes (23, 24). This multifactorial etiology may explain the heterogeneity in clinical presentations and the inconsistent findings regarding laboratory-based inflammatory markers observed across different studies.

This study has several noteworthy strengths. The use of an age- and sex-matched healthy control group enhances the reliability of the comparisons and reduces potential confounding. This study provides valuable insight to the existing literature by focusing exclusively on a pediatric population with idiopathic BP, whereas most prior studies have been conducted in adult or mixed-age cohorts using heterogeneous definitions and analytic approaches. By employing standardized formulas for both traditional and newly derived hematologic indices, reporting effect sizes alongside significance levels, and examining correlations with clinical severity, this study provides a more nuanced understanding of the inflammatory profile in pediatric BP. To our knowledge, while pediatric cohorts assessing SII and SIRI in BP have recently been reported, most prior pediatric studies have focused on traditional indices such as NLR, PLR, MPV, PDW and LMR. This study is the first pediatric analysis to evaluate ELR and RPR in relation to BP severity. Furthermore, the simultaneous evaluation of multiple indices within a single, well-defined cohort enhances interpretability and supports future meta-analytic integration. Together, these features strengthen the methodological quality and contextual contribution of the study to the pediatric neuroinflammation literature.

These findings should be interpreted in light of several limitations. The study is single-center and retrospective with limited power (n=23 vs 69; ~80% power for effects ≥ d≈0.67). Disease severity was mostly mild. Despite age/sex matching, hospital-based controls may differ on unmeasured factors. Although sampling occurred at presentation before steroids, onset-to-sampling times were not uniformly documented. With only a single patient in the HB grade 4 subgroup, precision for severity-stratified analyses was limited.

Conclusion

Modest reductions in MPV and PDW, together with a positive correlation between PCT and disease severity, suggest platelet-axis involvement in pediatric BP. Although their clinical utility remains limited, these readily available indices may provide supportive information when assessing initial disease severity or anticipating short-term recovery patterns. Incorporating such markers into routine evaluation may help refine early prognostic monitoring. From a clinical perspective, these results may cautiously support using MPV, PDW, and PCT as easily accessible, supplementary information when evaluating children with BP, while primary assessment and management should continue to rely on established clinical criteria. Larger, prospective studies are warranted to validate their role and clarify the underlying pathophysiological mechanisms.

Author Contributions

Study conception and design: **Meral Karadağ, Melek Yaman Ortaköylü;** data collection: **Meral Karadağ, Melek Yaman Ortaköylü;** analysis and interpretation of results: **Meral Karadağ, Melek Yaman Ortaköylü;** draft manuscript preparation: **Meral Karadağ, Melek Yaman Ortaköylü.** The authors reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Funding

The authors declare that the study received no funding.

Ethical Approval

This study was approved by the Malatya Turgut Özal University Health Sciences Scientific Research Ethics Committee (date: 05.08.2025, number: 2025-250)

REFERENCES

1. Gilden DH. Bell's palsy. N Engl J Med. 2004;351(13):1323-1331. https://doi.org/10.1056/NEJMcp041120
2. Lorch M, Teach SJ. Facial nerve palsy: etiology and approach to diagnosis and treatment. Pediatr Emerg Care. 2010;26(10):763-769. https://doi.org/10.1097/PEC.0b013e3181f3bd4a
3. Öztoprak Ü. İdiyopatik Periferik Fasiyal Paralizili Çocuklarda Hematolojik Parametrelerin ve Paralizi Derecesi ile İlişkilerinin Değerlendirilmesi. Van Tip Derg. 2021;28(3). https://doi.org/0.5505/vtd.2021. 34341
4. Liu J, Li G, Wu R, Qin X, Pan S, Liang P, Sun J. The systemic inflammatory response index as a novel diagnostic indicator for Bell's palsy. Br J Hosp Med (Lond). 2024;85(9):1-14. https://doi.org/10.12968/hmed.2024.0386
5. Sen HS, Aksoy F, Yıldız E. Evaluation of hematological parameters in children with idiopathic facial paralysis: A casecontrol study. Haseki Med Bull. 2024;62(1):36-41. https://doi.org/10.4274/haseki.galenos.2024.9363
6. Cayır S, Kılıçaslan C. Hematologic parameters as predictive markers in pediatric Bell's palsy. Eur Arch Otorhinolaryngol. 2021;278(5):1265-1272. https://doi.org/10.1007/s00405-020-06459-w
7. Imtiaz F, Shafique K, Mirza SS, Ayoob Z, Vart P, Rao S. Neutrophil lymphocyte ratio as a measure of systemic inflammation in prevalent chronic diseases in Asian population. Int Arch Med. 2012;5(1):2. https://doi.org/10.1186/1755-7682-5-2
8. Hu B, Yang XR, Xu Y, Sun YF, Sun C, Guo W, et al. Systemic immune-inflammation index predicts prognosis of patients after curative resection for hepatocellular carcinoma. Clin Cancer Res. 2014;20(23):6212-6222. https://doi.org/10.1158/1078-0432.CCR-14-0442
9. Forget P, Khalifa C, Defour JP, Latinne D, Van Pel MC, De Kock M. What is the normal value of the neutrophil-to-lymphocyte ratio? BMC Res Notes. 2017;10(1):12. https://doi.org/10.1186/s13104-016-2335-5
10. Öztürk B, Yaradılmış RM, Güneylioğlu MM, Güngör A, Bodur İ, Göktuğ A, et al. The role of systemic immuneinflammation index and systemic inflammation response index in the diagnosis of Bell's palsy. Behcet Uz Cocuk Derg. 2025;15(1):1-6. https://doi.org/10.4274/jbuch.galenos.2024.82574
11. Kara İS, Arslan YK. Hemogram Parameters of Children with Bell's Palsy at the Time of Admission and Their Findings in the Follow-up. Turk J Pediatr Dis. 2024:1-5.
12. Di Sarno L, Caroselli A, Graglia B, Causio FA, Gatto A, Pansini V, et al. Pediatric Bell's Palsy: Prognostic Factors, Management Strategy, and Treatment Outcomes. J Clin Med. 2024;14(1). https://doi.org/10.3390/jcm14010079
13. House JW, Brackmann DE. Facial nerve grading system. Otolaryngol Head Neck Surg. 1985;93(2):146-147. https://doi.org/10.1177/019459988509300202
14. Kim DH, Oh JH, Kim J, Cho CH, Lee JH. Predictive values of neutrophil-tolymphocyte ratio, platelet-tolymphocyte ratio, and other prognostic factors in pediatric patients with Bell's palsy. Ear Nose Throat J. 2021;100(10):720-725. https://doi.org/10.1177/0145561320922097
15. Oya R, Takenaka Y, Imai T, Sato T, Oshima K, Ohta Y, et al. Neutrophil-to-Lymphocyte Ratio and Platelet-to-Lymphocyte Ratio as Prognostic Hematologic Markers of Bell's Palsy: A Meta-analysis. Otol Neurotol. 2019;40(5):681-687. https://doi.org/10.1097/MAO.0000000000002166
16. Kim SJ, Lee HY. Hematological Findings in Patients with Acute Peripheral Facial Palsy. J Int Adv Otol. 2020;16(3):382-386. https://doi.org/10.5152/iao.2020.8748
17. Yıldız MG, Orhan İ, Kara İ, Sağıroğlu S, Bilal N, Doğaner A, et al. Bell Paralizisi Olan Hastalarda Hematolojik Parametrelerin Prognostik Değeri Var Mıdır? Ahi Evran Med J. 2022;6(3):304-309. https://doi.org/10.46332/aemj.1008500

18. Kınar A, Ulu Ş, Bucak A, Kazan E. Can Systemic Immune-Inflammation Index (SII) be a prognostic factor of Bell's palsy patients? Neurol Sci. 2021;42(8):3197-3201. https://doi.org/10.1007/s10072-020-04921-5

19. Çelik C, Tutar B, Berkiten G, Kumral TL, Atar Y, Sarı H, et al. Prognostic value of inflammatory and radiologic markers in pediatric bell's palsy: a retrospective study. Eur Arch Otorhinolaryngol. 2025;3. https://doi.org/10.1007/s00405-025-09653-w

20. Korniluk A, Koper-Lenkiewicz OM, Kamińska J, Kemona H, Dymicka-Piekarska V. Mean Platelet Volume (MPV): New Perspectives for an Old Marker in the Course and Prognosis of Inflammatory Conditions. Mediators Inflamm. 2019;2019:9213074. https://doi.org/10.1155/2019/9213074

21. Karadag-Oncel E, Ozsurekci Y, Kara A, Karahan S, Bulent Cengiz A, Ceyhan M. The value of mean platelet volume in the determination of community acquired pneumonia in children. Ital J Pediatr. 2013;39:16. https://doi.org/10.1186/1824-7288-39-16

22. Eliçora ŞŞ, Erdem D. Can blood parameters be used to assess the recovery time of Bell's palsy patients? A retrospective case-controlled study. Kulak Burun Bogaz Uygulamaları. 2017;5(1):1-6. https://doi.org/10.5606/kbbu.2017.07769

23. Lu J, Yin Z, Qiu Y, Yang Y, Chen Z, Wu J, et al. Exploring the Role of Inflammation and Metabolites in Bell's Palsy and Potential Treatment Strategies. Biomedicines. 2025;13(4):957. https://doi.org/10.3390/biomedicines13040957

24. Rajangam J, Lakshmanan AP, Rao KU, Jayashree D, Radhakrishnan R, Roshitha B, et al. Bell Palsy: Facts and Current Research Perspectives. CNS Neurol Disord Drug Targets. 2024;23(2):203-214. https://doi.org/10.2174/1871527322666230321120618

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):89-96
DOI: 10.29058/mjwbs.1734160

Evaluation of practical application of Tell-Hashomer, Livneh and Yalçınkaya-Özen criteria for diagnosis of Familial Mediterranean Fever in children

Çocuklarda Ailevi Akdeniz Ateşi tanısında Tel-Hashomer, Livneh ve Yalçınkaya-Özen kriterlerinin pratikteki uygulamasının değerlendirilmesi

Tuğba Kazlı[1], Müferet Ergüven[2]

[1]Department of Pediatrics, Pediatric Pulmonology, Zonguldak Bülent Ecevit University, School of Medicine, Zonguldak, Türkiye
[2]Department of Pediatrics, Division of Pediatric Rheumatology, Düzce University, School of Medicine, Düzce, Türkiye

Cite this article as: *Kazlı T and Ergüven M. Evaluation of practical application of Tell-Hashomer, Livneh and Yalçınkaya-Özen criteria for diagnosis of Familial Mediterranean Fever in children. Med J West Black Sea. 2026;10(1):89-96.*

This study was derived from Tugba Kazli's bachelor's thesis.

ABSTRACT

Aim: Familial Mediterranean Fever (FMF) is an autoinflammatory disease with abdominal pain, recurrent fever, pleuritis, arthritis, and skin lesions. Tel-Hashomer, Livneh and Yalçınkaya-Özen Criteria amongst various diagnostic criteria are introduced.

Material and Methods: This study is conducted to evaluate the Tel-Hashomer, Livneh, and Yalçınkaya-Özen Criteria among pediatric FMF patients for practical application. A total of 113 patients' records were reassessed according to three criteria, and the likelihood of an FMF diagnosis was calculated based on the logistic model developed by Yalçınkaya et al. Patients' likelihood of having FMF according to the logistic model, the rates of FMF diagnosis at presentation and after six months of colchicine treatment were calculated, and the rates of diagnosis after six months of colchicine treatment according to the Tel-Hashomer criteria were compared with the rates of diagnosis at presentation according to the Livneh and Yalçınkaya-Özen criteria.

Results: A statistically significant difference was observed between the criteria for making a positive diagnosis at presentation. Although diagnostic rates increased by the end of 6-months, statistical analysis still showed significant difference. Using the logistic model 85.84% of our patients (n=97) had a likelihood of FMF diagnosis above 80%.

Conclusion: Tel-Hashomer criteria should be preferred for their simplicity in diagnosing FMF but they appear less sensitive compared to the Livneh and Yalçınkaya-Özen Criteria. Therefore, for patients who do not receive a diagnosis with the Tel-Hashomer criteria at presentation or after six months of treatment, the more sensitive Livneh and Yalçınkaya-Özen Criteria can be applied in practice.

Keywords: Diagnostic criteria, Familial Mediterranean Fever, pediatric

ÖZ

Amaç: Ailevi Akdeniz Ateşi (AAA); karın ağrısı, tekrarlayan ateş, plevrit, artrit ve erizipel benzeri deri lezyonları ile karakterize olan otoinflamatuar bir hastalıktır. Tanısında Tel-Hashomer, Livneh, Yalçınkaya-Özen kriterleri gibi pekçok tanı kriteri tanımlanmıştır.

Gereç ve Yöntemler: Çalışmamızda pediatrik yaş grubundaki AAA hastalarında Tel-Hashomer, Livneh ve Yalçınkaya- Özen tanı kriterlerinin pratik uygulamadaki yerlerini değerlendirmek amaçlandı. 113 AAA hasta dosyası her üç tanı kriterine göre yeniden değerlendirildi ve Yalçınkaya ve arkadaşlarının geliştirdiği lojistik modele göre her hastanın AAA olma olasılığı hesaplandı. Hastaların Lojistik modele göre AAA olma olasılıkları, başvuru sırasında ve altı aylık kolşisin tedavisi sonrasında AAA tanısı alma oranları, Tel-Hashomer kriterlerine göre 6 aylık kolşisin tedavisi sonrasında tanı alma oranları ile Livneh ve Yalçınkaya-Özen kriterlerine göre başvuru sırasındaki tanı alma oranları kıyaslandı. Karşılaştırmalar istatistiksel olarak değerlendirildi.

Bulgular: Kriterler arasında geliş anında tanı koymaları yönünden istatistiksel olarak anlamlı fark saptandı. 6 aylık kolşisin tedavisinden sonra tanı oranları artmasına rağmen istatistiksel olarak kriterler arasında tanı koymaları yönünden anlamlı fark devam etti. Lojistik modele göre hastalarımızın %85,84'ünde (n=97),AAA olma ihtimali %80'in üzerindeydi.

Sonuç: Tel-Hashomer kriterleri FMF tanısını koymada daha basit dizayn edilmiş olması nedeni ile başvuru anında tercih edilebilir ancak Livneh ve Yalçınkaya-Özen kriterlerine göre daha az hassas görünmektedir. Bu nedenle yüksek klinik şüphe durumunda başvuru anında veya 6 aylık tedavi sonrasında tanı konulamayan hastalara Livneh ve Yalçınkaya-Özen kriterlerinin uygulanabileceği önerilebilir.

Anahtar Kelimeler: Tanı kriterleri, Ailevi Akdeniz Ateşi, pediatrik

Corresponding Author: Tuğba Kazlı ✉ tugba.karsan@gmail.com **Received:** 07.07.2025 **Revision:** 02.09.2025 **Accepted:** 05.01.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Tel-Hashomer, Livneh, and Yalçınkaya–Özen criteria were compared for practical diagnosis of pediatric FMF.
- Livneh and Yalçınkaya–Özen criteria showed significantly higher diagnostic yield at presentation.
- Tel-Hashomer criteria remained less sensitive even after 6 months of colchicine therapy.
- Tel-Hashomer criteria is practical to use in outpatient clinics.
- For children with high clinical suspicion of Familial Mediterranean Fever who remain undiagnosed according to the Tel-Hashomer criteria, the use of more sensitive diagnostic criteria may enable earlier diagnosis.

INTRODUCTION

Familial Mediterranean Fever (FMF) is a hereditary autoinflammatory disease associated with mutations in the *MEFV* gene located on chromosome 16. It predominantly affects populations originating from the Mediterranean region, including Turks, Armenians, Arabs, and Sephardic Jews. The disease is characterized by recurrent, self-limiting episodes of fever accompanied by serositis, such as peritonitis, pleuritis, and arthritis, as well as erysipelas-like skin lesions (1-3). The *MEFV* gene encodes pyrin, a protein involved in the regulation of the inflammatory response. Mutations affecting pyrin function lead to dysregulation of inflammation and result in the clinical manifestations of FMF (4). The manifestation of disease phenotype is influenced by environmental factors (5), the mRNA expression of the MEFV gene, and the number and type of mutations (6). Besides the classical autosomal recessive inheritance pattern, a vertical inheritance pattern has also been reported in some families from places with a high carrier frequency of MEFV variants (7). The most common disease-causing mutations are M694V, M694I, V726A and M680I located in exon 10 (8). Different from other autosomal recessive inherited diseases, variable clinical expression has been observed even in heterozygous individuals (9).

Currently, there is no single definitive laboratory test for FMF. During acute attacks, elevated levels of acute-phase reactants such as C-reactive protein, erythrocyte sedimentation rate, fibrinogen, and serum amyloid A can be observed. Among these, serum amyloid A has been suggested as a sensitive marker for detecting subclinical inflammation (10,11). The diagnosis of FMF involves a combination of clinical findings, family history, and molecular screening. Diagnosis of FMF is primarily based on clinical findings supported by genetic testing and family history. Several diagnostic criteria (Table 1) have been proposed over time, including the Tel-Hashomer (12), Livneh (13), and Yalçınkaya–Ozen criteria (14). These criteria differ in terms of sensitivity and specificity and may lead to variations in diagnostic outcomes, particularly in pediatric patients.

The present study aimed to evaluate the practical applicability of these three diagnostic criteria sets in pediatric FMF patients and to compare their diagnostic performance both at presentation and after colchicine treatment.

MATERIAL and METHODS

This retrospective study included pediatric patients diagnosed with FMF and followed at the Pediatric Rheumatology outpatient clinic of a tertiary care center. A total of 113 patients who met the predefined inclusion criteria and had complete medical records were included in the analysis. Approval for the study was obtained from the hospital' s Ethics Committee according to the Helsinki declarations.

All cases were re-evaluated according to the Tel-Hashomer, Livneh, and Yalçınkaya–Özen diagnostic criteria (Table 1). Patient data were collected from medical records and included demographic characteristics (Table 2), genetic mutation profiles (Table 3), clinical findings and laboratory parameters (Table 4) and diagnostic classification according to the three criteria sets (Table 5). The positive diagnostic yield of the Tel-Hashomer, Livneh, and Yalçınkaya-Ozen diagnostic criteria sets were compared both at presentation and after at least six months of colchicine therapy (Tables 6-8).

In addition, the probability of FMF for each patient was calculated using the logistic regression model developed by Yalçınkaya et al. (Table 9). This model incorporates key clinical variables including fever, abdominal pain, chest pain, arthritis, and family history.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics software (version 17.0). Continuous variables were expressed as mean ± standard deviation, while categorical variables were presented as numbers and percentages. Comparisons between diagnostic criteria were performed using the McNemar test. A p-value of <0.05 was considered statistically significant.

As this study included all eligible patients within the study period, no prior sample size calculation was performed.

Table 1: Diagnostic Criteria Sets for Familial Mediterranean Fever

Tel Hashomer Criteria	Livneh Criteria	Yalçınkaya-Özen Criteria
Major Criteria	**Major Criteria**	
1. Recurrent febrile episodes with serositis (peritonitis, synovitis, and pleuritis)	1. Peritonitis (generalized)	1. Fever (axillary temperature of > 38 °C, 6-72 h, ≥ 3 attacks)
2. Amyloidosis of AA type without a predisposing disease	2. Pleuritis (unilateral) or pericarditis	2. Abdominal pain (6–72 h duration, ≥ 3 attacks)
3. Favorable response to colchicine treatment	3. Monoarthritis (hip, knee, ankle)	3. Chest pain (6–2 h duration, ≥ 3 attacks)
	4. Fever alone	4. Arthritis (6–72 h duration, ≥ 3 attacks, oligoarthritis)
	5. Incomplete attacks involving abdomen • the temperature is normal or lower than 38°C • the attacks are longer or shorter than specified (but not shorter than 6 hours or longer than a week) • no signs of peritonitis are recorded during the abdominal attacks • the abdominal attacks are localized • the arthritis is in joints other than those specified	5. Family history of FMF
Minor Criteria	**Minor criteria** 1 or more of the following sites:	
1. Recurrent febrile episodes	1. Incomplete attacks involving chest	
2. Erysipelas-like erythema	2. Incomplete attacks involving joınts	
3. FMF in a first-degree relative	3. Exertional leg pain	
	4. Favorable response to colchicine	
	Supportive criteria:1- Family history of FMF, 2-Appropriate ethnic origin, 3-Age <20 years at disease onset, 4-Severe, requiring bed rest, 5-Spontaneous remission, 6-Symptom-free interval, 7-Transient inflammatory response, with 1 or more abnormal test result(s) for white blood cell count, erythrocyte sedimentation rate, serum amyloid A, and/or fibrinogen, 8-Episodic proteinuria/hematuria, 9-Unproductive laparotomy or removal of white appendix, 10-Consanguinity of parents	
Diagnosis	Diagnosis	Diagnosis
≥ 2 major or 1 major + 2 minor criteria	1 major criteria, or 2 minor criteria, or 1 minor criterion plus 5 supportive criteria, or 1 minor criterion plus 4 of the first 5 supportive criteria	≥ 2 criteria

The diagnostic criteria for Familial Mediterranean Fever (FMF) were derived from the original publications describing the Tel-Hashomer, Livneh, and Yalçınkaya–Özen criteria (1,13,14).

RESULTS

A total of 113 patients were included in the study, with a mean age of 9.48 ± 4.29 years. The cohort consisted of 63 females (55.8%) and 50 males (44.2%) (Table 2).

Genetic analysis revealed that the most common mutation was M694V, observed in both homozygous and heterozygous forms. A small proportion of patients had no detectable mutation or missing genetic data (Table 3).

The most frequently observed clinical manifestations were abdominal pain (87.6%), fever (76.1%), and peritonitis (74.3%). Other common findings included arthralgia (66.4%), myalgia (25.7%), and chest pain (25.7%). Laboratory indicators of inflammation were elevated in 62.8% of patients (Table 4).

Table 2: Demographic and Anthropometric Features of the Study Population

	Min-Max	Mean ± SD	
Age	**1.5 - 18 yrs**	**9.48 ± 4.29**	
		n	**%**
Age groups	< 5 yrs	18	15.93
	5 - 9 yrs	42	37.17
	10 - 14 yrs	29	25.66
	18- 14 yrs	24	21.24
Gender	Female	63	55.75
	Male	50	44.25

SD: Standard deviation; **SDS:** Standard deviation score. Categorical variables are presented as n (%), continuous variables as mean ± SD. Total number of patients: n=113.

Table 3. Distribution of MEFV Gene Mutations in Patients with Familial Mediterranean Fever

Mutation Type	Genotype / Mutation	n	%
Homozygous	M694V / M694V	24	21.2
	M680I / M680I	2	1.8
	E148Q / E148Q	1	0.9
Heterozygous	M694V / –	20	17.7
	V726A / –	7	6.2
	M680I / –	6	5.3
	E148Q / –	5	4.4
Compound Heterozygous	M694V / V726A	4	3.5
	M694V / M680I	7	6.2
	M694V / R761H	1	0.9
	V726A / M680I	1	0.9
	M694V / V726A / M680I (triple)	6	5.3
Complex / Mixed Forms	M694V / M694V; V726A / –	1	0.9
	V726A / –; M680I / M680I	2	1.8
No mutation detected	-	9	8.0
Not mentioned	-	17	15.0

Percentages were calculated based on the total number of patients (n=113). **MEFV:** Mediterranean fever gene.

Table 4. Clinical and Laboratory Features of the Patients

Clinical / Laboratory Finding	n	%
Abdominal pain	99	87.6
Fever	86	76.1
Peritonitis	84	74.3
Arthralgia	75	66.4
Laboratory findings supporting inflammation*	71	62.8
Chest pain	29	25.7
Myalgia	29	25.7
Monoarthritis	28	24.8
Episodic proteinuria / hematuria	22	19.5
Appendectomy history	11	9.7
Erysipelas-like erythema	8	7.1
Amyloidosis	1	0.9
Pleuritis	2	1.8

*Elevated ESR, CRP, fibrinogen, or WBC.

Table 5. FMF diagnostic rates at presentation and after 6-months long colchicine treatment(Tel Hashomer, Livneh and Yalçınkaya-Özen Criteria)

Name of criteria set	Number of criterion	FMF Diagnostic Rates At Presentation			FMF Diagnostic Rates After 6-Months Colchicine Treatment		
		Number of patients diagnosed (n)	Total number of patients diagnosed (n)	Persentage of patients diagnosed (%)	Number of patients diagnosed (n)	Total number of patients diagnosed (n)	Persentage of patients diagnosed (%)
Tel Hashomer, Criteria	1 major+2 minor	9	10	8.8	5	89	78.7
	2 major				46		
	2 major+1minor				32		
	2 major+2 minor	1			5		
	3 major+2 minor				1		
Livneh Criteria	≥1 major	107	110	97.3	107	111	98
	≥ 2 minor				2		
	1minor + ≥6 supporting criteria	2			1		
	1 minor criterion + ≥ 4 of the first 5 supportive criteria	1			1		
Yalçınkaya-Özen Criteria	2 criteria	32	104	92	32	104	92
	3 criteria	58			58		
	4 criteria	13			13		
	5 criteria	1			1		

Table 6a. Comparison of the Tel-Hashomer and Livneh criteria (at the time of presentation)

	Livneh Criteria(1)		p
Tel-Hashomer criteria(1)		positive	negative	0.001
	positive	10	-	
	negative	100	3	

(1): Diagnosis at the time of presentation

Table 6b. Comparison of the Tel-Hashomer and Yalçınkaya-Özen Criteria (at the time of presentation)

	Yalçınkaya-Özen Criteria(1)		P
Tel-Hashomer criteria(1)		positive	negative	0.001
	positive	10	-	
	negative	94	9	

(1): Diagnosis at the time of presentation

Table 6c. Comparison of the Livneh and Yalçınkaya-Özen Criteria (at the time of presentation)

	Livneh Criteria(1)		P
Yalçınkaya-Özen Criteria(1)		positive	negative	0.031
	positive	104	-	
	negative	6	3	

(1): Diagnosis at the time of presentation

Table 8a. Comparison of the Tel-Hashomer(after at least 6 months of colchicine therapy) and Livneh Criteria (at the time of presentation)

	Livneh Criteria(1)		p
Tel-Hashomer criteria (2)		positive	negative	0.001
	positive	88	1	
	negative	22	2	

(1): Diagnosis at the time of presentation (2): Diagnosis after at least 6 months of colchicine therapy

Table 8b. Comparison of the Tel-Hashomer(after at least 6 months of colchicine therapy) and Yalçınkaya-Özen Criteria (at the time of presentation)

	Yalçınkaya-Özen Criteria(1)		P
Tel-Hashomer criteria (2)		positive	negative	0.001
	positive	87	2	
	negative	17	7	

(1): Diagnosis at the time of presentation (2): Diagnosis after at least 6 months of colchicine therapy

Table 7a. Comparison of the Tel-Hashomer and Livneh criteria (after at least 6 months of colchicine therapy)

	Livneh Criteria(2)		P
Tel-Hashomer criteria (2)		positive	negative	0.001
	positive	88	1	
	negative	23	1	

(2): Diagnosis after at least 6 months of colchicine therapy

Table 7b. Comparison of the Tel-Hashomer and Yalçınkaya-Özen Criteria (after at least 6 months of colchicine therapy)

	Yalçınkaya-Özen Criteria(2)		P
Tel-Hashomer criteria (2)		positive	negative	0.001
	positive	87	2	
	negative	17	7	

(2): Diagnosis after at least 6 months of colchicine therapy

Table 7c. Comparison of the Livneh and Yalçınkaya-Özen Criteria (after at least 6 months of colchicine therapy)

	Livneh Criteria(2)		P
Yalçınkaya-Özen Criteria(2)		positive	negative	0.016
	positive	104	-	
	negative	7	2	

(2): Diagnosis after at least 6 months of colchicine therapy

Table 9. Distribution of FMF Probability According to the Logistic Model Developed by Yalcinkaya et al.

Probability Range (P)	**n**	**%**
≥ 0.99	43	38
0.95 – 0.989	19	16.8
0.90 – 0.949	12	10.6
0.80 – 0.899	23	20.3
0.70 – 0.799	3	2.7
0.50 – 0.699	5	4.4
0.30 – 0.499	3	2.7
< 0.30	5	4.4
Total	113	100.0

FMF probability for each patient was calculated using the logistic regression model developed by Yalçınkaya et al.
Formula: *A=exp(-2.55 + 1.99×fever + 2.09×abdominal pain + 1.49×chest pain + 1.35×arthritis + 3.36×positive family history); P(FMF)=A / (1 + A).*

At the time of presentation, diagnostic rates differed significantly among the criteria sets. The Tel-Hashomer criteria identified a smaller proportion of patients compared to the Livneh and Yalçınkaya–Özen criteria. Following six months of colchicine therapy, diagnostic rates increased across all criteria; however, statistically significant differences persisted (Table 5).

At the time of presentation, the diagnostic rates based on clinical findings were as follows:

- **Tel-Hashomer criteria:** 8.8% (n=10)
- **Livneh criteria:** 97.3% (n=110)
- **Yalçınkaya-Özen Criteria:** 92% (n=104)

The McNemar test revealed statistically significant differences among the criteria in terms of positive diagnosis at presentation (Tel-Hashomer vs. Livneh*, p=0.001; Tel-Hashomer vs. Yalçınkaya-Özen*, p=0.001; Livneh* vs. Yalçınkaya-Özen, p=0.031) (Table 6A-C).

After at least six months of colchicine treatment, the diagnostic rates were:

- **Tel-Hashomer criteria:** 78.7% (n=89)
- **Livneh criteria:** 98% (n=111)
- **Yalçınkaya-Özen Criteria:** 92% (n=104)

When comparing the Tel-Hashomer and Livneh criteria, 88 patients were positive and one patient was negative according to both criteria. The McNemar test showed a highly significant difference favoring the Livneh criteria (p=0.001) (Table 7A). When comparing the Tel-Hashomer and Yalçınkaya-Özen criteria, 87 patients were positive and 7 patients were negative according to both. A highly significant difference again favored the Yalçınkaya-Özen criteria (p=0.001) (Table 7B). In the comparison of the Yalçınkaya-Özen and Livneh criteria, 104 patients were positive and two patients were negative according to both. A statistically significant difference was also observed favoring the Livneh criteria (p=0.016) (Table 7C).

Following a six-month course of colchicine therapy, the Tel-Hashomer criteria continued to demonstrate significantly lower diagnostic yield compared with the baseline Livneh and Yalçınkaya-Özen assessments (Table 8A,B).

Finally**,** pairwise comparisons using the McNemar test demonstrated significant differences between all three criteria sets, with the Livneh and Yalçınkaya–Özen criteria showing higher diagnostic sensitivity. It is observed that a group of patients accounting for 10% (n=11) was non-responsive to colchicine therapy.

According to the logistic regression model, 85.8% of patients had a high probability (>80%) of FMF diagnosis (Table 9).

DISCUSSION

Since there is no specific laboratory test for FMF, diagnosis relies on clinical data, including ethnic background, family history, colchicine response, and diagnostic criteria (15). In recent years, mutation analyses have also been widely used to support clinical diagnosis. Because specific diagnostic criteria for childhood FMF are lacking, the criteria recommended for adults, like Tel-Hashomer and Livneh, are used. However, the validity of these criteria in children has not been thoroughly investigated. New criteria specifically for childhood FMF were developed by Yalçınkaya et al. (14).

When comparing Tel-Hashomer and Livneh criteria, the diagnostic rates at presentation were 8.8% and 97.3%, respectively, and after assessing colchicine response, the rates reached up to 78.7% and 98%. These differences in diagnostic rates both at presentation (p=0.001) and after colchicine response (p=0.001) suggesting that the Livneh criteria were more sensitive. The higher sensitivity of the Livneh criteria can be attributed to inclusion of more parameters, yielding complexity and long time requirement for outpatient settings. On the other hand, criteria such as ethnic background- in a Turkish cohort-, age of onset- in a pediatric group - and consanguinity—consanguineous marriages in our country are prevalent at a rate of 21% (16) results in reduced specificity in our cohort. Nevertheless, the lack of a control group limits our ability to assess sensitivity and specificity comprehensively. Additionally, it is wellknown that 5-10% of FMF patients exhibit colchicine resistance (17). In our study, this resistance rate was found to be 10% (n=11), highlighting one of the limitations of the Tel-Hashomer criteria.

When comparing Tel-Hashomer and Yalçınkaya-Özen Criteria, the diagnostic rates at presentation were 8.8% and 92%, and after colchicine response, the rates reached up to 78.7% and 92%. Statistically significant differences were observed between the criteria both at presentation (p=0.001) and after colchicine response (p=0.001). This result can be attributed to difficulties in detecting renal amyloidosis with the Tel-Hashomer criteria and false negative results due to colchicine resistance. While Yalçınkaya-Özen Criteria Set diagnosing more patients compared to Tel-Hashomer's, and they do not include as many parameters as the Livneh's which restrict their utilization in outpatient settings. Nonetheless, Yalçınkaya's study focused on homozygous and compound heterozygous patients, and sensitivity-specificity studies for patients without mutations or with mutations in only one allele have not yet been performed.

When comparing the Livneh and Yalçınkaya-Özen Criteria, the rates of diagnosis at presentation were 97.3% and 92%, respectively, while after assessing colchicine response, the rates were 98% and 92%. Statistically significant differences were observed between the two criteria both at presentation (p=0.031) and after colchicine response (p=0.016). This higher sensitivity of the Livneh criteria can be attributed to the broader range of parameters they include. However, the inclusion of criteria such as age of initial diagnosis, appropriate ethnic background and consanguinity in our pediatric cohort reduces its specificity. Additionally, Livneh

criteria are not widely used in our clinic, limiting its practical applicability.

The aim of the study was to address the question of whether we can use diagnostic criteria other than the Tel-Hashomer criteria, which require a six-month waiting period and interventional procedure like kidney biopsy to make a diagnosis sooner and simpler. The Tel-Hashomer criteria are relatively simple to apply for outpatient settings by involving fewer parameters. In contrast, assessing AA-type amyloidosis in every patient is often impractical and unnecessary, and even after six months of colchicine treatment, a negative response in 5-10% of cases does not necessarily indicate the absence of the disease (17). Major criteria such as the presence of amyloidosis and a positive colchicine response significantly impact the diagnostic yield of the Tel-Hashomer criteria.

While the Livneh criteria typically take longer to evaluate than Tel-Hashomer, their usage is limited in our clinic due to their complexity and the specificity reduction caused by some supporting criteria. Although colchicine response is also evaluated in Livneh criteria, as a minor criterion, it has no statistical significance (97.3%→98%). The Yalçınkaya Ozen criteria, being more practical due to their brief composition, simpler design and and child-oriented structure, offer an advantage by addressing the disease features more clearly. However, the limitation in sensitivity and specificity studies for patients with mutations in only one allele or no mutations restricts their use.

There was a statistically significant difference (p=0.001) between the positive diagnostic rate of the Tel-Hashomer criteria after six months of treatment (78.7%, n=89) and the positive diagnositic rate of the Livneh criteria at presentation (97.3%, n=110). This indicates that even after six months, the Tel-Hashomer criteria may diagnose fewer patients due to the reasons mentioned above.

In the McNemar analysis used to compare the diagnostic yields of the three criteria, some cells appeared empty. This finding is not primarily related to an insufficient sample size but rather reflects the asymmetric distribution of discordant pairs between the criteria. Because all 113 patients were evaluated simultaneously using different diagnostic sets, the majority of patients were consistently classified in the same direction (positive or negative) by both criteria being compared. As a result, discordant classifications were very few in certain comparisons, leading to zero-frequency cells. McNemar's test remains an appropriate statistical method for paired nominal data, yet this distributional characteristic should be taken into consideration when interpreting the significance of the results.

The main limitation of our study is being monocentric retrospective study of small sample size without a control group. Furthermore all patients were of Turkish origin. Hence, further multicenter prospective studies with large sample sizes, different ethnic groups with controls are needed.

Conclusion

Our study showed while the Tel-Hashomer criteria are simpler and should be preferred for initial diagnosis, the more sensitive criteria of Yalcinkaya Ozen and Livneh can be applied for patients who do not receive a diagnosis with Tel-Hashomer criteria at presentation or after six months of treatment.

Author Contributions

The author (s) confirm (s) contribution to the paper as follows: Study conception and design: **Muferet Erguven;** data collection: **Tugba Kazli;** analysis and interpretation of results: **Tugba Kazli;** draft manuscript preparation: **Tugba Kazli.** The author (s) reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Financial Support

The authors declare that the study received no funding.

Ethical Approval

The study protocol was approved by the Ethics Committee for Non-Interventional Clinical Research of Istanbul Goztepe Training and Research Hospital (date: 02.06.2009; approval no.: 57/Ç12). The work related to this manuscript was carried out in accordance with the World Medical Association Declaration of Helsinki.

REFERENCES

1. Sohar E, Gafni J, Pras M, Heller H. Familial Mediterranean fever: a survey of 470 cases and review of the literature. Am J Med. 1967;43 (2):227-253. https://doi.org/10.1016/0002-9343(67)90167-2
2. Ben-Chetrit E, Levy M. Familial Mediterranean fever. Lancet. 1998;351:659-664. https://doi.org/10.1016/S0140-6736(97)09408-7
3. Samuels J, Aksentijevich I, Torosyan Y, Centola M, Deng Z, Sood R, et al. Familial Mediterranean fever at the millennium: clinical spectrum, ancient mutations, and a survey of 100 American referrals to the National Institutes of Health. Medicine (Baltimore). 1998;77 (4):268-297. https://doi.org/10.1097/00005792-199807000-00005
4. Pras M. Familial Mediterranean fever: from the clinical syndrome to the cloning of the pyrin gene. Scand J Rheumatol. 1998;27 (2):92-97. https://doi.org/10.1080/030097498440949
5. The French FMF Consortium. A candidate gene for familial Mediterranean fever. Nat Genet. 1997;17 (1):25-31. https://doi.org/10.1038/ng0997-25
6. Barakat MH, Malhas LN, Gumaa KA. Catecholamine metabolism in recurrent hereditary polyserositis: pathogenesis of acute inflammation—the retention-leakage hypothesis. Biomed Pharmacother. 1989;43:763-769. https://doi.org/10.1016/0753-3322(89)90165-0

7. Booth DR, Gillmore JD, Lachmann HJ, Booth SE, Bybee A, Soytürk M, et al. The genetic basis of autosomal dominant familial Mediterranean fever. QJM. 2000;93 (4):217-221. https://doi.org/10.1093/qjmed/93.4.217
8. Ben-Zvi I, Herskovizh C, Kukuy O, Kassel Y, Grossman C, Livneh A. Familial Mediterranean fever without MEFV mutations: a case-control study. Orphanet J Rare Dis. 2015;10:34. https://doi.org/10.1186/s13023-015-0252-7
9. Ozdogan H, Ugurlu S. Familial Mediterranean fever. Presse Med. 2019;48 (1 Pt 2):e61-e76. https://doi.org/10.1016/j.lpm.2018.08.014
10. Shioya M, Taniguchi S, Masumoto J, Yasui K, Koike K, Komiyama A, et al. ASC, which is composed of a PYD and CARD, is up-regulated by inflammation and apoptosis in human neutrophils. Biochem Biophys Res Commun. 2002;293 (5):1314-1318. https://doi.org/10.1016/S0006-291X(02)00384-4
11. Bakkaloğlu A. Familial Mediterranean fever. Pediatr Nephrol. 2003;18 (9):853-859. https://doi.org/10.1007/s00467-003-1185-2
12. Daniels M, Shohat T, Brenner-Ullman A, Shohat VI. Familial Mediterranean fever: high gene frequency among the non-Ashkenazic and Ashkenazic Jewish populations in Israel. Am J Med Genet. 1995;55 (3):311-314. https://doi.org/10.1002/ajmg.1320550313
13. Livneh A, Langevitz P, Zemer D, Zaks N, Kees S, Lidar T, et al. Criteria for the diagnosis of familial Mediterranean fever. Arthritis Rheum. 1997;40 (10):1879-1885. https://doi.org/10.1002/art.1780401023
14. Yalçinkaya F, Özen S, Özçakar ZB, Aktay N, Çakar N, Düzova A, et al. A new set of criteria for the diagnosis of familial Mediterranean fever in childhood. Rheumatology (Oxford). 2009;48 (4):395-398. https://doi.org/10.1093/rheumatology/ken509
15. Tekin M, Yalçinkaya F, Tümer N, Çakar N, Koçak H, Özkaya N, et al. Familial Mediterranean fever: renal involvement by diseases other than amyloid. Nephrol Dial Transplant. 1999;14 (2):475-479. https://doi.org/10.1093/ndt/14.2.475
16. Tunçbilek E, Koç I. Consanguineous marriage in Turkey and its impact on fertility and mortality. Ann Hum Genet. 1994;58 (4):321-329. https://doi.org/10.1111/j.1469-1809.1994.tb00729.x
17. Zemer D, Revach M, Pras M, Modan B, Schor S, Sohar E, et al. A controlled trial of colchicine in preventing attacks of familial Mediterranean fever. N Engl J Med. 1974;291 (18):932-934. https://doi.org/10.1056/NEJM197410312911803

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):97-105
DOI: 10.29058/mjwbs.1795745

The effect of a surgical nursing course structured with role-playing and flipped learning methods on students' motivation and attitudes toward clinical practice

Rol oynama ve ters yüz öğrenme yöntemi ile yapılandırılmış cerrahi hemşireliği dersinin öğrencilerin motivasyonuna ve klinik uygulamaya yönelik tutumuna etkisi

Sibel Altıntaş, Sevim Çelik, Elif Karahan, Özge Uçar, Meryem Yücel

Department of Nursing, Bartın University, Faculty of Health Sciences, Bartın, Türkiye

Cite this article as: *Altıntaş S et al. The effect of a surgical nursing course structured with role-playing and flipped learning methods on students' motivation and attitudes toward clinical practice. Med J West Black Sea. 2026;10(1):97-105.*

ABSTRACT

Aim: Nursing education is a holistic process with both theoretical and practical dimensions. Students' motivation and attitudes toward clinical practice are crucial for success.

This study aimed to examine the effects of a surgical nursing course structured with role-playing and flipped learning methods on nursing students' motivation and clinical practice attitudes.

Material and Methods: This descriptive, cross-sectional, correlational study was conducted with second-year students enrolled in the Surgical Diseases Nursing course in the 2024–2025 academic year at a university in the Black Sea Region of Turkey (n=113). Based on Cohen's medium effect size, the minimum sample was 54, and 89 students participated. Role-playing and flipped learning were applied within the course. The data were collected using the Information Form, the Instructional Materials Motivation Scale, and the Korzul Clinical Practice Attitude Scale, and the level of statistical significance was set at $p < 0.05$.

Results: The mean age of participants was 20.96±2.66, and 74.2% were female (n=66). Students' motivation toward teaching materials was moderate (51.52±10.83), while attitudes toward clinical practice were high (84.57±12.88). A positive, moderate, and significant correlation was found between motivation and clinical practice attitude scores (rho=0.446, p<0.001).

Conclusion: The use of role-playing and flipped learning methods in nursing education enhances students' motivation and fosters more positive attitudes toward clinical practice. These methods should be encouraged among both academics and students.

Keywords: Nursing, clinical practice, motivation, student, role playing, flipped learning, attitude

ÖZ

Amaç: Hemşirelik eğitimi teorik ve uygulamalı boyutlarıyla bütüncül bir süreçtir. Öğrencilerin bu süreçte başarılı olabilmeleri için öğrenmeye yönelik motivasyonları ve klinik uygulamaya dair tutumları önemlidir. Bu araştırma, rol oynama ve ters yüz öğrenme yöntemi ile yapılandırılmış cerrahi hemşireliği dersinin hemşirelik öğrencilerinin motivasyonu ve klinik uygulama tutumlarına etkisini belirlemek amacıyla yapılmıştır.

Gereç ve Yöntemler: Tanımlayıcı, kesitsel ve ilişki arayıcı desende yürütülen araştırmanın evrenini, Türkiye'nin Karadeniz Bölgesi'nde bir üniversitenin Hemşirelik Bölümü'nde 2024-2025 akademik yılında Cerrahi Hastalıklar Hemşireliği dersini alan ikinci sınıf öğrenciler oluşturmuştur (n=113). Cohen'in orta etki büyüklüğüne göre en az 54 öğrenci ile tamamlanması planlanmış, araştırmaya 89 öğrenci katılmıştır. Ders kapsamında rol oynama ve ters yüz öğrenme yöntemleri kullanılmıştır. Veriler bilgi formu, Öğretim Materyali Motivasyon Ölçeği ve Korzul Klinik Uygulama Tutum Ölçeği ile toplanmış olup istatistikel anlamlılık düzeyi p<0,05 olarak kabul edilmiştir.

Bulgular: Öğrencilerin yaş ortalaması 20,96±2,66 olup, %74,2'si kadındır (n=66). Öğrencilerin öğretim materyali motivasyonları orta (51,52±10,83), klinik uygulama tutumları ise yüksek (84,57±12,88) düzeyde bulunmuştur. Ayrıca motivasyon ve klinik tutum puanları arasında pozitif yönde orta düzeyde anlamlı ilişki saptanmıştır (rho=0,446, p<0,001).

Sonuç: Rol oynama ve ters yüz öğrenme yöntemleri öğrencilerin motivasyonlarını artırmakta ve klinik uygulamalara yönelik olumlu tutum geliştirmelerine katkı sağlamaktadır. Bu yöntemlerin hem akademisyenler hem de öğrenciler tarafından yaygın olarak kullanılması önerilmektedir.

Anahtar Kelimeler: Hemşirelik, klinik uygulama, motivasyon, öğrenci, rol oynama, ters yüz öğrenme, tutum

Corresponding Author: Sibel Altıntaş ✉ sibel.alt@hotmail.com **Received:** 02.10.2025 **Revision:** 04.11.2025 **Accepted:** 07.01.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Role-playing and flipped learning methods enhance nursing students' motivation.
- These methods also positively influence students' attitudes toward clinical practice.
- Nursing students prefer that role-playing and flipped learning methods be used in other courses as well.

INTRODUCTION

Nursing education is continuously updated to meet the increasing demand for care services and the changing needs of society (1,2). Today, nursing education plays a key role in training nurses who can adapt to changes required by healthcare services, are open to self-improvement, can follow innovations, and provide quality care (2,3). The World Health Organization (WHO) recommends the use of innovative, interactive, and technological teaching methods in nursing education to enhance problem-solving, empathy, communication, clinical decision-making, and critical thinking skills for professionalization (4). Accordingly, nursing educators need to adopt multiple teaching strategies that support active learning and incorporate technology (5).

With the advancement of technology, nursing education has undergone a significant transformation, and the need for the use of information technologies in the educational process has been increasingly emphasized (6). Rapid developments in science and technology, the growing number of students, and limited educational resources have made it necessary to use innovative teaching methods in education (7). Notably, the Covid-19 pandemic has promoted a global shift toward blended learning to meet the theoretical and practical learning needs of nursing students (8).

The flipped learning method, an approach based on active learning, encourages students to take an active role in the learning process and combines the strengths of traditional and online learning (7,9). In this method, students can access materials such as videos, presentations, and audio recordings via online platforms using computers, tablets, or smartphones before the class, allowing them to review the content at their own learning pace (10). Thus, students participate in classes better prepared for new topics; moreover, those who cannot attend class do not face disadvantages, as they can keep up with the learning process of their peers (11).

Another active learning method, role-playing, is particularly important in nursing education, where the aim is to develop complex skills (12). This method involves students taking on specific roles and enacting scenario-based clinical situations, allowing theoretical knowledge to be reinforced through practical applications. Role-playing contributes to the development of students' communication, empathy, problem-solving, and active listening skills (13,14), thereby supporting their professional development in the nursing profession, which places a strong emphasis on human relationships (15). Indeed, a study involving bibliometric analysis revealed that the role-playing method has been increasingly recognized in nursing education and in the development of clinical skills (14).

In the literature, there are studies examining the effects of role-playing and flipped learning methods on students' skill acquisition, satisfaction levels, and communication skills (16-19). However, it is noteworthy that studies combining different teaching strategies and investigating their effects on students' motivation and attitudes toward clinical practice are limited. In this context, the present study aimed to examine the effect of a surgical nursing course structured with role-playing and flipped learning methods on students' motivation and attitudes toward clinical practice.

MATERIAL and METHODS

Type of the Study

The study was conducted in a descriptive, cross-sectional, and correlational design.

Population and Sample of the Study

The accessible population of the study consisted of second-year nursing students enrolled in the Surgical Nursing course at a state university located in the Western Black Sea Region of Türkiye during the 2024–2025 academic year (n=113). According to the power analysis calculated using G*Power 3.1.9.7 software, with a power of 95%, a 0.05 margin of error, and a medium effect size of 0.50 (Cohen, 1988), it was determined that a minimum of 54 students should participate in the study. The study was completed with 89 students who met the inclusion criteria (being enrolled in the Surgical Nursing course and taking it for the first time, volunteering to participate, not having worked or currently working as a nurse, fully attending classes conducted with the role-playing method, and having watched all flipped learning videos before the topic). The dependent variables of the study were students' motivation levels and attitudes toward clinical practice, while the independent variables were the students' demographic characteristics and the teaching methods used.

Data Collection Tools

In the study, the Student Information Form, Learning Material Motivation Scale, and Korzul Clinical Practice Attitude Scale were used.

Student Information Form: The information form was developed by the researcher based on the literature and consisted of 10 questions related to students' age, gender, and items evaluating role-playing and flipped learning methods (20-22).

Instructional Material Motivation Scale (IMMS): The scale was developed in Turkish by Tuğtekin and Dursun in 2022 to evaluate the motivation created by instructional materials on participants (23). The scale comprises 14 items under a single dimension and is scored on a 5-point Likert-type scale. There are no reverse items in the scale. Higher scores from the total or mean indicate higher positive motivation toward the evaluated material. The Cronbach's alpha reliability coefficient of the original scale was 0.95, while it was found to be 0.97 in this study.

Korzul Clinical Practice Attitude Scale (KCPAS): The scale was developed by Kanbay et al. in 2022 to determine the attitudes of students in health-related fields toward clinical practice. The scale consists of 25 items, including 12 reverse items, and has four subdimensions (adoption, adaptation, care, and clinical preference). The total score ranges between 25 and 125 (24). The Cronbach's alpha reliability coefficient of the original scale was 0.90, and it was 0.87 in this study.

Data Collection

During the spring semester of the 2024–2025 academic year, the teaching methods were explained to the students in detail during the first week of the Surgical Nursing course. All students were included in the flipped learning and role-playing methods during the semester. Course topics were presented by the instructor in accordance with the flipped learning method at least one week before the classroom session, and the lessons were recorded as videos (both verbal and visual). The recorded videos were uploaded to the university's distance education platform and made accessible to the students. Announcements were made through the same platform, asking students to watch the videos before class. Students had access to the videos throughout the semester. However, before the theoretical explanation of each topic in the classroom, students were required to watch the videos, and their viewing status was monitored by the instructor, who provided feedback during class sessions.

For the role-playing method, students were divided into groups of approximately 10, and each group was assigned a scenario preparation and classroom role-playing task based on a specific surgical system. The instructor provided guidance during the scenario development and performance stages. Following the course schedule, after the topic was presented by the instructor in class, students spent approximately four weeks preparing and then presented their role-playing performances to their peers in class. All role-playing activities were completed during the academic semester. Only the students who had watched all the videos before class and actively participated in role-playing activities were voluntarily included in the study.

The Student Information Form, IMMS, and KCPAS were prepared via Google Forms, and the survey link/QR code was shared with students through WhatsApp. Data collection was completed at the end of the semester with responses from 89 students.

Data Analysis

Data were analyzed using SPSS version 27.0. Descriptive statistics such as mean, standard deviation, frequency, and percentage were used to summarize the demographic characteristics of the students. The Shapiro-Wilk test was used to assess data normality, and since the data were not normally distributed, the Mann-Whitney U test and Spearman correlation test were applied. The level of statistical significance was set at $p<0.05$.

Ethical Considerations

Before starting the study, ethical approval was obtained from the Ethics Committee of Social and Human Sciences (Decision No: 2025-SBB-0673), and institutional permission was granted. Students were informed about the study in the classroom, and it was explained that the survey link/QR code would be shared with them at the end of the semester. At the beginning of the questionnaire, an explanatory note stated that no identifying information would be collected and that all data would be kept confidential. Students were asked to answer the question "Do you voluntarily agree to participate in this research?" with "yes" or "no." Only students who responded "yes" were allowed to complete the questionnaire. In addition, it was explained to the students that their course performance evaluation would not be affected if they chose not to participate. To prevent multiple submissions, participants were required to log into their Google accounts, but session data were anonymized in the settings. To avoid data loss, all survey questions were made mandatory. Permission to use the scale was obtained from the original authors.

RESULTS

The mean age of the students was 20.96±2.66 years, and 74.2% were female. Among the participants, 76.4% stated that they had chosen the department willingly, and 89.9% believed that theoretical courses affect clinical practice. Of the students, 62.9% reported that the role-playing method and 74.2% that the flipped learning method had been used

in previous courses; moreover, 61.8% wanted the role-playing method and 74.2% wanted the flipped learning method to be used in other courses as well. In addition, the mean satisfaction score for the role-playing method was 6.02±2.41 on a 0-10 scale, while the mean satisfaction score for the flipped learning method was 6.92±2.37. The mean score for the contribution of the role-playing method to course learning was 6.35±2.44, and for the flipped learning method, it was 7.03±2.30 (Table 1).

Table 2 presents the mean scores of the scales and their subdimensions. The total mean score of the IMMS was 51.52±10.23. The total mean score of the KCPAS was 84.57±12.88, with subdimension mean scores as follows: adoption=21.39±3.26, adaptation=30.69±5.49, care=13.26±2.29, and clinical preference=19.21±4.49 (Table 2).

Table 3 shows the relationships between student variables and scale scores. In the IMMS, scores were significantly higher among female students, those who had chosen the department willingly, and those who wanted the role-playing method to be used in other courses ($p<0.05$). In the KCPAS, female students had higher total scores and higher scores in the care and clinical preference subdimensions; students who had chosen the department willingly had higher scores in the adoption subdimension; and those who believed that the theoretical course affected clinical practice and who had not previously taken a course using the role-playing method had higher total scores and higher scores in the care and clinical preference subdimensions ($p<0.05$). Additionally, students who wanted the flipped learning method to be used in other courses had significantly higher total, adaptation, and clinical preference subdimension scores ($p<0.05$).

When findings regarding satisfaction and perceived learning contribution of the role-playing and flipped learning methods were examined, students who had not previously taken a course using the role-playing method had significantly higher satisfaction scores for role-playing and higher perceived learning contribution scores for the flipped learning method ($p<0.05$). Similarly, students who had not previously taken a course using the flipped learning method had significantly higher satisfaction and perceived learning contribution scores for the role-playing method ($p<0.05$). Moreover, students who wanted both methods to be used in other courses had significantly higher satisfaction and learning contribution scores for both methods ($p<0.05$).

Table 1: Socio-Demographic Characteristics of Students (n=89)

Characteristics		Mean	SD
Age		20.96	2.66
		n	%
Gender	Female	66	74.2
	Male	23	25.8
Choosing the program willingly	Yes	68	76.4
	No	21	23.6
Belief that theoretical course affects clinical practice	Yes	80	89.9
	No	9	10.1
Previous use of role-play method in the course	Yes	56	62.9
	No	33	37.1
Previous use of flipped classroom method in another course	Yes	66	74.2
	No	23	25.8
Desire to use role-play method in other courses	Yes	55	61.8
	No	34	38.2
Desire to use flipped classroom method in other courses	Yes	66	74.2
	No	23	25.8
Satisfaction with role-play method (0-10)		6.02	2.41
Contribution of role-play method to learning (0-10)		6.35	2.44
Satisfaction with flipped classroom method (0-10)		6.92	2.37
Contribution of flipped classroom method to learning (0-10)		7.03	2.30

n: Sample size, **Mean:** Average, **SD:** Standard Deviation

Table 2: Mean Scores of the Instructional Material Motivation Scale and Korzul Clinical Practice Attitude Scale, Including Subdimensions

Scales/Subdimensions		Mean	SD	Observed Min-Max	Scale Min-Max
	IMMS total score	51.52	10.83	14-70	14-70
	KCPAS total score	84.57	12.88	61-115	25-125
KCPAS Subdimensions	Adaptation Subdimension	21.39	3.26	13-30	6-30
	Adaptation Subdimension	30.69	5.49	21-45	9-45
	Care Subdimension	13.26	2.29	7-19	4-20
	Clinical Preference Subdimension	19.21	4.49	6-29	6-30

Mean: Average, **SD:** Standard Deviation, **Min:** Minimum Value, **Max:** Maximum Value, **IMMS:** Instructional Material Motivation Scale, **KCPAS:** Korzul Clinical Practice Attitude Scale

Table 3: Relationship Between Demographic Variables and Total/Subdimension Scores of the Scales

Scales	Instructional Material Motivation Scale (IMMS)	Korzul Clinical Practice Attitude Scale (KCPAS)					Satisfaction with Role-Playing Method	Contribution of Role-Playing Method to Learning	Satisfaction with Flipped Classroom Method	Contribution of Flipped Classroom Method to Learning
Variables	**Total Score**	**Adoption Subdimension**	**Adaptation Subdimension**	**Care Subdimension**	**Clinical Preference Subdimension**	**Total Score**				
	Mean±SD	**Mean±SD**	**Mean±SD**	**Mean±SD**	**Mean±SD**	**Mean±SD**	**Mean±SD**	**Mean±SD**	**Mean±SD**	**Mean±SD**
Gender										
Female	52.43±9.67	21.24±3.34	31.24±5.40	13.60±2.23	19.83±4.19	85.92±12.64	6.03±2.35	6.27±2.35	7.07±2.32	7.16±2.30
Male	48.91±13.52	21.82±3.05	29.13±5.58	12.30±2.22	17.43±4.91	80.69±13.05	6.00±2.62	6.60±2.72	6.47±2.50	6.65±2.32
p	**0.27**	0.56	0.09	**0.03**	**0.02**	**0.02**	0.99	0.56	0.32	0.27
Choosing the Department Willingly										
Yes	52.38±11.41	21.91±3.11	30.85±5.57	13.30±2.45	19.35±4.32	85.42±12.98	6.22±2.38	6.50±2.44	6.94±2.39	7.05±2.30
No	48.76±8.30	19.71±3.25	30.19±5.34	13.14±1.68	18.76±5.08	81.80±12.45	5.38±2.43	5.90±2.44	6.85±2.37	6.95±2.37
p	**0.02**	**0.00**	0.56	0.68	0.69	0.21	0.16	0.91	0.83	0.91
Thinking the Theoretical Course Affects Clinical Practice										
Yes	52.18±10.26	21.25±3.22	31.05±5.43	13.45±2.27	19.62±4.11	85.37±12.55	6.07±2.36	6.40±2.42	6.98±2.33	7.07±2.27
No	45.66±14.36	22.66±3.53	27.55±5.38	11.66±1.87	11.55±6.18	77.44±14.32	5.55±2.87	6.00±2.73	6.33±2.82	6.66±2.69
p	0.10	0.27	0.08	**0.01**	**0.01**	**0.03**	0.74	0.64	0.53	0.64
Previous Use of Role-Playing Method in Other Course										
Yes	50.53±11.69	21.10±3.28	29.51±5.12	12.94±2.22	18.51±4.38	82.08±12.26	5.64±2.46	6.05±2,52	6.60±2.44	6.71±2.33
No	53.21±9.09	21.87±3.22	32.69±5.60	13.81±2.32	20.39±4.49	88.78±12.99	6.66±2.20	6.87±2.24	7.45±2.18	7.57±2.19
p	0.42	0.28	0.00	**0.05**	**0.02**	**0.00**	**0.05**	0.08	0.08	**0.04**
Previous Use of Flipped Classroom in Other Course										
Yes	50.91±12.50	21.00±2.77	29.56±4.25	12.56±1.70	18.39±3.28	81.52±9.39	5.04±2.86	5.13±2.92	6.26±2.64	6.30±2.68
No	51.74±10.28	21.53±3.42	31.09±5.84	13.51±2.42	19.50±4.83	85.63±13.79	6.36±2.15	6.68±2.11	7.15±2.24	7.28±2.12
p	0.42	0.38	0.40	0.06	0.21	0.29	**0.04**	**0.00**	0.15	0.11
Willingness to Use Role-Playing Method in Other Courses										
Yes	55.58±7.82	21.82±3.62	30.55±5.87	13.64±2.24	20.05±4.19	86.08±14.16	7.00±2.05	7.29±1.86	7.02±2.19	7.20±2.02
No	49.01±11.70	21.12±3.03	30.78±5.30	13.03±2.30	18.69±4.62	83.63±12.06	5.41±2.43	5.78±2.59	6.85±2.49	6.92±2.47
p	**0.00**	0.41	0.58	0.25	0.31	0.54	**0.00**	**0.01**	0.93	0.88
Willingness to Use Flipped Classroom Method in Other Courses										
Yes	53.12±9.48	21.60±3.19	31.50±5.76	13.46±2.24	19.92±4.02	86.50±13.12	6.13±2.39	6.45±2.41	7.22±2.26	7.31±2.19
No	46.95±13.17	20.78±3.46	28.39±3.91	12.69±2.38	17.17±5.19	79.04±10.56	5.69±2.47	6.08±2.57	6.04±2.51	6.21±2.48
p	0.06	0.28	**0.05**	0.09	**0.03**	**0.01**	0.57	0.51	**0.05**	**0.05**

SD: Standard Deviation, **U/p:** Mann-Whitney U test/p- value, **IMMS:** Instructional Material Motivation Scale, **KCPAS:** Korzul Clinical Practice Attitude Scale

A positive, moderate, and statistically significant correlation was found between the total scores of the IMMS and the KCPAS (rho=0.446, p<0.001). In addition, significant correlations were determined among all subdimensions of the KCPAS (Table 4).

DISCUSSION

In this study, the effect of a surgical nursing course structured with role-playing and flipped learning methods on students' motivation and attitudes toward clinical practice was examined. The students' instructional material motivation was found to be moderate (51.52±10.83), and their KCPAS scores were found to be high (84.57±12.88). There are also studies in the literature showing that different teaching methods increase students' motivation. In the study conducted by Balkaya et al., it was found that the use of the cooperative learning method in the course increased students' motivation and self-confidence. In this study, 68.4% of the participating students stated that this method increased their self-confidence, 59.7% reported that it improved their academic achievement, and 67.4% stated that it made their knowledge more permanent (25). Molu and Keskin conducted a study with 72 nursing students taking a pharmacology course, using bingo and board games. As a result of the study, the total scores of the IMMS were high in both groups. It was found that 77.8% of the students in the bingo group played an educational game in class, 100% liked the bingo game, 97.2% thought that the game contributed to their learning, and they wanted to use game materials in other courses. In the board game group, 91.7% of the students played an educational game in class, 94.4% liked the board game, 91.7% thought that the game contributed to their learning, and 94.4% stated that they wanted to use game materials in other courses (26). In our study, the instructional material motivation scores of students who wanted to use the role-playing method in another course were found to be higher. In addition, the satisfaction scores and perceived contribution to learning of students who wanted to use the role-playing method again, as well as those who wanted to use the flipped learning method in another course, were significantly higher (p<0.05). In the study conducted by Ongor and Ulusoy, the effect of multimedia teaching materials on students' learning motivation in e-learning was examined. In the course conducted via e-learning, various WEB 2.0 tools were used in the experimental group. As a result of the study, the instructional materials motivation mean score of the experimental group was found to be higher than that of the control group (27). These results, parallel to our study, indicate that different teaching methods and web-supported instructional materials increase students' satisfaction and motivation.

In the study conducted by Erbuğa and Karakurt, the Clinical Practice Attitude Scale score of nursing students was similar to our study, found to be 88.98±16.33, and it was determined that the clinical practice attitudes of students who chose the department willingly were higher than those who did not (28). In our study, the Care and Clinical Preferences subscale scores were significantly higher among female students, and the Adoption subscale scores were higher among students who chose the department willingly. Additionally, the total scale score and Care and Clinical Preferences subscale scores were significantly higher

Table 4: Examination of the Relationship Between Total and Subdimension Scale Scores

Scale/Subdimensions			IMMS total score	KCPAS				
				Total score	Adoption Subdimension	Adaptation Subdimension	Care Subdimension	Clinical Preference Subdimension
IMMS total score		rho	1.000					
		p	-					
KCPAS total score		rho	0.446**	1.000				
		p	0.001	-				
KCPAS Subdimensions	Adoption Subdimension	rho	0.438**	0.578**	1.000			
		p	0.001	0.001	-			
	Adaptation Subdimension	rho	0.360**	0.878**	0.450**	1.000		
		p	0.001	0.001	0.001	-		
	Care Subdimension	rho	0.358**	0.830**	0.444**	0.667**	1.000	
		p	0.001	0.001	0.001	0.001	-	
	Clinical Preference Subdimension	rho	0.313**	0.834**	0.285**	0.714**	0.687**	1.000
		p	0.003	0.001	0.007	0.001	0.001	-

rho: Spearman's rank correlation coefficient; **p:** significance level; **IMMS:** Instructional Material Motivation Scale; **KCPAS:** Korzul Clinical Practice Attitude Scale. Significance: **p<0.05; ** indicates statistical significance.

among students who thought that the theoretical course affected the clinical course. According to the findings, students' willingness to choose the department affects their clinical practice attitudes. Therefore, it is important to diversify teaching methods to support students' willingness and motivation. Indeed, in our study, students stated at a high rate that they wanted the role-playing and flipped learning methods to be used in other courses as well. Moreover, the satisfaction and perceived learning contribution scores of these students were also found to be significantly higher.

In a scoping review including 17 studies, the effects of educational technologies used with nursing students on their engagement and motivation were examined. It was found that the most common innovative educational technology was gamification, and students reported increased motivation after gamification (29). In a study evaluating the effectiveness of the role-playing method among nursing students, the mean skill scores of the group using the role-playing method were found to be higher than those of the group that did not use the method (30). In a mixed-methods study evaluating the effectiveness of the role-playing method, it was revealed that the use of this method in teaching significantly increased students' self-esteem and satisfaction levels compared to traditional teaching methods (31). In a study examining the effect of scenario-based triage education through lectures and role-playing on nursing students' knowledge and practices, it was shown that the role-playing method was more effective than the traditional method (32). In our study, the satisfaction levels and perceived learning contribution of students using the role-playing method, as well as the proportion of those who wanted the method to be used again, were high.

In our study, a moderately positive and statistically significant correlation was found between the total scores of the IMMS and the KCPAS (r_s=0.446, $p<0.001$). In the literature, it has been found that the motivation levels of students using the flipped learning method are higher than those using the traditional method (33). In another study comparing nursing students in traditional and flipped classrooms, it was found that the motivation scores of the experimental group (flipped classroom) were higher than those of the control group (traditional classroom) (34). Park and Suh found a significant relationship between critical thinking, self-regulated learning, and learning satisfaction among nursing students educated with the flipped learning method (35). In the study of Li et al., it was reported that the academic achievement levels of nursing students increased through the flipped learning method (36). Lee and Park also found that the flipped learning method improved nursing students' self-management and problem-solving skills (37). It is crucial for nursing students to develop these skills before beginning their professional careers (38). Hence, promoting and disseminating such research is of considerable importance. When the literature is reviewed, consistent with our research findings, it is observed that the flipped learning method positively affects students' academic performance, analytical thinking, self-directed learning, and problem-solving skills.

Limitations

The absence of a control group in which the teaching methods were not applied limits the ability to draw definitive conclusions regarding the effects of the techniques used on motivation and clinical attitudes. In addition, only a post-test was administered to evaluate the effectiveness of the methods, and no pre-test/post-test comparison was conducted; this is also an important limitation of the study. Furthermore, the IMMS was administered by asking students to respond by considering both methods together. In future studies, it is recommended that, in research evaluating more than one method together, measurements for each method be conducted using separate scale applications.

Conclusion

This study revealed that the role-playing and flipped learning methods used in the surgical nursing course had positive effects on students' motivation and attitudes toward clinical practice. Students' instructional material motivation levels were found to be moderate, while their attitudes toward clinical practice were high; in particular, the satisfaction and learning contribution scores of students who wanted the role-playing and flipped learning methods to be used in other courses were found to be significantly higher. The findings are consistent with similar research in the literature and demonstrate that innovative teaching methods support students' motivation, engagement in learning, clinical practice attitudes, and academic achievement. Accordingly, it is recommended that teaching methods that enhance students' active participation in the learning process and strengthen their motivation and attitudes be diversified and utilized in nursing education.

Author Contributions

Study conception and design: **Sevim Celik, Elif Karahan;** data collection: **Ozge Ucar, Meryem Yucel;** analysis and interpretation of results: **Sibel Altıntas;** draft manuscript preparation: **Sevim Celik, Elif Karahan, Sibel Altıntas, Ozge Ucar, Meryem Yucel.** The authors reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Funding

The authors declare that the study received no funding.

Ethical Approval

The study was approved by the Bartın University Social and Human Sciences Ethics Committee (number: 2025-SBB-0673).

REFERENCES

1. Karasu F, Arapacı R, Öztürk E, Çam HH. The cooperative learning model in nursing education. İnönü University Journal of Vocational School of Health Services. 2020;8(3):973-988.
2. Uslusoy E, Ersoy S. The use of cooperative learning methods among nursing students in national literature: A systematic review. Süleyman Demirel University Journal of Health Sciences. 2022;13(2):201-211. https://doi.org/10.22312/sdusbed.1000222
3. Sarmasoğlu S, Dinç L, Elçin M. Nursing students' opinions about standardized patients and mannequins used in clinical skills training. Journal of Nursing Education and Research. 2016;13(2):107-115. https://doi.org/10.5222/HEAD.2016.107
4. World Health Organization (WHO). Nurse Educator Core Competencies. [Internet]. 2016 [Accessed 18.09.2025]. Available from: https://iris.who.int/bitstream/handle/10665/258713/9789241549622-eng.pdf
5. Barranquero-Herbosa M, Abajas-Bustillo R, Ortego-Maté C. Effectiveness of flipped classroom in nursing education: A systematic review of systematic and integrative reviews. Int J Nurs Stud. 2022;135:104327. https://doi.org/10.1016/j.ijnurstu.2022.104327
6. Şenyuva, E. Reflections on nursing education of technological developments. FNJN. 2019;27(1):79-90. https://doi.org/10.26650/FNJN322556
7. Oğurlu N. The use of the flipped classroom teaching method in nursing education. Journal of Nursing Science. 2023;6(2):123-129.
8. JJowsey T, Foster G, Cooper-Ioelu P, Jacobs S. Blended learning via distance in pre-registration nursing education: A scoping review. Nurse Educ Pract. 2020;44:102775. https://doi.org/10.1016/j.nepr.2020.102775
9. Talbert R, Mor-Avi A. A space for learning: An analysis of research on active learning spaces. Heliyon. 2019;5(12):e02967. https://doi.org/10.1016/j.heliyon.2019.e02967
10. Tan, C., Yue, W.-G., Fu, Y. Effectiveness of flipped classrooms in nursing education: Systematic review and meta-analysis. Chinese Nursing Research. 2017;4(4):192-200. https://doi.org/10.1016/j.cnre.2017.10.006
11. Hu R, Gao H, Ye Y, Ni Z, Jiang N, Jiang X. Effectiveness of flipped classrooms in Chinese baccalaureate nursing education: A meta-analysis of randomized controlled trials. Int J Nurs Stud. 2018;79:94-103. https://doi.org/10.1016/j.ijnurstu.2017.11.012
12. Azizah, S. Application of the role playing method to increase students' learning motivation. Sociale: International Journal of Social Studies. 2024;1(1):97-121. https://doi.org/10.19105/sociale.v1i1.12275
13. Cortés-Rodríguez AE, Roman P, López-Rodríguez MM, Fernández-Medina IM, Fernández-Sola C, Hernández-Padilla JM. Role-Play versus Standardised Patient Simulation for Teaching Interprofessional Communication in Care of the Elderly for Nursing Students. Healthcare (Basel). 2021;10(1):46. https://doi.org/10.3390/healthcare10010046
14. Terzi A, Yildirim Y, Kocakoc ID. The power of role play in nursing education: A bibliometric analysis and visualization. Nurse Education Today. 2025;149:106633. https://doi.org/10.1016/j.nedt.2025.106633
15. Novaes LMS, Paiva EMDC, O'Mahony A, Garcia ACM. Role-play as an Educational Strategy in Palliative Care: A Systematic Integrative Review. Am J Hosp Palliat Care. 2022;39(5):570-580. https://doi.org/10.1177/10499091211036703
16. Cansever Z, Avsar Z, Tastan K. Third Year Medical School Students' Experiences of Revealing Patients' Stories through Role Playing. Eurasian J Med. 2015;47(1):26-31. https://doi.org/10.5152/eajm.2014.53
17. Evans L, Vanden Bosch ML, Harrington S, Schoofs N, Coviak C. Flipping the Classroom in Health Care Higher Education: A Systematic Review. Nurse Educ. 2019;44(2):74-78. https://doi.org/10.1097/NNE.0000000000000554
18. Oliver R, Luther L. Flipping the Graduate Nursing Classroom: An Integrative Review. J Nurs Educ. 2020;59(6):305-310. https://doi.org/10.3928/01484834-20200520-02
19. Valizadeh L, Akbarzadeh B, Ghiyasvandiyan S, KuchakiNejad Z, Zamanzadeh V, Aghajari P, et al. The effects of role play simulation and demonstration on pediatric peripheral venous catheter insertion skill among nursing students: A three-group experimental study. Nursing and Midwifery Studies. 2021;10(1):1-6. https://doi.org/10.4103/nms.nms_94_18
20. Irmak AY, Metinoğlu M. The empathic tendency level of nursing students participating and not participating in case-based role play activities. Journal of Youth Research. 2017;5(12):153-167.
21. Akın S, Çakto P. Examination of the effect of instructional material on university students' motivation levels. Journal of Higher Education and Science. 2023;13(2):188-194. https://doi.org/10.5961/higheredusci.1234691
22. Acar MD, Altun EÇ. Investigation of scenario-based nursing teaching course in terms of students' motivation for learning materials. TOGÜ Journal of Health Sciences. 2025;4(1):36-49. https://doi.org/10.52369/togusagbilderg.1253512
23. Barut Tugtekin E, Dursun OO. Effect of animated and interactive video variations on learners' motivation in distance Education. Educ Inf Technol (Dordr). 2022;27(3):3247-3276. https://doi.org/10.1007/s10639-021-10735-5
24. Kanbay Y, Tektaş P, Işık E, Aslan Ö, Akçam A, Çınar S. Korzul Clinical Practice Attitude Scale validity and reliability study. AUHSJ. 2022;13(1):157-166. https://doi.org/10.31067/acusaglik.974347
25. Balkaya NA, Bilge Ç, Dönmez, S. The effect of cooperative learning and interactive teaching methods on nursing students' motivation and self-confidence. KASHED. 2022;8(1):14-24.
26. Molu B, Keskin AY. The effect of bingo games and board games applied to nursing students in pharmacology lessons on lesson motivation. BAUN Journal of Health Sciences. 2024;13(2):294-300. https://doi.org/10.53424/balikesirsbd.1355623
27. Ongor M, Uslusoy EC. The effect of multimedia-based education in e-learning on nursing students' academic success and motivation: A randomised controlled study. Nurse Educ Pract. 2023;71:103686. https://doi.org/10.1016/j.nepr.2023.103686
28. Erbuğa E, Karakurt P. Examination of nursing students' attitudes toward the nursing profession and clinical practice. EBYÜSBED. 2024;1(1):1-22.

29. Kowitlawakul Y, Tan JJM, Suebnukarn S, Nguyen HD, Poo DCC, Chai J, Wang W, Devi K. Utilizing educational technology in enhancing undergraduate nursing students' engagement and motivation: A scoping review. J Prof Nurs. 2022;42:262-275. https://doi.org/10.1016/j.profnurs.2022.07.015

30. Paul C, Bhokare N, Pathak S. Investigate the effect of role play on the skill among 2nd Basic B.Sc. students of SRMMCON regarding selected community procedure. Indian Journal of Forensic Medicine & Toxicology. 2020;14(4).

31. Mamaghani EA, Hosseinian E, Maghsoodi E. Role playing is an effective method for training physical examinations: A mixed-methods study. Teaching and Learning in Nursing. 2024;19(4):e672-e677. https://doi.org/10.1016/j.teln.2024.06.005

32. Delnavaz S, Hassankhani H, Roshangar F, Dadashzadeh A, Sarbakhsh P, Ghafourifard M, et al. Comparison of scenario based triage education by lecture and role playing on knowledge and practice of nursing students. Nurse Educ Today. 2018;70:54-59. https://doi.org/10.1016/j.nedt.2018.08.006

33. Naciri A, Hajji ME, Radid M, Kharbach A, Chemsi G. Exploring student motivation and performance in the flipped classroom: A case study of nursing students. Electronic Journal of General Medicine. 2022;19(3). https://doi.org/10.29333/ejgm/11796

34. Aksoy B, Pasli Gurdogan E. Examining effects of the flipped classroom approach on motivation, learning strategies, urinary system knowledge, and urinary catheterization skills of first-year nursing students. Jpn J Nurs Sci. 2022;19(2):e12469. https://doi.org/10.1111/jjns.12469

35. Park I, Suh Y. Meta-Analysis of Flipped Learning Effects in Nursing Education. Int J Environ Res Public Health. 2021;18(23):12814. https://doi.org/10.3390/ijerph182312814

36. Li BZ, Cao NW, Ren CX, Chu XJ, Zhou HY, Guo B. Flipped classroom improves nursing students' theoretical learning in China: A meta-analysis. PLoS One. 2020;15(8):e0237926. https://doi.org/10.1371/journal.pone.0237926

37. Lee MK, Park BK. Effects of Flipped Learning Using Online Materials in a Surgical Nursing Practicum: A Pilot Stratified Group-Randomized Trial. Healthc Inform Res. 2018;24(1):69-78. https://doi.org/10.4258/hir.2018.24.1.69

38. Kaş Güner C, Kutlutürkan S. The nurse's role in preventing ventilator-associated pneumonia. Med J West Black Sea. 2021;5(3):324-330. https://doi.org/10.29058/mjwbs.936579

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):106-115
DOI: 10.29058/mjwbs.1656011

Postoperative frequency of normocalcemic elevated parathormone in patients surgically treated for primary hyperparathyroidism: A single-center experience

Primer hiperparatiroidizm sebebiyle cerrahi tedavi uygulanan hastalarda ameliyat sonrası normokalsemik yüksek parathormon sıklığı: Tek merkez deneyimi

Gülşah Elbüken[1], Ogün İrem Bilen[1], Beyza Olcay Öztürk[1], Hünkar Ağgül[1], Sibel Özkan-Gürdal[2], Burcu Tokuç[3], Sayid Shafi Zuhur[1]

[1]Department of Internal Medicine, Tekirdağ Namık Kemal University Faculty of Medicine, Division of Endocrinology and Metabolism, Tekirdağ, Türkiye

[2]Department of General Surgery, Tekirdağ Namık Kemal University Faculty of Medicine, Tekirdağ, Türkiye

[3]Department of Medical Education, Tekirdağ Namık Kemal University Faculty of Medicine, Tekirdağ, Türkiye

Cite this article as: *Elbüken G et al. Postoperative frequency of normocalcemic elevated parathormone in patients surgically treated for primary hyperparathyroidism: A single-center experience. Med J West Black Sea. 2026;10(1):106-115.*

This study was partially presented as a poster at the 43rd Turkish Endocrine and Metabolic Disorders Congress on May 18-22, 2022 in Antalya, Turkey.

ABSTRACT

Aim: After parathyroidectomy, parathormone (PTH) levels may remain elevated despite normalization of serum calcium levels in some patients, which is known as "postoperative normocalcemic elevated PTH". In this study, we aimed to determine postoperative normocalcemic elevated parathormone frequency in patients undergoing parathyroidectomy due to primary hyperparathyroidism through a retrospective review.

Material and Methods: Preoperative, intraoperative, early postoperative (maximum 7-days, starting from the next day of surgery until discharge) and late postoperative (months 1, 3, 6 and 12) laboratory tests of the 127 patients were reviewed retrospectively. Corrected calcium levels were obtained using serum albumin levels.

Results: Laboratory values consistent with postoperative normocalcemic elevated PTH were identified in 56%, 44%, 42.3% and 31.9% of the patients at 1, 3, 6 and 12 months postoperatively. There are missing data in the 127 patients who underwent surgery because some patients did not attend their follow-up visits or were followed up in other centers. Serum calcium values were available and within normal range for all 91 patients followed for a mean duration of 44.0±27.2 months (median= 36, min= 15, max=144). The frequency of postoperative normocalcemic elevated parathormone was 34.4% among patients with serum calcium and PTH measurements available from the last visit.

Conclusion: The prevalence of postoperative normocalcemic elevated parathormone varies among different patient series depending on the criteria used. Whether the underlying secondary hyperparathyroidism contributes to postoperative normocalcemic elevated parathormone should be investigated and managed appropriately. Considering that postoperative normocalcemic elevated PTH may also represent primary hyperparathyroidism recurrence, regular follow-up is required for patients.

Keywords: Parathormone, parathyroidectomy, postoperative normocalcemic elevated parathyroid hormone, primary hyperparathyroidism

ÖZ

Amaç: Primer hiperparatiroidi nedeniyle paratiroidektomi yapılan bazı hastalarda, serum kalsiyum düzeyleri normale dönmesine rağmen parathormon (PTH) düzeyleri yüksek kalabilir. Bu durum "postoperatif normokalsemik yüksek PTH" olarak tanımlanır. Bu çalışmada, primer hiperparatiroidi nedeniyle ameliyat edilen hastalarda postoperatif normokalsemik yüksek PTH sıklığı araştırılmıştır.

Gereç ve Yöntemler: Çalışmada 127 hastanın ameliyat öncesi, ameliyat esnasındaki, erken postoperatif (ameliyattan sonraki gün ile taburculuk dahil ilk 7 gün) ve geç postoperatif (1., 3., 6. ve 12. aylar) laboratuvar değerleri geriye dönük olarak incelendi. Düzeltilmiş kalsiyum, serum albüminine göre hesaplandı.

Bulgular: Postoperatif normokalsemik yüksek PTH ile uyumlu laboratuvar bulguları: 1. ayda %56, 3. ayda %44, 6. ayda %42.3, 12. ayda %31.9 olarak bulundu. Ortalama 44.0±27.2 ay (ortanca 36, min 15 – maks 144) takip edilen 91 hastanın tümünde serum kalsiyum düzeyleri normaldi. Son kontrolde hem serum kalsiyum hem PTH değeri bulunan hastalarda postoperatif normokalsemik yüksek PTH sıklığı %34.4 olarak saptandı.

Sonuç: Postoperatif normokalsemik yüksek PTH sıklığı, kullanılan kriterlere ve hasta serilerine göre değişebilir. Altta yatan sekonder hiperparatiroidinin katkısı araştırılmalı ve yönetilmelidir. Ayrıca bu bulgunun primer hiperparatiroidi nüksünü gösterebileceği unutulmamalıdır. Hastaların düzenli takibi gereklidir.

Anahtar Kelimeler: Parathormon, paratiroidektomi, postoperatif normokalsemik yüksek parathormon, primer hiperparatiroidi

Corresponding Author: Gülşah Elbüken ✉ gelbuken@nku.edu.tr

Received: 13.03.2025 **Revision:** 09.08.2025 **Accepted:** 02.02.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Postoperative normocalcemic elevated parathormone (pNePTH) is a frequent finding after parathyroidectomy for primary hyperparathyroidism.
- The prevalence of pNePTH was 56%, 44%, 42.3% and 31.9% at 1, 3, 6 and 12 months postoperatively, respectively.
- Despite sustained normocalcemia, the long-term frequency of pNePTH was 34.4% in patients with extended follow-up.
- Secondary hyperparathyroidism mechanisms, including renal function impairment and possible active vitamin D deficiency, may contribute to postoperative pNePTH.
- As pNePTH may represent an early marker of recurrent primary hyperparathyroidism, long-term and regular postoperative follow-up is essential.

INTRODUCTION

Primary hyperparathyroidism (pHPT) is an endocrine disorder characterized by hypercalcemia and elevated serum parathyroid hormone (PTH) levels. The incidence of pHPT increases after the age of 50 in both men and women. Females are 2 to 4 times more frequently affected than males (1). Oftentimes, the presentation of pHPT is asymptomatic in parts of the world where serum calcium (sCa) levels are routinely assessed. In addition to mild hypercalcemia, pHPT may present with hypercalciuria and osteoporosis and even bone fractures and nephrolithiasis, which can be asymptomatic (2).

Primary hyperparathyroidism is caused by a single parathyroid adenoma in 80-85%, parathyroid hyperplasia in %10-15, multiple parathyroid adenomas in 2-3% and parathyroid carcinoma in 1% of the cases (3). Although most pHPT cases are sporadic, 10% of the cases have a genetic origin and may manifest as either neonatal severe pHPT, familial hyperthyroidism-jaw tumor syndrome or as a component of Multiple Endocrine Neoplasia (MEN type 1, MEN type 2A, MEN type 4) (3,4).

The definitive treatment of pHPT is surgical removal of parathyroid adenoma or hyperplasic parathyroid glands (parathyroidectomy). Parathyroidectomy is recommended in patients with symptoms and in those with asymptomatic disease who are at risk of progression or with evidence of subclinical end-organ sequelae (5). In pHPT, an increase in bone mineral density (BMD) and a decrease in kidney stone disease are observed following parathyroidectomy. Various pharmacological therapies can reduce sCa levels or increase BMD in patients for whom parathyroidectomy is not planned (3).

Following parathyroid surgery, sCa and PTH levels return to normal range after removal of the parathyroid adenoma or hyperplastic parathyroid glands. However, sometimes sCa and PTH levels do not return to normal range. In such cases, surgery may be considered as unsuccessful. At other times, sCa may be normal but PTH may rise again postoperatively despite normalization of sCa and PTH levels during or immediately after surgery. This condition is known as “postoperative normocalcemic elevated PTH” (pNePTH). There have been attempts to explain this condition by “secondary hyperparathyroidism” (sHPT) which involves stimulation of PTH secretion due to a relative reduction in sCa levels postoperatively compared to the preoperative period, or due to sCa deficit associated with replacement of Ca lost from the bone (hungry bone syndrome) postoperatively, or low preoperative levels or inadequate replacement of vitamin D, or impairment of active vitamin D production due to renal insufficiency (in all 4 cases, PTH increase is designated as sHPT). These postoperative hormonal changes are considered to represent an “adaptive state” and may be transient. Some authors argued that this adaptive state may last up to 6 months and normalization of Ca and PTH levels at 6 months after parathyroid surgery indicates that the parathyroid surgery is successful (6,7).

A decrease in serum ionized calcium (iCa) concentration stimulates parathyroid glands to secrete more PTH by activating the calcium-sensing receptors (CaSR) on the parathyroid glands. This known relationship between extracellular iCa concentration and PTH secretion has been defined as an inverse sigmoidal curve (8). Studies have shown that oral Ca supplementation reduces PTH levels in patients with pNePTH. In addition, by replacing deficient vitamin D, elevated PTH levels resulting from sHPT can be brought back to normal (7,9). pNePTH is defined as normal sCa in the face of persistently elevated PTH level despite appropriate Ca and/or vitamin D replacement (9,10). The reported prevalence of pNePTH is approximately 23.5% (min= 3%, max: 46%), although it varies across studies and case series (10).

In this study, we sought to determine both our surgical success rate and the pNePTH frequency in patients undergoing parathyroidectomy due to pHPT through a retrospective review.

MATERIAL and METHODS

Approval for the study was obtained from the Ethics Committee of Tekirdağ Namık Kemal University Medical Faculty on 25.01.2022 (No. 2022.07.01.07) and conducted in accordance with the principles laid out in the Declaration of Helsinki.

Patients: The data of 148 patients who were operated for PHPT between January 1, 2010 and January 5, 2022 by a single physician (Dr. SOG) at our hospital were reviewed retrospectively. Since 3 of 148 patients were operated for tertiary HPT, these patients were excluded from the analyses. Ultimately, 145 patients who underwent curative parathyroidectomy for pHPT were included in the data analysis. For these patients, age, sex, body mass index (BMI) and available data on sCa, phosphorus (P), PTH, 25-OH vit D3, serum alkaline phosphatase (ALP), estimated GFR (eGFR), 24-hour urine Ca excretion from preoperative, intraoperative, early postoperative (maximum 7-days, starting from the next day of surgery until discharge) and late postoperative (months 1, 3 and 12) periods were retrieved. In addition, for the patients who were not followed postoperatively at our hospital, postoperative sCa, P, albumin, PTH and 25-OH vit D3 levels were reviewed on the e-nabiz (e-pulse) application implemented by the Turkish Ministry of Health. sCa was corrected for hypoalbuminemia using the following formula: Corrected calcium (cCa) = [0.8 × (4.0 – albumin) + serum calcium].

Surgical Definitions: The indications for parathyroidectomy were determined as recommended by the "Guidelines from the Fourth International Workshop on the Management of Asymptomatic Primary Hyperparathyroidism" (5). The intraoperative PTH criterion used to predict successful parathyroidectomy was defined as a 50% decrease in serum PTH level in a peripheral blood sample drawn 10 minutes after removal of parathyroid adenoma or hyperplastic parathyroid glands compared to the preoperative PTH level (11). Surgical failure was defined as elevated serum calcium and PTH levels obtained immediately after parathyroidectomy. The time period (maximum 7 days) starting from the first day after the surgery until discharge was defined as the "early postoperative period". Recurrence was defined as hypercalcemia after normocalcemia for up to 6 months after succesful surgery. pNePTH was defined as elevated serum PTH (65 pg/ml or higher) after 6 months of surgery, despite normalization of sCa level. All causes of sHPT (such as hungry bone syndrome, low levels or inadequate replacement of 25-OH vitamin D3, or impairment of active vitamin D production due to renal insufficiency) were excluded.

Statistical Analyses

SPSS v22.0 (IBM Corp., Armonk, NY, USA) software was used for data analysis. The normality of data distribution was checked using Shapiro-Wilk test. Demographic and clinical characteristics of the patients were examined using descriptive statistics including number, percentage, mean, standard deviation. Although repeated-measures ANOVA was initially planned for the analysis, this method could not be applied due to missing data at several postoperative follow-up visits, which resulted in an unbalanced dataset. Because the assumptions of repeated-measures ANOVA (including sphericity and homogeneity of variances) were not met under these conditions, an alternative approach was required. Therefore, an independent samples t-test was used to evaluate the mean values at the available time points. The significance level was set at $p<0.05$ for all analyses.

RESULTS

The data of 148 patients who underwent parathyroid surgery were retrieved from hospital records. Three of these patients were operated for tertiary hyperparathyroidism (tHPT) and excluded from the study. Therefore, the data of the remaining 145 patients including 127 females (88%) and 18 (12%) males were reviewed retrospectively (Figure 1). Among the patients, 23 (16%) were operated before 2016 and 122 (84%) after 2016, the year we moved to our new modern hospital building. 127 (87.6%) patients had a solitary parathyroid adenoma, 2 (1.4%) had two parathyroid adenomas and 16 (11%) had parathyroid hyperplasia. The mean age of 145 patients undergoing surgery for pHPT was 57.6±11.5 years (min=23, max=81). There was no significant age difference between females and males (58.1±11.2 years vs 54.3 ±14.1 years, p=0.285). Preoperative sCa, P, albumin and PTH values were available for all 145 patients. Preoperative mean values were 11.5±0.7 mg/dl for cCa, 2.5±0.5 mg/dl for P and 252±235 pg/ml for PTH. Preoperatively, 25-OH vitamin D3 data were available for 122 patients, serum ALP data for 120 patients and eGFR data for 135 patients, with mean values of 20.6±13.8 ng/ml, 121±72 IU/L and 95.5±221.1 ml/min/1.73 m^2 respectively (Table 1). Preoperative 24-hour urine Ca excretion data were available for 113 patients with a mean value of 398±191 mg/24 h (min=101, max=1149) (not shown in table). Preoperative BMI values were available for 127 patients with a mean value of 31.0±5.7 kg/m^2 (min=20.5, max=57) (not shown in table).

Intraoperatively, PTH level was measured in 132 (91%) patients, sCa level in 58 (40%) patients and P level in 48 (33%) patients. In the early postoperative period, sCa and albumin concentrations were measured in 143 (99%) patients, P concentration in 131 (90%) patients and PTH level in 138 (95%) patients. Preoperative, intraoperative and postoperative laboratory values for patients undergoing surgery are shown in Table 1. Among 132 patients with PTH measurement intraoperatively, 112 (84.9%) showed a

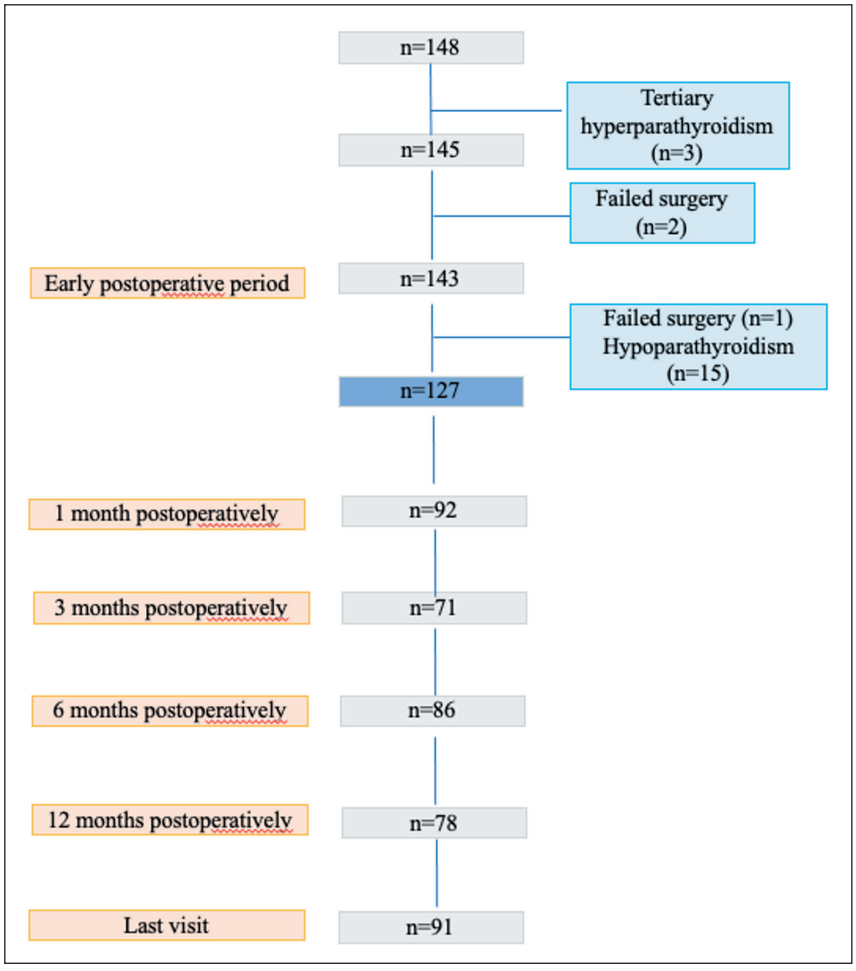


Figure 1: Evaluation of operated patients
n: number of patients with available data

Table 1: Preoperative, Intraoperative and Postoperative Laboratory Results

Parameter	Preoperatively	Intraoperatively	Early postoperative period	1 month postop	3 months postop	6 months postop	12 months / Last visit
cCa (mg/dl)	11.5 ± 0.7 (min=10.5, max=14.2) n=145	10.1 ± 0.8 (min=7.4, max=12) n=58	9.1 ± 0.8 (min=7.2, max=14.8) n=143	9.3 ± 0.6 (min=8.5, max=10.2) n=92	9.5 ± 0.4 (min=8.7, max=10.2) n=71	9.4 ± 0.4 (min=8.2, max=10.2) n=86	9.4 ± 0.4 (min=8.5, max=10.4) n=91
P (mg/dl)	2.5 ± 0.5 (min=1.2, max=3.9) n=145	2.7 ± 0.8 (min=1.4, max=4.7) n=48	3.2 ± 0.7 (min=1.3, max=4.8) n=131	3.2 ± 0.6 (min=1.4, max=5.1) n=76	3.2 ± 0.5 (min=1.7, max=4.3) n=66	3.2 ± 0.5 (min=2.1, max=5.1) n=78	3.4 ± 0.6 (min=2.0, max=4.7) n=78
PTH (pg/ml)	252 ± 535 (min=82, max=1667) n=145	46.4 ± 74.3 (min=12, max=530.3) n=132	35.7 ± 46.4 (min=2.1, max=269) n=138	89.8 ± 72.4 (min=17, max=316) n=79	83.5 ± 0 (min=15, max=307) n=71	67.3 ± 32.1 (min=17.1, max=316) n=78	65.9 ± 35.0 (min=18, max=212) n=62
25-OH vit D3 (ng/ml)	20.6 ± 13.8 (min=3, max=86.7) n=122	—	—	121 ± 89.8 n=75	91.6 ± 29.8 n=53	85.2 ± 21.8 n=49	77.6 ± 23.6 n=53
ALP (IU/L)	—	—	—	91.6 ± 29.8 (min=10, max=525) n=35	91.6 ± 29.8 (min=43, max=163) n=26	95 ± 24 (min=43, max=133) n=34	85.1 ± 19.4 (min=35, max=160) n=53
eGFR	95.5 ± 22.1 (min=37, max=140) n=135	—	93.8 ± 26.7 (min=40, max=135) n=58	—	—	95 ± 24 (min=40, max=136) n=55	85.1 ± 19.4 (min=42, max=121) n=42

reduction of at least 50% in PTH values. Calcium values showed a reduction of 45 (77.6%) in 58 patients with intraoperative sCa measurement. Normal cCa values were found in 76% of 58 patients with intraoperative sCa data available and 58.3% of 132 patients with intraoperative PTH measurement. During the early postoperative period, 110 (77%) patients had cCa values within normal range and 71 (51.5%) patients showed normal PTH values (Table 2).

The number of evaluable patients decreased to 143 after excluding 2 patients who were considered as surgical failure due to no reduction in intraoperative Ca and PTH values (Figure 1). Additionally, one patient with an early postoperative cCa value of 11.8 mg/dl was also considered as surgical failure and excluded from the study (n=142). Thus, during the time period from surgery until discharge, surgical failure was identified in a total of 3 (2%) patients out of 145 patients (Figure 1). Among the patients developing postoperative hypoparathyroidism and hypocalcemia, hypoparathyroidism and hypocalcemia persisted in 15 (10.35%) patients at the postoperative 1 month visit and subsequent follow-up visits. These patients were considered as having "persistent postoperative hypoparathyroidism" and excluded from further evaluations (n=142-15=127). Ultimately, 127 patients were left to be evaluated (Figure 1).

Table 2: Intraoperative and Early Postoperative cCa and PTH Responses

	Intraoperative (cCa n=58, PTH n=132) n (%)	Early postoperative (cCa n=143, PTH n=138) n (%)
cCa level within normal range	44 (76)	110 (77)
Elevated cCa level	13 (22.4)	6 (4)
Reduced cCa level	1 (1.6)	27 (19)
PTH level within normal range	77 (58.3)	71 (51.5)
Elevated PTH level	20 (15.2)	20 (14.5)
Reduced PTH level	35 (26.5)	47 (34)
cCa and PTH tested concurrently	57 (100)	138 (100)
Both cCa and PTH within normal range	23 (40)	110 (79.7)
Both cCa and PTH elevated	2 (3.5)	1 (0.7)
Normal cCa, reduced PTH	26 (46)	1 (0.7)
Both cCa and PTH reduced	1 (1.8)	15 (10.9)
Number of patients with pNePTH	5 (8.7)	11 (8)

n: number of patients with available data; **cCa:** albumin-corrected serum calcium level; **PTH:** intact parathormone level; **pNePTH:** postoperative normocalcemic elevated PTH.

Among these 127 patients, cCa measurements at 1, 3, 6 and 12 months postoperatively were available for 92, 71, 86 and 78 patients respectively. pHPT persistence/recurrence was observed in a total of 4/127 (3.15%) patients at 1, 3, 6 and 12 months postoperatively (n=1 at each timepoint) (Table 3). pNePTH manifestations became evident as of month 1 postoperatively. While the number of patients with laboratory workup varied at each postoperative follow-up visit, laboratory values consistent with postoperative pNePTH were identified in 56%, 44%, 42.3% and 31.9% of the patients at 1, 3, 6 and 12 months postoperatively (Table 3).

Patients with follow-up data available beyond 12 months postoperatively were evaluated separately. cCa measurements were available and within normal range for all 91 patients with a mean follow-up duration of 44.0±27.2 months (median= 36; min= 15, max= 144). Among these patients, pHPT recurrence was absent in 62 patients with PTH measurements. Considering the patients with available data, the frequency of pNePTH was 34.4% in 61 patients with both cCa and PTH measurements at the last visit (Table 3). Given the publications suggesting that pNePTH manifestations fully develop at about 6 months postoperatively, the postoperative follow-up data before and after that timepoint were reviewed for 33 patients with pNePTH at 6 months. Of these 33 patients, 30 had follow-up data at 1 and 3

Table 3: Changes in cCa and PTH Levels of Patients at Postoperative Follow-Up Visits

Time point	cCa normal n (%)	cCa elevated n (%)	PTH normal n (%)	PTH elevated n (%)	cCa & PTH tested n (%)	Both normal n (%)	Both elevated n (%)	pNePTH n (%)
1 month postop	89 (97)	3 (3)	32 (42)	45 (58)	75 (100)	32 (43)	1 (1)	42 (56)
3 months postop	70 (99)	1 (1)	32 (45)	39 (55)	71 (100)	39 (55)	1 (1)	31 (44)
6 months postop	85 (99)	1 (1)	44 (52)	40 (48)	78 (100)	44 (56.4)	1 (1.3)	33 (42.3)
12 months postop	77 (99)	1 (1.3)	48 (66.7)	24 (33.3)	72 (100)	48 (66.7)	1 (1.4)	23 (31.9)
Last visit	91 (100)	0	40 (65.6)	21 (34.4)	61 (100)	40 (65.6)	0	21 (34.4)

n: number of patients with available data; **cCa:** albumin-corrected serum calcium level; **PTH:** intact parathormone level; **pNePTH:** postoperative normocalcemic elevated PTH.

Table 4: 25-OH Vit D3 and PTH Levels

A. Mean 25-OH Vit D3 Level (ng/mL) ± SD					
Time point	**PTH elevated (mean±SD)**	**n**	**PTH normal (mean±SD)**	**n**	**P**
Intraoperative	19.9±12.8	20	20.1±13.7	77	0.967
Early postoperative	21.0±12.4	19	21.7±15.4	56	0.836
1 month postop	19.9±13.2	43	23.7±14.5	34	0.250
3 months postop	30.9±16.4	25	27.8±11.0	19	0.734
6 months postop	26.6±13.5	24	26.8±6.9	18	0.963
12 months postop	32.3±15.2	15	26.5±12.1	28	0.224
Last visit	27.1±11.0	17	30.2±12.4	35	0.393
B. Mean PTH Level (pg/mL) ± SD					
Time point	**Vit D elevated (mean±SD)**	**n**	**Vit D reduced (mean±SD)**	**n**	**P**
Intraoperative	28.36±19.02	21	49.55±36.94	76	0.057
Early postoperative	36.7±31.05	29	30.55±37.49	46	0.441
1 month postop	65.94±28.54	29	80.93±49.24	48	0.188
3 months postop	81.31±26.57	16	79.14±38.54	28	0.854
6 months postop	67.64±45.58	7	71.05±23.33	35	0.782
12 months postop	67.38±24.27	21	54.84±41.69	22	0.067
Last visit	63.20±37.70	31	74.81±44.04	21	0.360

n: number of patients with available data; **PTH:** Parathormone level; **SD:** Standard deviation.

*For calculations at intraoperative and early postoperative timepoints, preoperative 25-OH vit D3 levels were used. At other postoperative timepoints, 25-OH vit D3 levels measured at the corresponding month were used.

months postoperatively and pNePTH was present in 22/30 (73.33%) patients at 1 month and in 23/30 (76.66%) at 3 months after surgery. Beyond 6 months postoperatively, data from follow-up visit were available for 30/33 patients at 12 months postoperatively and 28/33 patients at the last visit. Persistence of pNePTH manifestations was observed in 16/30 (53.33%) patients at 12 months and 15/28 (53.57%) patients at the time of the last visit (data not shown).

Not all data were available for each postoperative follow-up visit and some data were missing for patients operated for pHPT. Looking at the available data from a different perspective, among patients with long-term follow-up data, those with pNePTH at the last visit (n=21) were evaluated. Retrospective review of the data for these patients revealed that 16/21 had PTH measurement at 12 months postoperatively and 11/16 (68.75%) had pNePTH. PTH values at 6 months postoperatively were available for all patients with pNePTH at the last visit, 15/21 of whom (71.42%) had pNePTH at 6 months after surgery. Among 21 patients with pNePTH at the time of the last visit, pNePTH was present in 12 of 16 (75%) patients at 3 months and in 10/14 (71.43%) patients at 1 month (data not shown).

Since relatively low levels of vitamin D postoperatively induces PTH secretion in the setting of sHPT, those with a PTH concentration of 65 pg/dl or higher were considered as having elevated PTH and those with a PTH less than 65 pg/dl were classified as having normal PTH. The two groups were compared to examine whether they differed in terms of 25-OH vit D3 levels. No difference was found between the groups with normal or elevated PTH in terms of 25-OH vit D3 levels (Table 4, Row A).

By the same token, the patients were divided into two groups as those with 25-OH vit D3 levels below 30 pg/ml and those with 25-OH vit D3 levels equal to or above 30 pg/ml, with the former classified as having sufficient vitamin D and the latter as having low vitamin D levels. No difference was found between the two groups with respect to PTH levels (Table 4, Row B).

Considering that progressive decrease in 1-25$(OH)_2$ vit D3 (calcitriol) in the setting of kidney failure stimulates PTH secretion, we compared mean eGFR between patients with or without pNePTH. Although eGFR values were lower in the pNePTH group, the difference was nonsignificant (Table 5, Row A). However, when we divided the patients into two groups as those with an eGFR ≥60 ml/min/1.73 m^2 and those with an eGFR <60 ml/min/1.73 m^2, higher PTH levels were observed in the latter. Although the number of patients with low eGFR was quite low, the difference between the two groups reached statistical significance especially for the PTH measurements at 1 month postoperatively and at the last visit (Table 5, Row B).

Table 5: eGFR and PTH Levels

A. Mean eGFR (ml/min/1.73 m²) ± SD					
Time point	**pNePTH (+) mean±SD**	**n**	**pNePTH (–) mean±SD**	**n**	**P**
1 month postop	93.88±24.95	38	101.50±19.98	29	0.170
3 months postop	88.19±28.65	29	102.11±25.68	20	0.082
6 months postop	92.93±29.71	27	94.94±23.77	21	0.796
12 months postop	85.70±25.63	16	99.29±23.68	35	0.083
Last visit	72.64±24.92	11	89.77±15.5	27	0.053
B. Mean PTH Level (pg/mL) ± SD					
Time point	**eGFR ≥ 60 mean±SD**	**n**	**eGFR < 60 mean±SD**	**n**	**P**
1 month postop	67.92±31.41	40	120.21±51.44	7	0.036
3 months postop	89.05±45.67	27	103.40±31.93	8	0.330
6 months postop	64.55±27.15	46	84.46±23.18	8	0.052
12 months postop	58.67±23.81	38	57.75±21.35	13	0.899
Last visit	53.32±17.99	31	98.71±40.62	7	0.025

n: number of patients with available data; **eGFR:** estimated glomerular filtration rate; **pNePTH:** postoperative normocalcemic elevated PTH; **PTH:** Parathormone; **SD:** Standard deviation; **25-OH vit D3:** 25-OH-Vitamin D3.

DISCUSSION

While the number of patients with available data at follow-up visits was variable in our series, laboratory values consistent with pNePTH were identified in 56%, 44%, 42.3% and 31.9% of the patients at 1, 3, 6 and 12 months postoperatively. There were 91 patients who had postoperative calcium levels measured, and 62 of these also had PTH values. None of these 62 patients had recurrence. The frequency of postoperative normocalcemic elevated parathormone was 34.4% among 61 patients with sCa and PTH measurements available from the last visit.

As in our study, pHPT is an endocrine disorder which predominantly affects middle-aged female population. The curative treatment of pHPT is parathyroidectomy which involves removal of autonomous parathyroid adenoma(s) or hyperplastic parathyroid (12,13). The exact cause and clinical significance of pNePTH which is characterized by elevated PTH levels despite normalization of sCa level postoperatively has not been fully elucidated. Several hypotheses were proposed to explain the pathogenetic mechanism of pNePTH. These hypotheses include desensitization of calcium-sensing receptors (CaSR) in the normal parathyroid gland due to exposure to high levels of sCa and PTH for years and the stimulation of PTH secretion by CaSR expression in response to a postoperative drop in sCa levels, or stimulation of PTH secretion due to sCa deficit as a result of replacement of Ca lost from the bone after surgery (hungry bone syndrome) (14,15). Other proposed mechanisms for pNePTH include sHPT states that occur due to low or insufficient replacement of vitamin D preoperatively, or impairment of active vitamin D production resulting from kidney failure (9,10).

The prevalence of pNePTH varies depending on the timing of patient assessment postoperatively. The reported pNePTH prevalence ranges from 12 to 43% in studies involving PTH measurement in less than 3 months after surgery (16). In the current study, the prevalence of pNePTH was 56% at 1 month and 44% at 3 months. Beyer et al. reported that the incidence of pNePTH at 3-6 months was 43% in patients treated with calcium carbonate versus 13% in those treated with calcitriol and calcium carbonate (9). In line with Beyer et al.'s findings, the incidence of pNePTH was 42.3% at 6 months in our study. There are studies reporting pNePTH rates ranging between 3% and 46% at 6 months and beyond postoperatively (9,17). In Beyer et al.'s study, some of the patients were treated with Ca tablets and others received calcitriol (9). In our daily practice, Ca supplementation is not given to patients with sCa levels within the normal range. Ca and calcitriol replacement is given only to patients with hypoparathyroidism, and therefore, in our series, only 15 patients with hypoparathyroidism identified at the postoperative follow-up received Ca tablets and calcitriol. These patients were considered "postoperative hypoparathyroidism" and thus, excluded from study assessments. In our series, considering that relatively low postoperative sCa levels may stimulate PTH secretion, patients with normal sCa levels were not given Ca tablets with the aim to lower PTH through Ca replacement. None of the evaluable patients received calcitriol replacement. Therefore, we hypothesize that the relatively low sCa and the lack of active vitamin D (1,25(OH)2 vit D3) may have led to the development of sHPT, which may have led to the increased prevalence of pNePTH in our study.

As known, it is 1,25 $(OH)_2$ vit D3 (calcitriol) or the metabolically active form of vitamin D that functions as vitamin D by binding to the receptors. PTH secretion is stimulated in the case of 1,25 $(OH)_2$ vit D3 deficiency which is associated with kidney failure. In our study sample, relatively higher PTH values were found among those with lower eGFR. Since 1,25 $(OH)_2$ vit D3 was not tested in any of the patients, it is not known whether the patients had adequate levels of active vitamin D. In many laboratories, 1,25 $(OH)_2$ vit D3 is not routinely tested. This confirms that the deficiency of 1,25 $(OH)_2$ vit D3 or active vitamin D rather than 25-OH vit D3 deficiency may have led to elevation of PTH. In general practice, 25-OH vit D3 level is measured to determine whether a patient has sufficient amount of vitamin D, and the decision for vitamin D replacement is based on the results of this test. Moreover, the replacement doses recommended by most of the guidelines are also based on 25-OH vit D3 results. In special patient populations such as those with kidney failure, 1,25 $(OH)_2$ vit D3 level is rarely tested. Similarly, at our hospital, the decision for vitamin D replacement is routinely made on the basis of 25-OH vit D3 level. In this retrospective study, none of the patients were tested for 1,25 $(OH)_2$ vit D3.

The definition of normocalcemic hyperparathyroidism requires a 25-OH vit D3 level of 30 ng/ml or higher (18). In our clinical practice, only pNePTH patients with a vitamin D level below 30 ng/ml received 25-OH vit D3 replacement to raise 25-OH vit D3 level above 30 ng/ml. Possibly for this reason, no difference was found in 25-OH vit D3 levels between the groups with elevated or normal PTH. Although no significant difference was observed comparing PTH levels between patients with or without a low vitamin D level based on preoperative 25-OH vit D3 level, patients with lower 25-OH vit D3 levels showed higher intraoperative PTH levels. This suggests that intraoperative elevation of PTH may occur due to sHPT in patients with low preoperative vitamin D levels. However, since all patients with a postoperative 25-OH vit D3 level less than 30 ng/dl received vitamin D replacement, this difference seems to disappear. It should be kept in mind that the aforementioned comparisons did not include all patients but could only be made among patients with 25-OH vit D3 and PTH measurements available.

Comparing the mean eGFR between patients with or without elevated PTH, the difference was nonsignificant. However, when we divided the patients into two groups as those with an eGFR ≥60 ml/min/1.73 m^2 and those with an eGFR <60 ml/min/1.73 m^2, relatively higher PTH levels were observed in the latter although the number of patients was small in that group. The difference reached statistical significance for the PTH levels measured at 1 month posoperatively and at the last visit. This suggests that, in the setting of kidney failure, sHPT may contribute to the development of pNePTH. However, the number of patients with an eGFR <60 ml/min/1.73 m^2 was quite small, and further studies involving a larger sample are warranted to draw a firm conclusion. In addition, sHPT may occur due to chronic kidney insufficiency preventing the synthesis of calcitriol via 1α-hydroxylation of 25 OH vit D3, which may be manifested by elevated PTH levels in combination with normal sCa. As these patients are not tested for calcitriol, it is not known whether they have adequate levels of active vitamin D which is responsible for biological activity. Consistently, in a study by Cao et al., while no difference was found between the groups with or without pNePTH in 25-OH vit D3 levels, 1,25 $(OH)_2$ vit D3 levels were lower in patients with pNePTH (6). Had calcitrol been tested, the difference that could not be shown between patients with or without pNePTH in terms of 25-OH vit D3 levels could perhaps be demonstrated for calcitriol levels.

Among our patients with available data, PHPT recurrence occurred in 4 (3.15%) patients within 12 months after surgery. Those with PHPT recurrence at 1 month and 3 months showed high cCa and PTH values at 6 months as well. Elevated cCa and PTH levels were observed in one patient at 6 months and at 12 months in another patient. In some publications, persistent hyperparathyroidism is defined as the occurrence of hypercalcemia within 6 months of parathyroidectomy, whereas recurrent hyperparathyroidism is defined as hypercalcemia developing after 6 months (19). According to these criteria, these 3 of these 4 patients may be classified as having persistent PHPT because of elevated cCa and PTH within the first 6 months (3/127=2.36%) and the remaining one (1/127=0.79%) patient can be considered as having recurrent PHPT due to high cCa and PTH at 12 months.

Of the patients with available follow-up data for more than 12 months, none had PHPT recurrence. In patients with PHPT, the reported recurrence rates after parathyroidectomy vary from 2.5% to 10% depending on the characteristics of the patient population studied, the type of parathyroidectomy and the duration of follow-up (20,21). In a case series of Mallick et al., recurrence was reported in 28 (10.7%) of 261 patients with a mean time to recurrence of 77 months (range 13-170) (22). Lou et al. reported a 10-year recurrence rate of 14.8% among 196 patients. The median time to recurrence was 6.3 years (interquartile range, 3.4-10.8 years) and 34.5% of all recurrences occurred within 10 years after surgery (23). In addition, double adenomas were more common in patients with disease recurrence (19,23,24). In a study by Nordenström et al., a recurrence rate of 16% was found in patients with PHPT over a follow-up duration, and the significance of long-term follow-up in surgically treated patients was underscored (25). The presence of pNePTH is an important factor that predicts PHPT recurrence. Recur-

rent PHPT is present in 3-7% of patients with pNePTH (26). In our sample, pNePTH was also present in the previous follow-up visits in 3 patients with persistent disease who developed PHPT within 6 months postoperatively.

In some publications, it is mentioned that parathyroid hyperplasia may lead to persistent/recurrent PTH elevation in the postoperative period (19,27). In our case series, elevated PTH may be caused by PHPT associated with hyperplastic parathyroid gland. However, in this retrospective study, the available data were insufficient to confirm this argument. Classification of persistence and recurrence was based on laboratory results only, and the diagnosis could not be confirmed due to incomplete clinical and radiological data of the patients. During long-term follow-up of patients after parathyroidectomy, sCa and PTH levels should be followed at 6 and 12 months and annually therafter, with serum ionized calcium and 25-OH vit D3 levels monitored regularly if PTH elevation is detected.

The major limitation of our study was its retrospective nature and the lack of complete long-term follow-up data at all months postoperatively for the entire sample of 127 patients. For example, 91 of 127 patients were followed for longer than 12 months and had normal cCa levels. Of these 91 patients, 61 were tested for PTH which was normal in 40 (65.6%), and 21 (34.4%) had pNePTH. However, sCa and PTH levels were unknown for 36 patients. Among the patients with follow-up data, pHPT recurrence was not observed but for these 36 patients, it is unclear whether they had disease recurrence and if so, how many of them were affected.

Conclusion

Although the frequency of pNePTH in case series varies depending on the criteria used, it remains high. Whether underlying sHPT contributes to pNePTH should be thoroughly investigated and factors associated with persistence and recurrence treated. Considering that pNePTH may also represent PHPT recurrence, long-term follow-up is required for patients.

Author Contributions

Study conception and design: **Gülşah Elbüken, Sibel Özkan Gürdal, Sayid Shafi Zuhur;** data collection: **Ogün Bilen, Beyza Olcay Öztürk, Hünkar Ağgül;** analysis and interpretation of results: **Burcu Tokuç;** draft manuscript preparation: **Gülşah Elbüken.** The authors reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Financial Support

The authors declare that the study received no funding.

Ethical Approval

Approval for the study was obtained from the Ethics Committee of Tekirdağ Namık Kemal University Medical Faculty on 25.01.2022 (No. 2022.07.01.07).

REFERENCES

1. Wermers RA, Griebeler ML, Thapa P, Hathcock MA, Kearns AE. Survival in primary hyperparathyroidism over five decades (1965-2010): a population-based retrospective study. Bone. 2021;152:116099. https://doi.org/10.1016/j.bone.2021.116099
2. Walker MD, Silverberg SJ. Primary hyperparathyroidism. Nat Rev Endocrinol. 2018;14(2):115-125. https://doi.org/10.1038/nrendo.2017.104
3. Kearns AE, Thompson GB. Medical and surgical management of hyperparathyroidism. Mayo Clin Proc. 2002;77(1):87-91. https://doi.org/10.4065/77.1.87
4. Ahmed FW, Majeed MS, Kirresh O. Multiple endocrine neoplasias type 4. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022.
5. Bilezikian JP, Brandi ML, Eastell R, Silverberg SJ, Udelsman R, Marcocci C, et al. Guidelines for the management of asymptomatic primary hyperparathyroidism: summary statement from the Fourth International Workshop. J Clin Endocrinol Metab. 2014;99:3561-3569. https://doi.org/10.1210/jc.2014-1413
6. Cao S, Hu Y, Zhao Y, Su Z, Xu Z, Gao X, et al. A retrospective study of elevated post-operative parathormone in primary hyperparathyroid patients. Oncotarget. 2017;8(60):101158-101164. https://doi.org/10.18632/oncotarget.20416
7. Carty SE, Roberts MM, Virji MA, Haywood L, Yim JH. Elevated serum parathormone level after "concise parathyroidectomy" for primary sporadic hyperparathyroidism. Surgery. 2002;132(6):1086-1092. https://doi.org/10.1067/msy.2002.128479
8. Brown EM. 4-parameter model of the sigmoidal relationship between parathyroid hormone release and extracellular calcium concentration in normal and abnormal tissue. J Clin Endocrinol Metab. 1983;56(3):572-581. https://doi.org/10.1210/jcem-56-3-572
9. Beyer TD, Solorzano CC, Prinz RA, Babu A, Nilubol N, Patel S. Oral vitamin D supplementation reduces the incidence of eucalcemic PTH elevation after surgery for primary hyperparathyroidism. Surgery. 2007;141:777-783. https://doi.org/10.1016/j.surg.2007.01.025
10. de la Plaza Llamas R, Ramia Ángel JM, Arteaga Peralta V, Garcia Amador C, López Marcano AJ, Medina Velasco AA, et al. Elevated parathyroid hormone levels after successful parathyroidectomy for primary hyperparathyroidism: a clinical review. Eur Arch Otorhinolaryngol. 2018;275(3):659-669. https://doi.org/10.1007/s00405-017-4836-9
11. Barczynski M, Konturek A, Hubalewska-Dydejczyk A, Cichon S, Nowak W. Evaluation of Halle, Miami, Rome, and Vienna intraoperative iPTH assay criteria in guiding minimally invasive parathyroidectomy. Langenbecks Arch Surg. 2009;394:843-849. https://doi.org/10.1007/s00423-009-0510-z

12. Hessman O, Westerdahl J, Al-Suliman N, Christiansen P, Hellman P, Bergenfelz A. Randomized clinical trial comparing open with video-assisted minimally invasive parathyroid surgery for primary hyperparathyroidism. Br J Surg. 2010;97:177-184. https://doi.org/10.1002/bjs.6810
13. Roman SA, Sosa JA, Mayes L, Desmond E, Boudourakis L, Lin R, et al. Parathyroidectomy improves neurocognitive deficits in patients with primary hyperparathyroidism. Surgery. 2005;138:1121-1128. https://doi.org/10.1016/j.surg.2005.08.033
14. Kifor O, Moore FD, Wang P, Goldstein M, Vassilev P, Kifor I, et al. Reduced immunostaining for the extracellular Ca2+-sensing receptor in primary and uremic secondary hyperparathyroidism. J Clin Endocrinol Metab. 1996;81:1598-1606. https://doi.org/10.1210/jcem.81.4.8636374
15. Gogusev J, Duchambon P, Hory B, Giovannini M, Goureau Y, Sarfati E, et al. Depressed expression of calcium receptor in parathyroid gland tissue of patients with hyperparathyroidism. Kidney Int. 1997;51:328-336. https://doi.org/10.1038/ki.1997.41
16. Yamashita H, Noguchi S, Moriyama T, Takamatsu Y, Sadanaga K, Uchino S, et al. Reelevation of parathyroid hormone level after parathyroidectomy in patients with primary hyperparathyroidism: importance of decreased renal parathyroid hormone sensitivity. Surgery. 2005;137:419-425. https://doi.org/10.1016/j.surg.2004.12.009
17. Mandal AK, Udelsman R. Secondary hyperparathyroidism is an expected consequence of parathyroidectomy for primary hyperparathyroidism: a prospective study. Surgery. 1998;124:1021-1026. https://doi.org/10.1067/msy.1998.92004
18. Zavatta G, Clarke BL. Normocalcemic hyperparathyroidism: a heterogeneous disorder often misdiagnosed? JBMR Plus. 2020;4(8):e10391. https://doi.org/10.1002/jbm4.10391
19. Mazotas IG, Yen TWF, Doffek K, Shaker JL, Carr AA, Evans DB, et al. Persistent/recurrent primary hyperparathyroidism: does the number of abnormal glands play a role? J Surg Res. 2020;246:335-341. https://doi.org/10.1016/j.jss.2019.08.007
20. Shirali AS, Wu SY, Chiang YJ, Graham PH, Grubbs EG, Lee JE, et al. Recurrence after successful parathyroidectomy: who should we worry about? Surgery. 2022;171(1):40-46. https://doi.org/10.1016/j.surg.2021.06.035
21. Pavlidis ET, Pavlidis TE. Update on the current management of persistent and recurrent primary hyperparathyroidism after parathyroidectomy. World J Clin Cases. 2023;11(10):2213-2225. https://doi.org/10.12998/wjcc.v11.i10.2213
22. Mallick R, Nicholson KJ, Yip L, Carty SE, McCoy KL. Factors associated with late recurrence after parathyroidectomy for primary hyperparathyroidism. Surgery. 2020;167(1):160-165. https://doi.org/10.1016/j.surg.2019.05.076
23. Lou I, Balentine C, Clarkson S, Schneider DF, Sippel RS, Chen H. How long should we follow patients after apparently curative parathyroidectomy? Surgery. 2017;161(1):54-61. https://doi.org/10.1016/j.surg.2016.05.049
24. Alhefdhi A, Schneider DF, Sippel R, Chen H. Recurrent and persistent primary hyperparathyroidism occurs more frequently in patients with double adenomas. J Surg Res. 2014;190:198-202. https://doi.org/10.1016/j.jss.2014.02.024
25. Nordenström E, Westerdahl J, Bergenfelz A. Long-term follow-up of patients with elevated PTH levels following successful exploration for primary hyperparathyroidism. World J Surg. 2004;28:570-575. https://doi.org/10.1007/s00268-004-7124-y
26. Oltmann SC, Maalouf NM, Holt S. Significance of elevated parathyroid hormone after parathyroidectomy for primary hyperparathyroidism. Endocr Pract. 2011;17:57-62. https://doi.org/10.4158/EP10324.RA
27. Goldfarb M, Gondek S, Irvin GL 3rd, Lew JI. Normocalcemic parathormone elevation after successful parathyroidectomy: long-term analysis of parathormone variations over 10 years. Surgery. 2011;150:1076-1084. https://doi.org/10.1016/j.surg.2011.09.017

ÖZGÜN ARAŞTIRMA

Batı Karadeniz Tıp Dergisi
Medical Journal of Western Black Sea

Med J West Black Sea 2026;10(1):116-121
DOI: 10.29058/mjwbs.1801109

Adneksiyal kitlesi olan kadınlarda siberkondri düzeyini etkileyen faktörler: Tek merkezli kesitsel çalışma

Factors affecting the level of cyberchondria in women with adnexal masses: A single-center cross-sectional study

Hande Nur Öncü[1], Gökçen Ege[1], Neslihan Öztürk[1], Candost Hanedan[1], Vakkas Korkmaz[2]

[1]Jinekolojik Onkoloji Kliniği, Etlik Şehir Hastanesi, Ankara, Türkiye
[2]Kadın Hastalıkları ve Doğum Kliniği, Lokman Hekim Üniversitesi, Ankara, Türkiye

Bu makaleye yapılacak atıf: *Öncü HN ve ark. Adneksiyal kitlesi olan kadınlarda siberkondri düzeyini etkileyen faktörler: tek merkezli kesitsel çalışma. Med J West Black Sea. 2026;10(1):116-121.*

ÖZ

Amaç: Bu çalışmada, adneksiyal kitlesi olan kadınlarda siberkondri düzeylerini Siberkondri Ciddiyet Ölçeği-Kısa Formu (SCÖ-12) kullanarak değerlendirmeyi amaçladık.

Gereç ve Yöntemler: Bu çalışma, tek merkezde yürütülen kesitsel bir araştırma olarak planlandı. Jinekolojik onkoloji kliniğine başvuran ve adneksiyal kitle tanısı alan toplam 134 kadın çalışmaya dahil edildi. Tüm katılımcılara SCÖ-12 uygulandı. SCÖ-12 toplam ve alt ölçek skorları; medeni durum, eğitim düzeyi, çocuk sahibi olma durumu, CA-125 düzeyi ve menopoz durumu değişkenlerine göre karşılaştırıldı.

Bulgular: Katılımcı grubun median yaşı 38 (18–75), median vücut kütle indeksi (VKİ) 30 (20–42), median parite sayısı 2 (0–8) ve median CA-125 değeri 45 (11–2350) olarak saptandı. Toplam SCÖ-12 skoru ve dört alt boyut karşılaştırıldığında, genç (≤35 yaş), çocuğu olmayan ve lise ve üzeri eğitim düzeyine sahip kadınlarda toplam SCÖ-12 skorları anlamlı olarak daha yüksek bulundu (p<0,05). Alt ölçek analizinde de aşırılık ve sıkıntı alt boyutlarının bu gruplarda anlamlı olarak daha yüksek olduğu saptandı (p<0,05).

Sonuç: Adneksiyal kitle tanısı alan kadınlarda siberkondri, önemli bir sorun olarak öne çıkmaktadır. Genç yaş, çocuk sahibi olmama ve yüksek eğitim düzeyi, siberkondri şiddetiyle anlamlı olarak ilişkili bulunmuştur. Hekimlerin bu hasta grubuna uygun bilgilendirme ve yönlendirme yapması önemlidir.

Anahtar Kelimeler: Adneksiyal kitle, anksiyete, internet, kadın sağlığı, siberkondri

ABSTRACT

Aim: This study aimed to evaluate the level of cyberchondria among women with adnexal masses using the Cyberchondria Severity Scale–Short Form (CSS-12).

Material and Methods: This study was designed as a single-center, cross-sectional study. A total of 134 women diagnosed with adnexal masses who presented to the gynecologic oncology clinic were included in the study. All participants completed the CSS-12 questionnaire. Total and subscale scores of CSS-12 were compared according to marital status, educational level, parity, CA-125 levels, and menopausal status.

Results: The median age of participants was 38 years (18–75), the median body mass index (BMI) was 30 (20–42), the median parity was 2 (0–8), and the median CA-125 level was 45 IU/mL (11–2350). The total CSS-12 score was significantly higher among younger women (≤35 years), nulliparous women, and those with a high school or higher educational level (p<0.05). Subgroup analysis revealed that excessiveness and distress scores were also significantly higher in these groups (p<0.05).

Conclusion: Cyberchondria is a significant issue among women diagnosed with adnexal masses. Younger age, nulliparity, and higher education levels are strongly associated with greater severity of cyberchondria. Healthcare providers should provide relevant information and guidance specific to this patient group.

Keywords: Adnexal mass, anxiety, cyberchondria, internet, women's health

Sorumlu Yazar: Hande Nur Öncü ✉ handekayahan@gmail.com **Geliş Tarihi:** 11.10.2025 **Revizyon Tarihi:** 04.11.2025 **Kabul Tarihi:** 02.02.2026

Telif hakkı © 2026 Yazar(lar). Zonguldak Bülent Ecevit Üniversitesi tarafından yayımlanmıştır. Açık erişimli bu makale, orijinal çalışmaya uygun şekilde atıfta bulunulması koşuluyla, herhangi bir ortamda veya formatta sınırsız kullanım, dağıtım ve çoğaltmaya izin veren Creative Commons Atıf Lisansı (CC BY) ile dağıtılmıştır.

Öne Çıkan Noktalar

- Bu çalışma, adneksiyal kitle tanısı alan kadınlarda siberkondri düzeyini değerlendiren ilk klinik çalışmalardan biridir.
- Siberkondri düzeyi genç (≤35 yaş), çocuğu olmayan ve lise ve üzeri eğitim düzeyine sahip kadınlarda anlamlı olarak daha yüksektir.
- Aşırılık (excessiveness) ve sıkıntı (distress) alt boyutları bu hasta gruplarında belirgin olarak artmıştır.
- Adneksiyal kitle gibi sık görülen bir klinik durumda, siberkondri açısından yüksek riskli hastaların belirlenmesi, uygun hasta bilgilendirmesi ve danışmanlığı açısından önem taşımaktadır.

GİRİŞ

Siberkondri, bireylerin sağlıklarıyla ilgili olarak internet üzerinden aşırı veya tekrarlayıcı biçimde bilgi aramaları sonucu kaygı ve endişe düzeylerinin artmasıyla karakterize bir durumdur (1). İnternetin hızlı, ücretsiz, anonim ve kolay erişilebilir olması, onu sağlıkla ilgili bilgi edinmede en sık başvurulan kaynaklardan biri hâline getirmiştir. Amerika Birleşik Devletleri'nde 2017 yılında yapılan bir araştırmada, bireylerin %74,4'ünün internet üzerinden sağlık bilgisi aradığı bildirilmiştir (2).

Çevrimiçi sağlık aramaları bazı durumlarda bireylerin hastalıkları, tedavi seçenekleri ve önleme yöntemleri hakkında bilgi edinmesini sağlayarak sağlık okuryazarlığını artırabilir (3). Ancak internet ortamında yer alan bilgilerin doğruluk düzeyinin değişken olması ve nadir fakat ciddi hastalıkları da içeren olasılıkların sunulması, bireylerde yanlış yorumlama ve aşırı kaygıya neden olabilmektedir (4). Bu durum tekrarlayan internet aramalarına, yanlış tanılara ve artan sağlık anksiyetesine yol açabilir. McElroy ve Shevlin tarafından 2014 yılında geliştirilen Siberkondri Ciddiyet Ölçeği (SCÖ) bireylerdeki bu davranışın düzeyini değerlendirmeyi amaçlamaktadır (5). Daha sonra, ölçeğin geçerlilik ve güvenilirlik özellikleri korunarak 12 maddeden oluşan kısa formu (SCÖ-12) geliştirilmiştir (6).

Adneksiyal kitle benign, borderline veya malign özellikte olabilen lezyonları kapsayan ve jinekolojik pratikte sık karşılaşılan bir klinik bulgudur. Premenopozal kadınların yaklaşık %8 ila %35'inde, postmenopozal kadınların ise %3 ila %17'sinde tespit edilebilir (7-9). Adneksiyal kitlelerin prevalansı yüksek olmasına karşın, over kanseri geliştirme riski yaklaşık %1'dir (10). Çoğu adneksiyal kitle benign olsa da malignite olasılığı hastalarda kaygı yaratabilmektedir. Görüntüleme yöntemleri kitleyi karakterize etmek için yaygın olarak kullanılsa da kesin tanı genellikle histopatolojik inceleme ile konur.

Adneksiyal kitle tanısı alan kadınlar, malignite riski veya cerrahi gereklilik konusunda yeterli bilgiye sahip olmadan önce önemli düzeyde endişe yaşayabilmektedir. Tanısal belirsizlik süreci, kanser olasılığına dair korku ve üreme sağlığına ilişkin kaygılar bu hastalarda internet üzerinden yoğun bilgi arama davranışını tetikleyebilmekte ve bu durum siberkondri gelişimine zemin hazırlayabilmektedir. Bu nedenle birçok hasta tanı sürecinde internetten bilgi arayışına yönelmektedir. Ancak bu durum, yanlış veya olumsuz bilgilere maruz kalmasına ve anksiyete düzeyinin daha da artmasına neden olabilir. Hekimlere sık başvuru, tedavi kararını erteleme veya farklı tedavi arayışları gibi sonuçlara da yol açabilmektedir. Adneksiyal kitlelerin neden olduğu tanısal belirsizlik ve olası malignite kaygısı göz önüne alındığında, bu hasta grubunda internet kullanımına bağlı anksiyetenin değerlendirilmesi klinik yönetim açısından önemlidir.

Bu çalışmanın amacı, adneksiyal kitlesi olan kadınlarda siberkondri düzeyini SCÖ-12 kullanarak değerlendirmek ve buna etki eden klinik ve demografik faktörleri belirlemektir.

GEREÇ ve YÖNTEMLER

Bu çalışma, lokal etik kurul onayı alındıktan sonra, Mayıs 2025–Eylül 2025 tarihleri arasında, jinekolojik onkoloji kliniğine başvuran ve adneksiyal kitle nedeniyle operasyon planlanan 134 kadın üzerinde gerçekleştirildi. İnternet erişimi olmayan, okuma-yazması bulunmayan, daha önce adneksiyal kitle nedeniyle operasyon geçirmiş olan, psikiyatrik hastalığı veya düzenli psikiyatrik ilaç kullanımı olan ve çalışmaya katılmayı reddeden hastalar çalışma dışında bırakıldı. 18–90 yaş aralığında, ilk kez adneksiyal kitle tanısı nedeniyle ameliyat planlanmış, günde en az bir saat internet kullandığını ve çevrim içi bilgi arama davranışına sahip olduğunu beyan eden (aktif internet kullanıcısı olan) ve çalışmaya katılmayı kabul eden hastalar dahil edildi. Anket uygulanmadan önce katılımcılara çalışmanın amacı açıklanarak yazılı onamları alındı.

Siberkondri düzeyi, McElroy ve ark. tarafından geliştirilen SCÖ-12 kullanılarak değerlendirildi. SCÖ-12, 12 maddeden oluşan, 5'li Likert tipi (1=Asla, 5=Her zaman) bir ölçektir ve dört alt boyuttan oluşmaktadır: (i) Zorlama: Yoğun çevrimiçi aramaların günlük yaşamı kesintiye uğratması; (ii) Aşırılık: Tekrarlayıcı veya aşırı tıbbi bilgi arama davranışı; (iii)

Sıkıntı: Olumsuz aramaların neden olduğu uykusuzluk ve depresyon gibi duygusal tepkiler; (iv) Güvence arayışı: Tıp uzmanları veya diğer kaynaklardan güvence arama eğilimi. Toplam puan 12–60 arasında değişmekte olup, yüksek puan daha yüksek siberkondri düzeyini göstermektedir (6).

Siberkondri Ciddiyet Ölçeği-12'nin Türkçe geçerlilik ve güvenirlik çalışması, 2023 yılında Türkiye'de gerçekleştirilmiş olup, ilgili çalışmada ölçeğin iç tutarlılık katsayısı (Cronbach alfa) 0,80 olarak bildirilmiştir (11). Mevcut çalışmamızda SCÖ-12'nin iç tutarlılığı yeniden değerlendirilmiş ve toplam ölçek için Cronbach alfa değeri 0,76 olarak hesaplanmıştır.

Katılımcıların yaş (35 yaş altı veya üstü), çocuk sahibi olma durumu (nullipar veya multipar), menopoz durumu (premenopozal veya postmenopozal), eğitim düzeyi (lisenin altı veya lise ve üstü), medeni durumu (evli veya evli olmayanlar) ve CA-125 (35 altı veya üstü) değerine göre SCÖ-12 toplam ve alt ölçek skorları karşılaştırıldı.

İstatistiksel Analiz

İstatistiksel analizler IBM SPSS Statistics for Mac, sürüm 22.0 (Armonk, NY, ABD) programı kullanılarak gerçekleştirildi. Sürekli değişkenler medyan (minimum–maksimum) değerleriyle belirtildi. Sürekli değişkenlerin normal dağılım göstermemesi nedeniyle, gruplar arasındaki karşılaştırmalarda Mann–Whitney U testi kullanıldı.

Örneklem büyüklüğü, çalışmanın başlangıcında a priori güç analizi ile belirlendi. Güç analizi G*Power yazılımı kullanılarak, birincil sonuç değişkeni olan SCÖ-12 toplam skoruna dayalı olarak yapıldı. Analizde α=0,05 ve istatistiksel güç (1–β)=0,80 kabul edildi ve minimum örneklem büyüklüğü 128 olarak hesaplandı. Çalışma 134 katılımcı ile tamamlandı. Tüm istatistiksel analizlerde $p<0,05$ değeri anlamlı kabul edildi.

BULGULAR

Median yaş 38 (18–75), median vücut kütle indeksi (VKİ) 30 (20–42), median parite sayısı 2 (0–8) ve median CA-125 değeri 45 (11–2350) olarak tespit edildi. Hastaların %22,4'ü (30/134) menopozda, %20,1'i (27/134) nullipar, %70,9'u (95/134) lise ve üzeri eğitim seviyesinde, %71,6'sı (96/134) evliydi ve iki hasta (%1,5) gebeydi. Hastaların demografik ve klinik özellikleri Tablo 1'de gösterilmektedir.

Hastaların SCÖ-12 toplam skorları ve alt ölçekleri karşılaştırıldığında, yaşı ≤35 olan hastalarda median toplam SCÖ-12 skoru 36 (12–57) iken, >35 yaş grubunda 33 (12–56) olarak bulundu (p=0,010). Nullipar kadınların median toplam skoru 37 (13–58), multipar kadınlarınki 33 (12–50) idi (p=0,003). Lise ve üzeri eğitim düzeyine sahip kadınlarda median toplam skor 35 (12–58), daha düşük eğitim düzeyindekilerde ise 31 (13–50) olarak saptandı (p=0,020).

Alt ölçek analizinde, ≤35 yaş olan kadınlarda aşırılık ve sıkıntı alt boyut skorları sırasıyla p=0,015 ve p=0,022; nullipar kadınlarda p=0,008 ve p=0,011; lise ve üzeri eğitim düzeyine sahip kadınlarda ise p=0,030 ve p=0,025 olarak anlamlı derecede daha yüksek bulundu. Menopoz durumu, medeni durum ve CA-125 düzeyi ile SCÖ-12 toplam veya alt ölçek skorları arasında anlamlı bir ilişki saptanmadı (p>0,05) (Tablo 2).

TARTIŞMA

Bu çalışmada, adneksiyal kitlesi olan kadınlarda genç yaş, çocuğu olmama ve yüksek eğitim düzeyi ile siberkondri düzeyi arasında anlamlı bir ilişki olduğu saptandı. Jinekolojik pratikte sık karşılaşılan adneksiyal kitle hastalarında siberkondriye etki eden faktörlerin belirlenmesi, bu hasta grubunun bilgilendirilmesi ve yönetimi açısından klinik olarak önemli olduğunu düşünmekteyiz.

Siberkondri, özellikle COVID-19 pandemisi döneminde, artan belirsizlik ve internet üzerinden bilgiye erişimin kolaylaşması ile daha çok gündeme gelmiş ve bu konu ile ilgili birçok çalışma yayımlanmıştır (12-14). Ancak bu çalışmaların büyük bölümü hasta olmayan, öğrenci veya çalışan bireylerden oluşan popülasyonlarda yapılmıştır.

Son yıllarda siberkondri düzeyi, belirli hastalık gruplarında da araştırılmaya başlanmıştır. Sucu ve ark.'ı polikistik over sendromu (PKOS) olan adölesanlarda toplam SCÖ skorlarının aynı yaş grubundaki sağlıklı bireylere kıyasla anlamlı

Tablo 1: Çalışma Grubunun Demografik ve Klinik Özellikleri

Değişkenler	
Hasta sayısı	134
Yaş (yıl), medyan (min-max)	38 (18-75)
VKI (kg/m²), medyan (min-max)	30 (20-42)
Parite, medyan (min-max)	2 (0-8)
CA-125 (IU/mL), medyan (min-max)	45 (11- 2350)
Eğitim düzeyi, n (%)	
Lise ve üzeri	95 (70,9)
Lise altı	39 (29,1)
Menopoz durumu, n (%)	
Premenopoz	104 (77,6)
Postmenopoz	30 (22,4)
Medeni durum, n (%)	
Evli	96 (71,6)
Bekar	38 (28,4)
Gebe, n (%)	2 (1,5)

Veriler medyan (minimum–maksimum) veya sayı (%) olarak ifade edilmiştir. **VKI:** vücut kütle indeksi.

Tablo 2: Hasta Özelliklerine Göre SCÖ-12 Toplam ve Alt Grup Skorlarının Karşılaştırılması

Değişken / Grup	n	Toplam SCÖ-12	p	Aşırılık	p	Sıkıntı	p	Zorlama	p	Güven arayışı	p
Yaş											
≤35 yaş	49	36 (12–57)	0,010	10 (3–15)	0,015	9 (3-14)	0,022	8 (3–14)	0,410	9 (3–14)	0,280
>35 yaş	85	33 (12–56)		8 (3–15)		8 (3–14)		8 (3–13)		8 (3–14)	
Parite											
Nullipar	27	37 (13–58)	0,003	10 (3–15)	0,008	9 (4–14)	0,011	7 (3–15)	0,370	9 (3–14)	0,295
Multipar	107	33 (12–60)		9 (3–15)		8 (3–15)		7 (3–15)		8 (3–15)	
Eğitim seviyesi											
Lise ve üstü	95	35 (12–58)	0,020	10 (3–14)	0,030	9 (3–15)	0,025	8 (3–15)	0,410	9 (3–14)	0,340
Lise altı	39	31 (13–50)		8 (3–12)		8 (3–13)		8 (4–13)		8 (3–12)	
Menopoz durumu											
Premenopoz	104	34 (13–57)	0,180	9 (3–14)	0,190	8 (3–14)	0,210	7 (4–15)	0,480	8 (3–14)	0,510
Postmenopoz	30	33 (13–55)		9 (3–14)		8 (3–13)		7 (4–15)		8 (3–13)	
Medeni durum											
Evli	96	34 (12–58)	0,440	9 (3–14)	0,460	7 (3–14)	0,480	8 (3–15)	0,500	8 (3–15)	0,530
Bekâr	38	35 (12–58)		9 (3–15)		7 (3–14)		8 (3–15)		8 (3–14)	
CA-125 seviyesi											
Normal (≤35 IU/mL)	92	34 (13–58)	0,410	9 (3–14)	0,420	8 (3–15)	0,430	7 (4–15)	0,450	8 (3–14)	0,470
Yüksel (>35 IU/mL)	42	34 (12–59)		9 (3–15)		8 (3–14)		7 (3–15)		8 (3–15)	

Veriler medyan (minimum–maksimum) olarak sunulmuştur.
p-değerleri Mann–Whitney U testi kullanılarak hesaplanmıştır; p<0,05 istatistiksel olarak anlamlı kabul edilmiştir.
İstatistiksel olarak anlamlı değerler kalın olarak gösterilmiştir.

olarak daha yüksek olduğunu bildirmiştir (15). Benzer şekilde Çakır, HPV (humanpapilloma virus) pozitif ve anormal uterin kanaması olan kadınların siberkondri düzeylerini karşılaştırmış ve HPV pozitif olanlarda SCÖ skorlarının anlamlı olarak yüksek olduğunu göstermiştir (16). HPV alt tipleri üzerine yapılan bir diğer çalışmada ise, HPV 16/18 pozitif kadınlarda toplam SCÖ ve aşırılık alt boyutu skorlarının diğer yüksek riskli HPV gruplarına kıyasla anlamlı olarak daha yüksek olduğu tespit edilmiştir (17). Bizim çalışmamızda da benzer şekilde, genç, çocuğu olmayan ve eğitim düzeyi yüksek kadınlarda aşırılık ve sıkıntı alt grupları anlamlı olarak yüksek bulunmuştur. Bu bulgular, bu hasta grubunda olası fertilite kaybı veya toplumsal beklentilere bağlı kaygıların siberkondriyi artırabileceğini düşündürmektedir.

Gebeler üzerinde yapılan çalışmalarda da benzer sonuçlar bildirilmiştir. Üniversite mezunu ve doğum yapmamış gebelerde SCÖ skorlarının daha yüksek olduğu ve siberkondri ile sağlık anksiyetesi arasında pozitif ilişki olduğu gösterilmiştir (18). Sostaric ve ark.'ı ise tıbbi komplikasyonu olan gebelerde, komplikasyonu olmayanlara kıyasla siberkondri düzeyinin anlamlı olarak yüksek olduğunu bildirmiştir (19). Genel popülasyonda yapılan bir araştırmada, siberkondri ile sağlık anksiyetesi arasında anlamlı bir ilişki olduğu ve bu ilişkinin bireylerin sağlık okuryazarlığı düzeyine göre değişebildiği gösterilmiştir. Sağlık okuryazarlığı arttıkça, bireyler internetten edindikleri bilgileri daha eleştirel değerlendirebilmekte, bu da kaygı düzeyini etkileyebilmektedir (20). Bu durum, eğitim düzeyi yüksek kadınlarda siberkondri düzeylerinin neden farklı şekilde seyrettiğini açıklayabilir.

İlginç olarak, çalışmamızda CA-125 düzeyi ile siberkondri şiddeti arasında anlamlı bir ilişki saptanmadı. Bu beklenmedik bulgunun nedeni net değildir. Çalışma kriterlerimiz gereği tüm katılımcıların aktif internet kullanıcısı olması, CA-125 düzeyi yüksek olan ancak internet kullanmayan bireylerin etkisini dışlamış olabilir.

Çalışmamızda genç ve nullipar hastalarda siberkondri düzeylerinin daha yüksek olması dikkat çekicidir. Üreme çağındaki kadınlarda fertilite kaybı korkusu, doğurganlık potansiyelinin cerrahi sonrası etkilenebileceğine dair endişeler ve toplumsal beklentiler, internetten yoğun bilgi arama davranışını tetikleyebilir. Bu durum, mevcut tanısal belirsizlik ve malignite olasılığına ilişkin kaygılarla birleştiğinde anksiyete düzeyinin artmasına ve siberkondri davranışının pekişmesine yol açabilir. Bu nedenle hekimlerin, özellikle genç ve çocuğu olmayan kadınlarda güvenilir çevrim içi kaynaklar önermesi ve yeterli bilgilendirme yapması büyük önem taşımaktadır.

Bu çalışma bildiğimiz kadarıyla, adneksiyal kitle tanısı alan hastalarda siberkondri düzeyini değerlendiren ilk çalışma-

dır. Çalışmamızın bazı sınırlılıkları bulunmaktadır. Kesitsel tasarımı nedeniyle nedensel ilişki kurulması mümkün değildir. Hasta sayısının sınırlı olması nedeniyle çok değişkenli analiz yapılamamıştır; bu nedenle faktörlerin bağımsız etkileri değerlendirilememiştir. Ayrıca, çalışmanın homojen bir klinik popülasyonda yürütülmesi, genellenebilirliği kısıtlamaktadır. Gelecekte yapılacak daha geniş örneklemli araştırmalarda çok değişkenli istatistiksel analizlerin uygulanması, siberkondri düzeyini etkileyen faktörlerin bağımsız katkılarının daha net biçimde ortaya konulmasına katkı sağlayacaktır.

Sonuç

Adneksiyal kitle tanısı alan genç, çocuğu olmayan ve eğitim düzeyi yüksek kadınlarda siberkondri düzeyleri daha yüksektir. Malignite olasılığına dair belirsizlik ve cerrahi gereklilik konusundaki kaygılar, internetten bilgi arama davranışını artırabilir. Özellikle genç ve çocuğu olmayan kadınlarda, fertilite kaybı korkusu ve toplumsal beklentiler bu davranışı ve beraberinde gelişen anksiyeteyi pekiştirebilir. Hekimlerin bu hasta grubunu doğru bilgilendirme ve yönlendirmesi hastaların kaygı düzeyini azaltabilir. Daha geniş ve farklı popülasyonlarda yapılacak ileri çalışmalar, bu bulguların doğrulanmasına katkı sağlayacaktır.

Yazar Katkı Beyanı

Çalışma tasarımı ve kavramsallaştırma: **Hande Nur Öncü, Vakkas Korkmaz;** veri toplama: **Gökçen Ege, Neslihan Öztürk, Candost Hanedan;** sonuçların analizi ve yorumlanması: **Hande Nur Öncü;** aday makalenin hazırlanması: **Hande Nur Öncü.** Yazarlar sonuçları gözden geçirdi ve makalenin son halini onayladı.

Çıkar Çatışması

Yazarlar, herhangi bir çıkar çatışması olmadığını beyan eder.

Finansal Destek

Yazar(lar), çalışmanın herhangi bir finansman almadığını beyan eder.

Etik Kurul Onayı

Çalışma, Ankara Etlik Şehir Hastanesi Klinik Araştırmalar Etik Kurulu tarafından onaylanmıştır (Tarih: 13.05.2025, sayı: AEŞH-BADEK2-2025-067).

KAYNAKLAR

1. Starcevic V, Berle D. Cyberchondria: towards a better understanding of excessive health-related Internet use. Expert Rev Neurother. 2013;13(2):205-213. https://doi.org/10.1586/ern.12.162
2. Schuster AM, Ghaiumy Anaraky R, Cotten SR. Online health information seeking and the association with anxiety among older adults. Front Public Health. 2023;11:1076571. https://doi.org/10.3389/fpubh.2023.1076571
3. McMullan RD, Berle D, Arnáez S, Starcevic V. The relationships between health anxiety, online health information seeking, and cyberchondria: systematic review and meta-analysis. J Affect Disord. 2019;245:270-278. https://doi.org/10.1016/j.jad.2018.11.037
4. White RW, Horvitz E. Cyberchondria: studies of the escalation of medical concerns in Web search. ACM Trans Inf Syst. 2009;27(4):23. https://doi.org/10.1145/1629096.1629101
5. McElroy E, Shevlin M. The development and initial validation of the Cyberchondria Severity Scale (CSS). J Anxiety Disord. 2014;28(2):259-265. https://doi.org/10.1016/j.janxdis.2013.12.007
6. McElroy E, Kearney M, Touhey J, Evans J, Cooke Y, Shevlin M. The CSS-12: development and validation of a short-form version of the Cyberchondria Severity Scale. Cyberpsychol Behav Soc Netw. 2019;22(5):330-335. https://doi.org/10.1089/cyber.2018.0624
7. Pavlik EJ, Ueland FR, Miller RW, Ubellacker JM, DeSimone CP, Elder J, et al. Frequency and disposition of ovarian abnormalities followed with serial transvaginal ultrasonography. Obstet Gynecol. 2013;122(2 Pt 1):210-217. https://doi.org/10.1097/AOG.0b013e318298def5
8. Borgfeldt C, Andolf E. Transvaginal sonographic ovarian findings in a random sample of women 25-40 years old. Ultrasound Obstet Gynecol. 1999;13(5):345-350. https://doi.org/10.1046/j.1469-0705.1999.13050345.x
9. Castillo G, Alcázar JL, Jurado M. Natural history of sonographically detected simple unilocular adnexal cysts in asymptomatic postmenopausal women. Gynecol Oncol. 2004;92(3):965-969. https://doi.org/10.1016/j.ygyno.2003.11.029
10. National Cancer Institute. Surveillance, Epidemiology, and End Results (SEER) Program: Cancer Stat Facts—Ovarian Cancer [Internet]. Bethesda (MD): National Cancer Institute; 2025 [cited 2026 Mar 11]. Available from: https://seer.cancer.gov/statfacts/html/ovary.html
11. Yorgancıoğlu Tarcan G, Karahan A, Sebik NB. Kısa form siberkondri (CSS-12) ciddiyet ölçeğinin geçerlik ve güvenirliği: sağlık bilişimi alanına özel bir uygulama. Sağlık ve İnovasyon Dergisi. 2023;26(1):207-218.
12. Karakaş N, Tekin Ç, Bentli R, Demir E. Cyberchondria, COVID-19 phobia, and well-being: a relational study on teachers. Med Lav. 2022;113(3):e2022027. https://doi.org/10.23749/mdl.v113i3.12661
13. Varer Akpinar C, Mandiracioglu A, Ozvurmaz S, Kurt F, Koc N. Cyberchondria and COVID-19 anxiety and internet addiction among nursing students. Curr Psychol. 2023;42(3):2406-2414. https://doi.org/10.1007/s12144-022-04057-z
14. Serra-Negra JM, Paiva SM, Baptista AS, Cruz AJ, Pinho T, Abreu MHNG. Cyberchondria and associated factors among Brazilian and Portuguese dentists. Acta Odontol Latinoam. 2022;35(1):45-50. https://doi.org/10.54589/aol.35/1/45
15. Topkara Sucu S, Kolomuç Gayretli T, Küçükkayıkçı AS, et al. Cyberchondria levels in adolescent patients with polycystic ovary syndrome in the digital age. J Pediatr Adolesc Gynecol. 2024;37(6):569-573. https://doi.org/10.1016/j.jpag.2024.08.006
16. Turhan Cakir A. Cyberchondria levels in women with human papilloma virus. J Obstet Gynaecol Res. 2022;48(10):2610-2614. https://doi.org/10.1111/jog.15354

17. Öncü HN, Ege G, Sucu ST, Öztürk N, Hanedan C, Korkmaz V. The impact of human papillomavirus 16/18 positivity on cyberchondria levels in women undergoing colposcopy. Rev Assoc Med Bras (1992). 2025;71(5):e20241692. https://doi.org/10.1590/1806-9282.20241692

18. Gulec Satir D, Bakir S. The relationship between cyberchondria and health anxiety in pregnant women. J Consum Health Internet. 2024;28(2):91-103. https://doi.org/10.1080/15398285.2024.2355608

19. Šoštarić M, Jokić-Begić N, Vukušić Mijačika M. Can't stop, won't stop—understanding anxiety's role in cyberchondria among pregnant women. Women Health. 2024;64(2):185-194. https://doi.org/10.1080/03630242.2024.2308525

20. Sansakorn P, Mushtaque I, Awais-E-Yazdan M, Dost MKB. The relationship between cyberchondria and health anxiety and the moderating role of health literacy among the Pakistani public. Int J Environ Res Public Health. 2024;21(9):1168. https://doi.org/10.3390/ijerph21091168

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):122-132
DOI: 10.29058/mjwbs.1798429

Sensory perception profiles in children with epilepsy: A prospective controlled study

Epilepsili çocuklarda duyusal algı profilleri: Prospektif kontrollü bir çalışma

Seda Ayaz Taş[1], Merve Başol Göksülük[2], Özge Karanlık Özcan[1], Seda Yakıt Yeşilyurt[3], Sedef Kestellioğlu[4], Sezen Tezcan[1], Ayşegül Danış[5]

[1]Department of Physiotherapy and Rehabilitation, Bolu Abant İzzet Baysal University, Faculty of Health Sciences, Bolu, Türkiye
[2]Department of Biostatistics, Sakarya University, Faculty of Medicine, Sakarya, Türkiye
[3]Department of Physiotherapy and Rehabilitation, İzmir University of Economics, Faculty of Health Sciences, İzmir, Türkiye
[4]Department of Pediatrics, Bolu Abant İzzet Baysal University, Faculty of Medicine, Bolu, Türkiye
[5]Department of Pediatric Neurology, Bolu Abant İzzet Baysal University, Faculty of Medicine, Bolu, Türkiye

Cite this article as: *Ayaz Taş S et al. Sensory perception profiles in children with epilepsy: A prospective controlled study. Med J West Black Sea. 2026;10(1):122-132.*

ABSTRACT

Aim: Epilepsy is a chronic neurological disorder that disrupts the functional organization of the developing brain, with consequences extending beyond recurrent seizures to cognitive and sensory processes. This study aims to elucidate the effects of epilepsy on sensory perception by comparing the sensory profiles of children with epilepsy to those of their typically developing peers.

Material and Methods: In this a prospective controlled study, the sample size was determined based on an a priori power analysis, and a total of 49 children aged 5-12 years (25 with epilepsy and 24 typically developing controls), selected using convenience sampling, were included in the study. Demographic and clinical data were collected from the participants, and sensory perception was evaluated using the validated Turkish version of the Sensory Processing Measure-Home Form. The scale comprises the domains of social participation, vision, hearing, touch, taste and smell, body awareness, balance and motion, and planning and ideas. Group differences were analyzed using appropriate statistical methods with adjustments for potential confounders.

Results: No significant differences were observed between children with epilepsy and healthy controls in total sensory processing scores ($p>0.05$). Deficits in taste and smell perception were significantly associated with an increased body mass index ($p=0.010$). Sensory profiles were found to be primarily influenced by clinical sub-components, including structural brain abnormalities, antiepileptic drug use, and seizure timing. These findings indicate a heterogeneous sensory structure shaped by specific clinical and metabolic determinants.

Conclusion: Sensory processing and perception in pediatric epilepsy are influenced by structural, metabolic, and clinical variables rather than the diagnosis alone. It is recommended that individualized strategies considering these parameters be adopted in clinical management.

Keywords: Antiepileptic drugs, children, epilepsy, perception, sensory

ÖZ

Amaç: Epilepsi, gelişmekte olan beynin işlevsel organizasyonunu bozan ve yalnızca tekrarlayan nöbetlerle sınırlı kalmayıp bilişsel ve duyusal süreçleri de etkileyen kronik bir nörolojik bozukluktur. Bu çalışma, epilepsili çocukların duyusal algı profillerini tipik gelişen akranlarıyla karşılaştırarak epilepsinin duyusal algı süreçleri üzerindeki etkisini ortaya koymayı amaçlamaktadır.

Gereç ve Yöntemler: Bu prospektif kontrollü çalışmada örneklem büyüklüğü ön güç analizine dayalı olarak belirlenmiş; 5-12 yaş aralığında, kolayda örnekleme yöntemiyle seçilen toplam 49 çocuk (25 epilepsi tanılı, 24 tipik gelişen kontrol) çalışmaya dahil edildi. Katılımcılardan demografik ve klinik veriler toplandı; duyusal algı, Duyusal İşleme Ölçeği-Ev Formu'nun Türkçe geçerli-güvenilir versiyonuyla değerlendirildi. Ölçek; sosyal katılım, görme, işitme, dokunma, tat-koku, vücut farkındalığı, denge-hareket ve planlama alanlarını içermektedir. Grup farkları, karıştırıcılar için düzeltilerek uygun istatistiksel yöntemlerle analiz edildi.

Bulgular: Epilepsili çocuklar ile kontroller arasında toplam duyusal işlemleme skorlarında anlamlı fark saptanmamıştır ($p>0.05$). Tat ve koku algısındaki zayıflık artmış vücut kitle indeksi ile ilişkili bulunmuştur ($p=0.010$). Duyusal profillerin; yapısal beyin farklılıkları, ilaç kullanımı ve nöbet zamanlaması gibi klinik alt bileşenlerden etkilendiği saptanmıştır. Bulgular, klinik faktörlere göre şekillenen heterojen bir duyusal yapıya işaret etmektedir.

Sonuç: Çocukluk çağı epilepsisinde duyusal işlemleme ve algı, tanıdan ziyade yapısal, metabolik ve klinik değişkenlerden etkilenmektedir. Klinik yönetimde bu parametreleri dikkate alan bireyselleştirilmiş stratejilerin benimsenmesi önerilmektedir.

Anahtar Kelimeler: Antiepileptik ilaçlar, çocuklar, epilepsi, algı, duyusal

Corresponding Author: Özge Karanlık Özcan ✉ ozgekaranlik@gmail.com **Received:** 07.10.2025 **Revision:** 09.12.2025 **Accepted:** 06.02.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Overall sensory processing did not differ between children with epilepsy and typically developing peers; however, taste and smell perception was associated with body mass index.
- Structural brain abnormalities and seizure distribution play a decisive role in shaping sensory profiles in children with epilepsy.
- Detailed assessment of sensory profiles is essential for clinical follow-up and the planning of individualized management strategies.

INTRODUCTION

Epilepsy is a chronic disorder characterized by recurrent seizures, arising from abnormal and synchronized neuronal discharges, which have two causes: neuronal hyperactivity, neurotransmitter imbalances, and impairments in energy metabolism in the brain (1-3). Epileptic seizures disrupt the functional organization of the developing brain, leading to various functional limitations in motor coordination, cognitive functions, sensory processing, and social interaction (4,5). The involvement of sensory cortices such as the parietal, temporal, or occipital lobes in seizures may lead to disruptions in the central processing and perception of sensory information. This indicates that in children with epilepsy, effects are felt not only in the responses to stimuli but also in the perceptual components—such as recognition, discrimination, and interpretation of sensory stimuli (6-8).

Sensory perception is a neurodevelopmental process that enables the central nervous system to regulate and integrate a wide range of environmental inputs: visual, auditory, tactile, proprioceptive, vestibular, gustatory, and olfactory, thereby rendering them meaningful (9,10). Visual perception plays a central role in cognitive processes such as object recognition, spatial orientation, and motor planning; these processes are mediated by different neural networks, enabling the individual to interact harmoniously with their environment (11-14). Auditory perception is a determining factor in distinguishing environmental sounds, language development, and social communication, and its deficiencies may lead to limitations in language and communication skills (14,15). Tactile and proprioceptive perception contribute to the maintenance of body awareness, postural control, and motor coordination, whereas vestibular perception supports balance and spatial orientation processes. Deficiencies in these domains may adversely affect motor control and daily life activities (16,17). Gustatory and olfactory perception play a role in food choices, emotional memory, and the recognition of environmental stimuli, therefore, disruptions in these perceptual processes may negatively affect eating behaviors, mood, and social interaction (18-20).

The healthy functioning of these sensory perception processes is critically important for children's independence in daily life activities, harmony in social interactions, and cognitive development; however, neurological disorders such as epilepsy may impair the efficiency of these mechanisms, leading to various limitations. In children with epilepsy, this process can be affected by reductions in inhibitory circuits along with alterations in synaptic and cortical networks, resulting in behavioral problems in daily life, excessive or insufficient responses to sensory stimuli, and difficulties in adapting to social life (21-24). Therefore, identifying sensory perception impairments in children with epilepsy may contribute to creating a more holistic approach to co-occurring neurodevelopmental problems and the development of individualized intervention strategies. The aim of this study is to reveal the impact of epilepsy on sensory perception processes through a comparison of profiles of children with epilepsy and their typically developing peers.

MATERIAL and METHODS

Study Design

This study was designed as a prospective controlled study.

Sample Size Calculation

No study directly assessing sensory perception in the target population was identified in the literature. However, previous research indicates that sensory perception and sensory modulation are closely related constructs and may produce comparable outcomes. Therefore, a study assessing sensory modulation was used as a reference. In that study, a large effect size of d = 1.31 was reported between patient and control groups (25). Based on this effect size, the required sample size was calculated for an independent samples t test with 90 percent power and a two sided significance level of 5 percent. The analysis indicated that at least 14 participants per group were required. Assuming a potential dropout rate of 20 percent, the sample size was increased to 17 participants per group, resulting in a minimum total sample size of 34 participants. Sample size calculations were performed using G*Power version 3.1.9.7.

Patients

A total of 49 children aged 5-12 years were included in this study, comprising an epilepsy group (n=25) and a typically developing control group (n=24). The participants were recruited through convenience sampling based on their willingness to participate. Epilepsy diagnosis was established by a pediatric neurologist based on clinical evaluation and electroencephalography (EEG) confirmation, in accordance with ILAE criteria. Children were required to have had at least two unprovoked epileptic seizures to meet the diagnostic criteria. The inclusion criteria for children with epilepsy were being within the 5-12 age range and having a documented history of epileptic seizures. For the control group, eligibility required being in the same age range and including clinically healthy children with no neurological, developmental, sensory, or chronic medical conditions, as confirmed through parent-reported medical history and available health records. During the research process, children were excluded if they were found to have missing data or neuromuscular or neurodevelopmental disorders (e.g., cerebral palsy, hydrocephalus, autism spectrum disorder). Convenience sampling continued throughout the study period and involved recruiting children with epilepsy and typically developing peers who fulfilled the inclusion criteria.

Data Collection Process

The data collection process was carried out between February 2025 and July 2025 by the research team from the Department of Physiotherapy and Rehabilitation at Bolu Abant İzzet Baysal University. All assessments were conducted face-to-face in a quiet, distraction-free room within the university's clinical practice area, with parents present during each stage of the procedure.

Demographic information (age, height, body weight, and sex), medical history, seizure characteristics, seizure frequency, and antiepileptic medication use were obtained through parent interviews and by reviewing the children's hospital medical records. No additional clinical examinations, laboratory tests, or neurological assessments were performed for the purposes of the study; only height and weight measurements were taken on-site using standard procedures.

Sensory perception levels were assessed using the Sensory Processing Measure-Home Form, completed by parents. During the administration of the questionnaire, the research team provided clarification when needed to ensure parents fully understood the items; however, all responses were given independently by the parents. All forms were completed within the assessment environment, and the research team verified the accuracy and completeness of the collected data.

Ethical Considerations

Ethical approval for the study was obtained from the Clinical Research Ethics Committee of Bolu Abant İzzet Baysal University (Decision No: 2024/387, Date: 07.01.2025). The study was registered at ClinicalTrials.gov with the registration number NCT06818357. In accordance with the principles of the Declaration of Helsinki, the research was conducted between February 2025 and July 2025 in the Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Bolu Abant İzzet Baysal University. After information about the study was provided, verbal and written informed consent was obtained from the participating children and their parents.

Measurements

Additional Information on Children's Demographic and Epilepsy Characteristics

In addition to general demographic information such as age, sex, body weight, and height, a parents completed a structured questionnaire in order to comprehensively assess the clinical characteristics of children with epilepsy. This involved collecting data on the use of antiepileptic drugs, the name of the medication used, treatment response status (whether seizure control was achieved), seizure frequency within the past year, the time of day when seizures were more common, and their relationship to sleep (occurring exclusively during sleep, exclusively during wakefulness, or in both states). Furthermore, EEG and magnetic resonance imaging (MRI) findings of children with epilepsy were classified as either "normal" or "abnormal" by a pediatric neurologist.

Sensory Processing and Perception

The Sensory Processing Measure, developed by Parham and Ecker, is a norm-referenced assessment tool designed to evaluate sensory processing, planning, and social participation in children aged 5-12 years, based on parent or teacher reports. The scale consists of structured questions that provide a holistic perspective on the child's sensory functioning within home, school, and community settings. In the present study, children's sensory processing and perception characteristics were assessed using parents' responses to the Home Form of the Sensory Processing Measure. The scale includes a total of 75 items, each rated on a four-point Likert scale ranging from "never" (1) to "always" (4), based on parental observations.

In the Sensory Processing Measure, higher scores indicate greater sensory processing difficulties, whereas lower scores reflect more typical sensory functioning. Raw scores are converted into standardized T-scores, with values of 40-59 considered typical, 60-69 indicating mild to moderate difficulties, and scores ≥70 reflecting definite sensory process-

ing dysfunction. Each sensory domain—Social Participation, Vision, Hearing, Touch, Taste and Smell, Body Awareness, Balance and Motion, and Planning and Ideas—generates its own raw score and corresponding T-score. Because each subscale is evaluated independently, domain-specific interpretations can be made rather than relying solely on a total sensory score. In this study, between-group comparisons were conducted using the scores obtained from all relevant subdomains (26,27). The tool used in our study was Turkish version of the Sensory Processing Measure-Home Form, which has established validity and reliability (28).

Statistical Analysis

Data analyses were performed using IBM SPSS Statistics ver 26 and R ver 4.0.5. For descriptive statistics, numerical variables were summarized as mean ± standard deviation (SD) or as median and minimum-maximum values, while categorical variables were presented as frequencies and percentages. The normality of the data was evaluated using both graphical methods (e.g., Q-Q plots, histograms) and hypothesis testing (Shapiro-Wilk test). Comparisons of numerical variables between two groups were conducted with either the independent samples t-test or the Mann-Whitney U test, depending on distributional assumptions. Group comparisons for categorical variables were performed using appropriate chi-square tests (Pearson's chi-square test, Fisher's exact test, etc.). To examine differences between groups in total and subscale scores while accounting for potential confounders (age, sex, and BMI), a linear regression model was employed. Model assumptions were checked using residual plots, including residuals versus fitted values and normal Q-Q plots. Robust methods were applied to estimate the standard error. Effect sizes were reported alongside p-values. For independent samples t-tests, Cohen's d was used and interpreted as 0.2=small, 0.5=medium, and 0.8=large. For the Mann-Whitney U test, effect size r was calculated as $r=z/\sqrt{n}$ and interpreted as 0.10=small, 0.30=medium, and 0.50=large. For categorical data analyses, Phi or Cramer's V coefficients were reported, with the same interpretation thresholds (0.10=small, 0.30=medium, 0.50=large). In R, the following packages were used for analyses: lmtest (29), emmeans (30), and sandwich (31). A p-value of<0.05 was considered statistically significant.

RESULTS

Patient Selection

A total of 49 children aged between 5 and 12 (mean= 8.67±2.11) years were included in the study. The patient group consisted of 25 children diagnosed with epilepsy, and the control group, 24 typically developing children. No statistically significant difference was observed between groups in terms of age and sex distribution (p=0.131 and p=0.058, respectively). However, the effect size values indicated approaching moderate difference for age (d=0.439) and sex (Φ=0.271). Children in the patient group had significantly higher mean weight and height compared to the control group (p=0.016 and p=0.037, respectively), with medium-to-large effect sizes (d=0.717 and d=0.614, respectively). No significant difference was observed in body mass index (BMI) between groups (p=0.077), although the effect size suggested a moderate difference (d=0.514). A significant difference was also found with respect to children's place of residence (p=0.014): 87.5% in the control group lived in the city center, compared with 56% of the patient group. The participants' demographic and clinical characteristics are presented in Table 1.

Sensory Processing and Perception

The comparison results of the sensory processing subscales and the total sensory system scores between groups are presented in Table 2. Group comparisons were performed both without adjustment and after adjusting for potential confounding variables, including age, sex, and BMI and place of residence. Due to the small number of participants residing in villages, the district and village categories were combined and included in the model as 1 = province and 2 = district plus village. Both unadjusted and adjusted analyses, no statistically significant differences were observed between groups in subscale and total scores. However, BMI was negatively associated with taste and smell scores, and this relationship reached statistical significance (b = -0.138, (95% CI: -0.241; -0.034); p = 0.010).

The comparisons of sensory processing subscale scores and total sensory system scores according to different clinical characteristics of the patient group are presented in Table 4. A moderate difference was observed between patients using antiepileptic medication and those not using medication in terms of Body Awareness and Balance and Motion scores, with higher scores in patients receiving antiepileptic treatment (r = 0.312, p = 0.132 and r = 0.292, p = 0.154, respectively). A moderate difference was also found between patients who responded to antiepileptic treatment and those who did not with respect to Hearing scores, with higher scores observed in patients with a continued treatment response (r = 0.342, p = 0.107). In addition, a moderate difference was identified between sleep related seizure groups in terms of Social Participation scores, with lower Social Participation scores observed in both the asleep and awake groups (r = 0.342, p = 0.107). Furthermore, moderate differences were observed between patients with normal and abnormal MRI findings in Vision, Taste and Smell, and Planning and Ideas scores, with higher scores in the non normal MRI group. Moreover, a statistically significant difference was found between MRI groups with respect to Total Sensory Systems scores, with significantly higher scores in the non normal group.

Table 1: Demographic and Clinical characteristics of participants (n=49)

	Group			
	Patient (n=25)	**Control (n=24)**	**ES**	**p**
Age	9.12 ± 2.07	8.21 ± 2.08	0.439	0.131*
Sex				
Female	11 (44.0)	17 (70.8)	0.271	0.058**
Male	14 (56.0)	7 (29.2)		
Weight (kg)	40.83 ± 17.78	30.69 ± 8.91	0.717	0.016*
Height (cm)	140.92 ± 15.80	131.29 ± 15.53	0.614	0.037*
BMI (kg/cm^2)	19.63 ± 5.46	17.47 ± 2.20	0.514	0.077*
Place of Residence				
Province	14 (56.0)	21 (87.5)	0.377	0.014**
District	7 (28.0)	3 (12.5)		
Village	4 (16.0)	0 (0.0)		
Type of Epilepsy				
Focal	11 (44.0)			
Generalize	14 (56.0)			
Use of Epileptic Medication				
Yes	21 (84.0)			
No	4 (16.0)			
Medication (n=21)[A]				
Antiexcitatory drugs	3 (14.3)			
Sodium channel-blocking AED				
Drugs enhancing inhibition				
AED acting on GABA	2 (9.6)			
AED with multiple mechanisms of action	6 (28.6)			
Other mechanisms	6 (28.6)			
Response to Antiepileptic Medication				
Under Control	22 (88.0)			
Continue	3 (12.0)			
Seizure Status in the Last Year				
Yes	13 (52.0)			
No	12 (48.0)			
Relationship Between Seizure and Sleep				
Asleep	9 (36.0)			
Awake	11 (44.0)			
Both	5 (20.0)			
EEG Results				
Normal	9 (36.0)			
Non-Normal	16 (64.0)			
MRG Results (n=24)				
Normal	20 (80.0)			
Non- Normal	4 (16.0)			

Descriptive statistics were defined as mean±standard deviation or median (minimum - maximum) values for numerical variables, and frequency (percentage) for categorical variables. **AED:** antiepileptic drugs. * Independent t test (with Cohen'd effect size as d=0.2 small, d=0.5 medium, and d ≥ 0.8 large), ** Chi-square tests (Pearson, Yates Correction or Fisher's Exact test wirh Phi/Cramer V effect size as d=0.1 small, d=0.3 medium, and d ≥ 0.5 large). *** Mann-Whitney U test (with formula as d=0.1 small, d=0.3 medium, and d ≥ 0.5 large), **BMI:** body mass index, **GABA:** gamma-aminobutyric acid, **MRI:** magnetic resonance imaging, **ES:** effect size, **[A]:** Multiple Selection

DISCUSSION

This study aimed to compare the sensory processing and perception profiles of children with epilepsy to those of their typically developing peers, identifying the neurological and clinical determinants underlying these processes. Our findings demonstrate that children with epilepsy exhibit performance levels comparable to their peers in both total sensory

Table 2: Comparison of sensory processing subscale scores and total sensory system scores between groups

	Group			
	Patient (n=25)	**Control (n=24)**	**p^{unAdj}**	**p^{Adj}**
Social participation	35.36 ± 5.22	34.5 ± 5.57	0.588	0.629
Vision	14.96 ± 3.32	15.04 ± 3.91	0.939	0.579
Hearing	10.28 ± 2.65	10.42 ± 2.95	0.868	0.484
Touch	15.56 ± 3.75	14.87 ± 3.78	0.538	0.621
Taste and Smell	6.12 ± 1.39	5.46 ± 0.83	0.053	0.116
Body Awareness	12.76 ± 3.31	13.13 ± 4.51	0.754	0.983
Balance and Motion	15.80 ± 3.07	17.04 ± 4.67	0.289	0.105
Planning and Ideas	12.08 ± 3.90	11.58 ± 4.77	0.699	0.594
Total Sensory Systems	75.48 ± 12.97	75.35 ± 16.64	0.976	0.658

PunAdj: Unadjusted regression model results, **PAdj:** Adjusted regression model results for age, sex, BMI, and place of residence

systems and sensory subdomains. This similarity suggests that epilepsy does not invariably lead to widespread sensory impairment in children, who may maintain functionality levels close to their peers across various sensory domains. Notably, however, deficits in taste and smell perception were found to be associated with an increase in Body Mass Index (BMI), independent of the epilepsy diagnosis itself. Furthermore, it was determined that the sensory profiles are primarily shaped by clinical sub-components—such as structural brain differences, antiepileptic drug (AED) use, and the timing of seizures—rather than the diagnosis of epilepsy alone. These results confirm that sensory processes in pediatric epilepsy do not follow a homogeneous course but instead present a heterogeneous structure tailored to each child's clinical presentation.

Sub-domain analyses revealed that impairments in taste and smell perception were correlated with increased BMI. These findings align with existing literature reporting that visceral adiposity and high BMI values weaken odor discrimination and perception skills (32-34). In the literature, structural mechanisms such as a reduced number of taste buds and deficiencies in taste identification abilities accompanying obesity in children and adolescents have been highlighted (35,36) Moreover, impaired taste perception may trigger a cycle of high-energy food consumption and obesity by delaying the sensation of satiety (37). In light of these data, we suggest that impairments in taste and smell perception identified in childhood epilepsy should be evaluated within the framework of the general metabolic profile and physical growth parameters rather than as a direct reflection of seizure activity.

Another significant finding of our study is that children with abnormal brain MRI findings had significantly higher total sensory system scores compared to the group with normal findings. This suggests that sensory processing and perception disorders observed in childhood epilepsy may be closely linked to structural differences and neuroanatomical abnormalities in the central nervous system. Existing literature demonstrates that structural volume losses in sensory integration centers such as the thalamus, insula, and putamen, along with the disruption of functional connectivity between these regions and cortical networks, impair sensory organization (38-40). Given that thalamocortical and insular networks serve as primary sensory relay stations where environmental stimuli are filtered and integrated (41), it is expected that abnormalities in these structures lead to impairments in sensory processing and perception. The high sensory scores—indicating dysfunction—found in children with MRI abnormalities suggest that these patients are insufficient in processing sensory information and that structural damage exerts a holistic disruptive effect on sensory processes. Thus, we believe that radiological anomalies are a critical parameter to consider in clinical evaluations as they may play a restrictive role in an individual's sensory performance.

Clinical subgroup analyses revealed that children using antiepileptic drugs (AEDs) had higher body awareness, balance, and motion scores compared to the non-medicated group. The effects of AEDs on neurotransmitter systems and brain regions where sensory information is regulated are well-documented (42,43). While it is theoretically anticipated that these neurochemical changes might negatively impact the processing of body awareness and vestibular perception systems (44), the literature remains inconclusive. Our findings suggest that AED use may have restrictive effects on the processing and perception of sensory stimuli in these two specific domains. Although current data are limited for establishing a direct cause-and-effect relationship, these differences highlight the clinical importance of monitoring perceptual processes in children receiving AED therapy. Consequently, further research with larger samples is needed to isolate the specific effects of medications on sensory-perceptual systems.

Table 3: Regression Coefficients and 95% Confidence Intervals for sensory processing subscale scores and total sensory system scores

	Factors	Coefficient	95 % Confidence Interval for Coefficient		P
			LL	UL	
Social participation	Group	0.968	-3.045	4.981	0.629
	Age	0.074	-0.781	0.928	0.863
	Gender	1.160	-2.607	4.927	0.538
	BMI	0.098	-0.295	0.419	0.618
	Place of Res.	-2.208	-8.029	3.614	0.449
Vision	Group	-0.848	-3.908	1.211	0.579
	Age	-0.084	-0.749	0.582	0.801
	Gender	0.533	-1.936	3.002	0.666
	BMI	0.001	-0.374	0.375	0.998
	Place of Res.	2.220	-1.158	5.598	0.192
Hearing	Group	-0.664	-2.563	1.235	0.484
	Age	0.075	-0.466	0.617	0.781
	Gender	0.839	-1.053	2.732	0.376
	BMI	-0.197	-0.398	0.004	0.055
	Place of Res.	2.071	-0.417	4.560	0.100
Touch	Group	-0.636	-3.217	1.944	0.621
	Age	0.533	-0.067	1.132	0.080
	Gender	1.757	-0.481	3.995	0.121
	BMI	-0.224	-0.474	0.026	0.077
	Place of Res.	2.191	-1.141	5.523	0.182
Taste and Smell	Group	0.607	-0.156	1.371	0.116
	Age	0.117	-0.043	0.277	0.147
	Gender	0.066	-0.608	0.740	0.845
	BMI	-0.138	-0.241	-0.034	**0.010**
	Place of Res.	0.709	-0.274	1.691	0.153
Body Awareness	Group	-0.029	-2.765	2.707	0.983
	Age	-0.301	-0.952	0.350	0.357
	Gender	-0.822	-3.332	1.687	0.512
	BMI	-0.283	-0.581	0.015	0.062
	Place of Res.	2.417	-1.875	6.709	0.262
Balance and Motion	Group	-2.031	-4.501	0.439	0.105
	Age	0.128	-0.442	0.699	0.652
	Gender	-0.133	-2.322	2.056	0.903
	BMI	-0.079	-0.363	0.203	0.573
	Place of Res.	2.786	-0.827	6.405	0.128
Planning and Ideas	Group	-0.867	-4.122	2.388	0.594
	Age	0.091	-0.542	0.724	0.774
	Gender	1.574	-1.701	4.850	0.338
	BMI	0.053	-0.331	0.437	0.781
	Place of Res.	2.367	-2.669	7.402	0.349
Total Sensory Systems*	Group	0.039	-0.136	0.214	0.658
	Age	-0.013	-0.054	0.028	0.520
	Gender	-0.051	-0.194	0.093	0.482
	BMI	0.014	-0.002	0.031	0.087
	Place of Res.	-0.158	-0.373	0.057	0.146

LL: lower limit, **UP:** upper limit, **BMI:** body mass index. * Values were transformed using the reciprocal (1/y) transformation.

While no direct relationship was found between seizure timing and general sensory processing, a notable differentiation was observed in social participation. In our study, children who experienced seizures both during sleep and wakefulness had lower social participation scores (reflecting higher functional capacity in the scale used) compared

Table 4: Comparison several epilepsy features based on the sensory processing subscale scores and total sensory system scores

	SOC	VIS	HEA	TOU
Type of Epilepsy				
Focal (n=11)	37 (16 - 40)	14 (11 - 24)	10 (8 - 19)	15 (11 - 23)
Generalize (n=14)	36 (30 - 40)	14.5 (12 - 22)	9 (8 - 14)	13.5 (12 - 24)
ES	0.205	0.110	0.179	0.133
p	0.317	0.609	0.403	0.536
Use of Epileptic Medication				
Yes (n=21)	37 (28 - 40)	14 (11 - 24)	9 (8 - 13)	15 (11 - 24)
No (n=4)	31.5 (16 - 40)	15.5 (14 - 19)	11.5 (8 - 19)	12.5 (12 - 23)
ES	0.277	0.232	0.190	0.180
p	0.177	0.262	0.369	0.409
Response to Antiepileptic Medication				
Under Control (n=22)	37 (16 - 40)	14 (11 - 24)	9 (8 - 19)	14 (11 - 24)
Continue (n=3)	34 (28 - 37)	14 (11 - 22)	11.5 (8 - 19)	18 (13 - 18)
ES	0.270	0.008	**0.342**	0.178
p	0.206	0.969	0.107	0.398
Seizure Status in the Last Year				
Yes (n=13)	37 (16 - 40)	14 (11 - 22)	9 (8 - 19)	17 (11 - 24)
No (n=12)	37 (28 - 40)	14 (12 - 24)	9.5 (8 - 14)	13.5 (12 - 23)
ES	0.078	0.055	0.022	0.160
p	0.728	0.810	0.936	0.437
Relationship Between Seizure and Sleep				
Asleep (n=9)	37 (16 - 40)	15 (12 - 22)	9 (8 - 19)	17 (12 - 23)
Awake (n=11)	37 (32 - 40)	13 (11 - 24)	10 (8 - 13)	14 (11 - 24)
Both (n=5)	33 (28 - 38)	14 (14 - 17)	9 (8 - 14)	13 (12 - 18)
ES	**0.105**	0.004	0.076	0.077
p	0.116	0.385	0.851	0.862
EEG Results				
Normal (n=9)	37 (30 - 40)	14 (12 - 17)	9 (8 - 14)	14 (12 - 18)
Non-Normal (n=16)	37 (16 - 40)	14.5 (11 - 24)	9.5 (8 - 19)	14.5 (11 - 24)
ES	0.046	0.108	0.012	0.160
p	0.846	0.598	0.978	0.452
MRG Results				
Normal (n=20)	37 (28 - 40)	14 (11 - 24)	9 (8 - 14)	13.5 (11 - 24)
Non- Normal (n=4)	36 (16 - 39)	17.5 (13 - 22)	11 (9 - 19)	17.5 (13 - 23)
ES	0.094	**0.297**	0.269	0.251
p	0.682	0.157	0.210	0.241

	TNS	BOD	BAL	PLA	TOT
Type of Epilepsy					
Focal (n=11)	5 (5 - 8)	12 (10 - 19)	15 (11 - 21)	11 (9 - 17)	73 (56 - 109)
Generalize (n=14)	6 (5 - 10)	11 (10 - 20)	15 (12 - 24)	11 (9 - 22)	70.5 (62 - 100)
ES	0.081	0.039	0.011	0.129	0.055
p	0.727	0.851	0.979	0.536	0.809
Use of Epileptic Medication					
Yes (n=21)	6 (5 - 10)	11 (10 - 20)	15 (11 - 20)	11 (9 - 22)	72 (56 - 100)
No (n=4)	5.5 (5 - 8)	14.5 (11 - 19)	18.5 (14 - 24)	14 (9 - 21)	78 (67 - 109)
ES	0.039	**0.312**	**0.292**	0.190	0.111
p	0.858	0.132	0.154	0.369	0.592
Response to Antiepileptic Medication					
Under Control (n=22)	6 (5 - 10)	11 (10 - 20)	15.5 (11 - 24)	11 (9 - 22)	70.5 (56 - 109)
Continue (n=3)	6 (5 - 7)	12 (10 - 16)	15 (14 - 16)	12 (9 - 12)	75 (74 - 84)

Table 4: Continued

ES	0.035	0.172	0.084	0.000	0.201
p	0.906	0.906	0.723	>0.999	0.353
Seizure Status in the Last Year					
Yes (n=13)	6 (5 - 10)	11 (10 - 20)	15 (11 - 21)	11 (9 - 22)	73 (56 - 109)
No (n=12)	6 (5 - 8)	12 (10 - 19)	15.5 (13 - 24)	10.5 (9 - 21)	71 (62 - 91)
ES	0.040	0.050	0.203	0.095	0.005
p	0.852	0.810	0.320	0.650	0.979
Relationship Between Seizure and Sleep					
Asleep (n=9)	6 (5 - 9)	11 (10 - 19)	16 (12 - 21)	12 (9 - 22)	74 (68 - 109)
Awake (n=11)	6 (5 - 10)	12 (10 - 20)	15 (11 - 18)	9 (9 - 16)	70 (56 - 100)
Both (n=5)	5 (5 - 7)	11 (10 - 16)	16 (14 - 24)	11 (11 - 21)	73 (65 - 88)
ES	0.047	0.083	0.004	0.085	0.048
p	0.620	0.914	0.387	0.144	0.623
EEG Results					
Normal (n=9)	6 (5 - 9)	12 (10 - 18)	16 (14 - 24)	10 (9 - 22)	70 (67 - 88)
Non-Normal (n=16)	6 (5 - 10)	11 (10 - 20)	15 (11 - 21)	11 (9 - 19)	72.5 (56 - 109)
ES	0.048	0.116	0.274	0.017	0.005
p	0.846	0.598	0.187	0.934	>0.999
MRG Results					
Normal (n=20)	5 (5 - 10)	11 (10 - 20)	15 (11 - 24)	10.5 (9 - 21)	70 (56 - 100)
Non- Normal (n=4)	7 (6 - 9)	16.5 (10 - 19)	18 (14 - 21)	14.5 (10 - 22)	84.5 (75 - 109)
ES	**0.413**	0.277	0.255	**0.367**	**0.419**
p	0.056	0.210	0.241	0.081	**0.037**

Descriptive statistics were defined as median (minimum - maximum) Mann-Whitney U test (with formula $r=z/\sqrt{n}$ as d=0.1 small, d=0.3 medium, and d ≥ 0.5 large), **ES:** effect size,**SOC:** social participation**, VIS:** vision**, HEA:** hearing**, TOU:** touch, **EEG:** electroencephalography, **MRG:** magnetic resonance imaging

to those whose seizures were limited to a specific time frame. This suggests higher social adaptation capacity in this group. While the literature often reports that the unpredictable and continuous nature of seizures deepens social restriction, isolation, and perceived stigma (45,46), our results may be related to the adaptation and tolerance developed by families. The distribution of seizures throughout the day may have prompted families to be constantly prepared and adapt to varying conditions. This suggests that by learning to live with the reality of seizures, families may adopt a more flexible attitude, including the child in social life rather than isolating them. Therefore, high social participation despite continuous seizure activity may stem from the supportive role of family coping skills in preserving the child's social functionality.

Strengths

This study provides a significant contribution to the limited evidence base regarding sensory processing and perception profiles in childhood epilepsy. The utilization of the Sensory Profile-Home Form, a validated and reliable assessment instrument, underscores the methodological rigor of the research. The inclusion of an age-matched control group facilitated structured and robust comparisons between children with epilepsy and their typically developing peers. Furthermore, the prospective design allowed for the systematic collection of clinical and demographic data, thereby enhancing the consistency of study procedures. Notably, by shifting the focus from the diagnosis alone to clinical sub-components—such as medication use and MRI findings—this study offers a more nuanced understanding of the individual variations inherent in sensory processes.

Limitations

Despite its contributions, several limitations of this study should be acknowledged. The use of convenience sampling may limit the representativeness of the cohort and the generalizability of the findings to broader populations. Additionally, the relatively small sample size may have constrained the statistical power of the results. Sensory processing and perception were assessed exclusively through parent-report scales; while this is a standard approach in pediatric research, it remains susceptible to reporting bias. Furthermore, the variability in antiepileptic drug types and dosages presents a potential confounding factor that may influence sensory outcomes. Future research involving larger, multi-center samples, objective clinical assessment tools, and more homogeneous medication groups is warranted to provide a more comprehensive understanding of sensory functions in pediatric epilepsy.

Conclusion and Recommendations

This study demonstrates that sensory processing and perception in childhood epilepsy present a heterogeneous structure shaped by metabolic and structural factors. The findings indicate that an epilepsy diagnosis does not result in universal perceptual impairment; however, individual clinical variables significantly alter the clinical picture. In particular, structural brain differences and seizure distribution play a decisive role. In conclusion, epilepsy exerts multifaceted effects on sensory processing and developmental functions beyond seizure activity. Detailed assessment of sensory profiles is critical for clinical follow-up and the planning of individualized treatment strategies. Healthcare professionals are encouraged to adopt a holistic approach that encompasses the child's social participation and sensory-perceptual functionality.

Acknowledgments

The authors thank the children and their families for participating in the study.

Author Contributions

Study conception and design: **Seda Ayaz Taş;** data collection: **Sedef Kestellioğlu, Ayşegül Danış;** analysis and interpretation of results: **Merve Başol Göksülük;** draft manuscript preparation: **Seda Ayaz Taş, Merve Başol Göksülük, Özge Karanlık Özcan, Seda Yakıt Yeşilyurt, Sedef Kestellioğlu, Sezen Tezcan, Ayşegül Danış.** The author(s) reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Financial Support

The authors declare that the study received no funding.

Ethical Approval

The study protocol was approved by the Clinical Research Ethics Committee of Bolu Abant Izzet Baysal University (Decision No: 2024/387, Date: 07 January 2025). This study was registered at ClinicalTrials.gov (Identifier: NCT06818357). Informed assent was obtained from the participating children, and written informed consent was obtained from their parents or legal guardians. The privacy rights of all participants were protected.

REFERENCES

1. Green JL, dos Santos WF, Fontana ACC. Role of glutamate excitotoxicity and glutamate transporter EAAT2 in epilepsy: opportunities for novel therapeutics development. Biochem Pharmacol. 2021;193:114786. https://doi.org/10.1016/j.bcp.2021.114786
2. Boison D, Steinhäuser C. Epilepsy and astrocyte energy metabolism. Glia. 2018;66(6):1235-1243. https://doi.org/10.1002/glia.23247
3. Anwar H, Khan QU, Nadeem N, Pervaiz I, Ali M, Cheema FF. Epileptic seizures. Discoveries (Craiova). 2020;8(2):e110. https://doi.org/10.15190/d.2020.7
4. Wu J, Hu D, Zheng R, Jiang T, Gao F, Cao J. Information flow-based brain network analysis of healthy and epileptic syndromes in children. IEEE Sens J. 2024;24(12):19417-19427. https://doi.org/10.1109/JSEN.2024.3394046
5. Feng X, Piper RJ, Prentice F, Clayden JD, Baldeweg T. Functional brain connectivity in children with focal epilepsy: a systematic review of functional MRI studies. Seizure. 2024;117:164-173. https://doi.org/10.1016/j.seizure.2024.03.014
6. Stenroos P, Guillemain I, Tesler F, Montigon O, Collomb N, Stupar V, et al. EEG-fMRI in awake rats and whole-brain simulations show decreased brain responsiveness to sensory stimulations during absence seizures. eLife. 2024;12:e90318. https://doi.org/10.7554/eLife.90318
7. Panarese A, Vissani M, Meneghetti N, Vannini E, Cracchiolo M, Micera S, et al. Disruption of layer-specific visual processing in a model of focal neocortical epilepsy. Cereb Cortex. 2023;33(7):4173-4187. https://doi.org/10.1093/cercor/bhac335
8. Studer F, Laghouati E, Jarre G, David O, Pouyatos B, Depaulis A. Sensory coding is impaired in rat absence epilepsy. J Physiol. 2019;597(3):951-966. https://doi.org/10.1113/JP277117
9. Rodríguez-Armendariz E, Vela-Romero M, Galiana A. Sensory processing challenges in children with neurodevelopmental disorders and genetic conditions: an observational study. NeuroSci. 2024;5(3):339-353. https://doi.org/10.3390/neurosci5030023
10. Gigliotti F, Giovannone F, Belli A, Sogos C. Atypical sensory processing in neurodevelopmental disorders: clinical phenotypes in preschool-aged children. Children (Basel). 2024;11(7):875. https://doi.org/10.3390/children11070875
11. Galiana-Simal A, Vela-Romero M, Romero-Vela VM, Oliver-Tercero N, García-Olmo V, Benito-Castellanos PJ, et al. Sensory processing disorder: key points of a frequent alteration in neurodevelopmental disorders. Cogent Med. 2020;7(1):1787680. https://doi.org/10.1080/2331205X.2020.1787680
12. Cascio CJ. Somatosensory processing in neurodevelopmental disorders. J Neurodev Disord. 2010;2(2):62-69. https://doi.org/10.1007/s11689-010-9046-3
13. Monaco S, Malfatti G, Culham JC, Cattaneo L, Turella L. Decoding motor imagery and action planning in the early visual cortex. Neuroimage. 2020;218:116972. https://doi.org/10.1016/j.neuroimage.2020.116972
14. Alkhlouf AM, Che Omar M, Kazakzeh M. The role of early intervention in developing auditory perception skills in first-stage children. Int J Acad Res Bus Soc Sci. 2025;15(3):1100-1106. https://doi.org/10.6007/IJARBSS/v15-i3/22643
15. Liu P, Zhou J, Zhang L, Ji H, Xu J, Xu Q, et al. A longitudinal study on auditory processing development in preschool children with autism spectrum disorder: study protocol. BMC Psychiatry. 2024;24:723. https://doi.org/10.1186/s12888-024-06086-3
16. Lackner JR, DiZio P. Vestibular, proprioceptive, and haptic contributions to spatial orientation. Annu Rev Psychol. 2005;56:115-147. https://doi.org/10.1146/annurev.psych.55.090902.142023
17. Cullen KE. Vestibular processing during natural self-motion: implications for perception and action. Nat Rev Neurosci. 2019;20:346-363. https://doi.org/10.1038/s41583-019-0153-1

18. Loureiro M, Achargui R, Flakowski J, Van Zessen R, Stefanelli T, Pascoli V, et al. Social transmission of food safety depends on synaptic plasticity in the prefrontal cortex. Science. 2024;384:eadh1234. https://doi.org/10.1126/science.adh1234
19. Wu B, Eldeghaidy S, Ayed C, Fisk ID, Hewson L, Liu Y. Mechanisms of umami taste perception: from molecular level to brain imaging. Crit Rev Food Sci Nutr. 2022;62:7015-7024. https://doi.org/10.1080/10408398.2021.1902175
20. Boesveldt S, Parma V. The importance of the olfactory system in human well-being, through nutrition and social behavior. Cell Tissue Res. 2021;383:559-567. https://doi.org/10.1007/s00441-020-03367-7
21. Ianov Vitanov R, Achterberg J, Akarca D, Astle DE, Baker K. Synaptic function and sensory processing in ZDHHC9-associated neurodevelopmental disorder: a mechanistic account. bioRxiv. 2024. https://doi.org/10.1101/2024.03.28.587155
22. Bozzi Y, Casarosa S, Caleo M. Epilepsy as a neurodevelopmental disorder. Front Psychiatry. 2012;3:19. https://doi.org/10.3389/fpsyt.2012.00019
23. Lobato-Ruiz V, Romero-Ayuso D, Toledano-González A, Triviño-Juárez JM. Quality of life and parental stress related to executive functioning, sensory processing, and activities of daily living in children and adolescents with neurodevelopmental disorders. PeerJ. 2025;13:e19326. https://doi.org/10.7717/peerj.19326
24. Hornix BE, Havekes R, Kas MJH. Multisensory cortical processing and dysfunction across the neuropsychiatric spectrum. Neurosci Biobehav Rev. 2019;97:138-151. https://doi.org/10.1016/j.neubiorev.2018.11.013
25. Shahar E, Zlotnik S, Ravid S, Engel-Yeger B. Sensory processing disabilities in childhood-onset generalized epilepsy. J Pediatr Neurol. 2013;11(2):83-88. https://doi.org/10.3233/JPN-130569
26. Parham LD, Ecker C, Kuhaneck H, Henry D, Glennon T. Sensory Processing Measure (SPM) Manual. Los Angeles: Western Psychological Services; 2007.
27. Brown T, Almiento L, Yu ML, Bhopti A. The Sensory Processing Measure-Second Edition: a critical review and appraisal. Occup Ther Health Care. 2024;38(3):842-875. https://doi.org/10.1080/07380577.2023.2294174
28. Ozbakir M. Duyusal İşleme Ölçeği – Ev Formu'nun Türkiye koşullarına uyarlanması [master's thesis]. İstanbul: Marmara Üniversitesi; 2010.
29. Zeileis A, Hothorn T. Diagnostic checking in regression relationships. R News. 2002;2(3):7-10.
30. Lenth R. emmeans: Estimated marginal means, aka least-squares means [Internet]. 2025. Available from: https://rvlenth.github.io/emmeans/
31. Zeileis A, Köll S, Graham N. Various versatile variances: an object-oriented implementation of clustered covariances in R. J Stat Softw. 2020;95:1-36. https://doi.org/10.18637/jss.v095.i01
32. Velluzzi F, Deledda A, Onida M, Loviselli A, Crnjar R, Sollai G. Relationship between olfactory function and body mass index in normal weight healthy subjects and patients with overweight or obesity. Nutrients. 2022;14(6):1262. https://doi.org/10.3390/nu14061262
33. Melis M, Tomassini Barbarossa I, Crnjar R, Sollai G. Olfactory sensitivity is associated with body mass index and polymorphism in the voltage-gated potassium channel Kv1.3. Nutrients. 2022;14(23):4986. https://doi.org/10.3390/nu14234986
34. Fernandez-Garcia JC, Alcaide J, Santiago-Fernandez C, Roca-Rodriguez MM, Aguera Z, Baños R, et al. An increase in visceral fat is associated with a decrease in the taste and olfactory capacity. PLoS One. 2017;12(2):e0171204. https://doi.org/10.1371/journal.pone.0171204
35. Rohde K, Schamarek I, Blüher M. Consequences of obesity on the sense of taste: taste buds as treatment targets? Diabetes Metab J. 2020;44(4):509-528. https://doi.org/10.4093/dmj.2019.0254
36. Mameli C, Cattaneo C, Panelli S, Comandatore F, Sangiorgio A, Bedogni G, et al. Taste perception and oral microbiota are associated with obesity in children and adolescents. PLoS One. 2019;14(9):e0221656. https://doi.org/10.1371/journal.pone.0221656
37. Fathi M, Javid AZ, Mansoori A. Effects of weight change on taste function: a systematic review. Nutr J. 2023;22:50. https://doi.org/10.1186/s12937-023-00852-7
38. Alpcan A, Bayar Muluk N, Inal M, Kömürcü Erkmen SP, Şencan Z. Peripheral and central smell regions in children with epilepsy: an MRI evaluation. J Clin Neurosci. 2022;95:99-105. https://doi.org/10.1016/j.jocn.2021.11.041
39. Zhang Z, Liao W, Chen H, Mantini D, Ding JR, Xu Q, et al. Altered functional-structural coupling of large-scale brain networks in idiopathic generalized epilepsy. Brain. 2011;134(10):2912-2928. https://doi.org/10.1093/brain/awr223
40. Nuyts S, D'Souza W, Bowden SC, Vogrin SJ. Structural brain abnormalities in genetic generalized epilepsies: a systematic review and meta-analysis. Epilepsia. 2017;58(12):2025-2037. https://doi.org/10.1111/epi.13928
41. Mazzola L, Royet JP, Catenoix H, Montavont A, Isnard J, Mauguière F. Gustatory and olfactory responses to stimulation of the human insula. Ann Neurol. 2017;82(3):360-370. https://doi.org/10.1002/ana.25022
42. Brodie MJ, Besag F, Ettinger AB, Mula M, Gobbi G, Comai S, et al. Epilepsy, antiepileptic drugs, and aggression: an evidence-based review. Pharmacol Rev. 2016;68(3):563-602. https://doi.org/10.1124/pr.115.012021
43. Sills GJ, Rogawski MA. Mechanisms of action of currently used antiseizure drugs. Neuropharmacology. 2020;168:107966. https://doi.org/10.1016/j.neuropharm.2020.107966
44. Yuan X, Fu Z, Ji P, Guo L, Al-Ghamdy AO, Alkandiri A, et al. Selenium nanoparticles pre-treatment reverses behavioral, oxidative damage, neuronal loss and neurochemical alterations in pentylenetetrazole-induced epileptic seizures in mice. Int J Nanomedicine. 2020;15:6339-6353. https://doi.org/10.2147/IJN.S259188
45. Kanemura H. Therapeutic strategies in children with epilepsy: a quality-of-life-related perspective. J Clin Med. 2024;13(2):405. https://doi.org/10.3390/jcm13020405
46. Carter B, Cook G, Bray L, Collingwood A, Saron H, Rounce-field-Swales A. Parents' and caregivers' fears and concerns about their child's epilepsy: a scoping review. PLoS One. 2022;17(9):e0274001. https://doi.org/10.1371/journal.pone.0274001

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):133-141
DOI: 10.29058/mjwbs.1849779

Effect of intraoperative magnesium sulfate infusion on early postoperative opioid consumption in laparoscopic hysterectomy: A randomized, double-blind, controlled trial

Laparoskopik histerektomide intraoperatif magnezyum sülfat infüzyonunun erken postoperatif opioid tüketimine etkisi: Randomize, çift kör, kontrollü çalışma

İlke Tamdoğan[1], Dilek Yeniay[1], Fatma Alkan Bayburt[1], Tuna Albayrak[1], Azime Bulut[2]

[1]Department of Anesthesiology and Reanimation, Giresun University School of Medicine, Giresun, Türkiye
[2]Department of Anesthesiology and Reanimation, Medipol Bahçelievler University Hospital, İstanbul, Türkiye

Cite this article as: *Tamdoğan İ et al. Effect of intraoperative magnesium sulfate infusion on early postoperative opioid consumption in laparoscopic hysterectomy: A randomized, double-blind, controlled trial. Med J West Black Sea. 2026;10(1):133-141.*

ABSTRACT

Aim: To investigate whether intraoperative intravenous magnesium sulfate, added to a standard multimodal analgesia regimen, reduces cumulative opioid consumption during the first 24 hours after laparoscopic hysterectomy.

Material and Methods: In this single-center, prospective, randomized, double-blind, placebo-controlled trial, 40 adults scheduled for laparoscopic hysterectomy were randomized (1:1) to receive magnesium sulfate (20 mg/kg IV loading dose followed by a continuous infusion of 20 mg/kg/h until the end of surgery) or an equal volume of 0.9% saline. Postoperative analgesia was standardized using a predefined multimodal regimen and fentanyl based patient controlled analgesia; tramadol was administered as rescue analgesia when the Numeric Rating Scale (NRS) pain score was ≥ 4. The primary outcome was total opioid consumption within the first 24 postoperative hours, expressed as morphine milligram equivalents (MME). Secondary outcomes included NRS pain scores at rest and with movement at prespecified time points, postoperative nausea and vomiting (PONV), Ramsay sedation scores, rescue analgesic requirement, and intraoperative hemodynamic variables.

Results: Total fentanyl consumption (255 [182.5–392.5] μg vs 465.0 [262.5–500.0] μg; p=0.008) and total opioid consumption MME (34.5 [26.9–42.9] mg vs 50.0 [31.3–69.1] mg; p=0.011) were significantly lower in the magnesium group. NRS scores, PONV incidence, and Ramsay sedation scores did not differ at any time point. Mean arterial pressure was higher at minute 20 in the magnesium group (p=0.011), while hypotension and bradycardia rates were comparable.

Conclusion: Intraoperative magnesium sulfate reduced 24 h postoperative opioid requirements without improving pain scores, supporting magnesium as an opioid-sparing adjunct within multimodal analgesia.

Keywords: Laparoscopy, hysterectomy, magnesium sulfate, patient-controlled analgesia, postoperative pain.

ÖZ

Amaç: Standart multimodal analjezi rejimine eklenen intraoperatif intravenöz magnezyum sülfatın, laparoskopik histerektomi sonrası ilk 24 saatte kümülatif opioid tüketimini azaltıp azaltmadığını araştırmak.

Gereç ve Yöntemler: Bu tek merkezli, prospektif, randomize, çift kör, plasebo kontrollü çalışmada laparoskopik histerektomi planlanan 40 erişkin hasta 1:1 oranında randomize edilerek magnezyum sülfat (20 mg/kg IV yükleme dozu, ardından cerrahi sonuna kadar 20 mg/kg/saat sürekli infüzyon) veya eş hacimde %0.9 NaCl aldı. Postoperatif analjezi, önceden tanımlanmış multimodal rejim ve fentanil bazlı hasta kontrollü analjezi ile standartlaştırıldı. Sayısal Derecelendirme Skalası (NRS) ağrı skoru ≥ 4 olduğunda kurtarma analjezisi olarak tramadol uygulandı. Birincil sonlanım noktası, postoperatif ilk 24 saatteki toplam opioid tüketiminin morfin miligram eşdeğeri (MME) idi. İkincil sonlanım noktaları; belirlenmiş zaman noktalarında istirahatte ve hareketle NRS ağrı skorları, postoperatif bulantı-kusma (PONV), Ramsay sedasyon skorları (RSS), kurtarma analjezik gereksinimi ve intraoperatif hemodinamik değişkenlerdi.

Bulgular: Toplam fentanil tüketimi (255 [182.5–392.5] μg'ye karşı 465.0 [262.5–500.0] μg p=0.008) ve toplam opioid tüketimi MME (34.5 [26.9–42.9] mg'ye karşı 50.0 [31.3–69.1] mg p=0.011) magnezyum grubunda anlamlı olarak daha düşüktü. NRS skorları, PONV insidansı ve RSS skorları hiçbir zaman noktasında farklılık göstermedi. Ortalama arter basıncı magnezyum grubunda 20. dakikada daha yüksekti (p=0.011), hipotansiyon ve bradikardi oranları ise benzerdi.

Sonuç: İntraoperatif magnezyum sülfat, ağrı skorlarında iyileşme sağlamaksızın postoperatif 24 saatlik opioid gereksinimini azaltmıştır. Bu bulgu, multimodal analjezi içinde magnezyumu opioid azaltıcı bir adjuvan ajan olarak desteklemektedir.

Anahtar Kelimeler: Laparoskopi, histerektomi, magnezyum sülfat, hasta kontrollü analjezi, postoperatif ağrı.

Corresponding Author: İlke Tamdoğan ✉ drilkeipek@gmail.com
Received: 26.12.2025 **Revision:** 10.02.2026 **Accepted:** 17.02.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Intraoperative intravenous magnesium sulfate infusion significantly reduced cumulative opioid consumption during the first 24 hours after laparoscopic hysterectomy.
- The opioid-sparing effect of magnesium was achieved without changes in pain scores at rest or during movement.
- No between-group differences were observed in postoperative nausea and vomiting, sedation level, or rescue analgesic requirement.
- The administered magnesium sulfate regimen was well tolerated, with an acceptable safety profile and preserved intraoperative hemodynamic stability.

INTRODUCTION

Hysterectomy is one of the most commonly performed major gynecological surgical procedures worldwide. With advances in minimally invasive surgical techniques, laparoscopic hysterectomy (LH) has become increasingly preferred in contemporary practice because of its clear advantages over open surgery, including reduced blood loss, shorter length of hospital stay, and faster recovery (1). However, even minimally invasive approaches do not completely eliminate the risk of acute postoperative pain, which can impair patient comfort and delay early mobilization. This pain comprises components related to port-site incisions, deep visceral pain, and referred pain that is often perceived in the shoulder; therefore, effective analgesia is of critical importance for patient satisfaction and for preventing the development of chronic postoperative pain (2).

Current Enhanced Recovery After Surgery (ERAS) recommendations advocate a multimodal analgesia strategy in gynecologic surgery to minimize opioid consumption. This approach involves the combination of non-opioid systemic agents with regional and local analgesic techniques (3,4). Surgical trauma generates intense nociceptive input to the central nervous system, leading to central sensitization. Central sensitization is characterized by increased excitability of neurons in the spinal cord, resulting in allodynia (pain elicited by a normally non-painful stimulus) and hyperalgesia (an exaggerated pain response to a painful stimulus). One of the key molecular targets in this mechanism is the N-methyl-D-aspartate (NMDA) receptor, which is activated by glutamate, the principal excitatory neurotransmitter (5,6). Excessive activation of NMDA receptors in response to surgical stimulation contributes to a reduced pain threshold and the development of opioid tolerance (6).

Magnesium sulfate ($MgSO_4$) is an endogenous cation that physiologically blocks the ion channels of NMDA receptors. Through this mechanism, it exhibits antinociceptive and antihyperalgesic properties (6). This pharmacologic effect offers the potential to reduce postoperative pain intensity and produce an opioid-sparing effect by preventing postoperative central sensitization (5,6). Magnesium sulfate infusion has been shown to significantly reduce opioid consumption and prolong the time to the first analgesic requirement across various surgical procedures, particularly those associated with high pain intensity, such as spinal, cardiac, and major abdominal surgeries (7,8).

Despite being a minimally invasive technique, LH still requires effective analgesia because of visceral and somatic pain components. However, the results of studies investigating the contribution of magnesium sulfate to postoperative analgesia after LH remain inconsistent, and the optimal administration protocol has not been clearly established (9). This inconsistency may be attributable to differences in study design, variations in magnesium sulfate dosing and timing of administration, heterogeneity in concomitant multimodal analgesic regimens, and variability in patient populations (10,11). Consequently, definitive conclusions regarding the effectiveness and clinical benefits of magnesium sulfate in specific gynecological laparoscopic procedures, such as LH, remain unclear.

The primary aim of this study was to evaluate the effect of intraoperative intravenous (IV) MgSO4 infusion, in addition to a standard multimodal analgesia regimen, on opioid consumption during the first 24 h in patients undergoing LH. The secondary aims were to investigate postoperative pain severity, time to reach a Modified Aldrete Score (MAS) > 8 time to first analgesic requirement, incidence of postoperative nausea and vomiting.

MATERIAL and METHODS

Study Design

This single-center, prospective, randomized, double-blind clinical trial was conducted at the Department of Anesthesiology and Reanimation at Giresun Maternity and Child Health Training and Research Hospital. The study was designed in accordance with the principles of the Declaration of Helsinki and CONSORT recommendations for randomized controlled trials (Figure 1). The study protocol was approved by the Gümüşhane University Clinical Re-

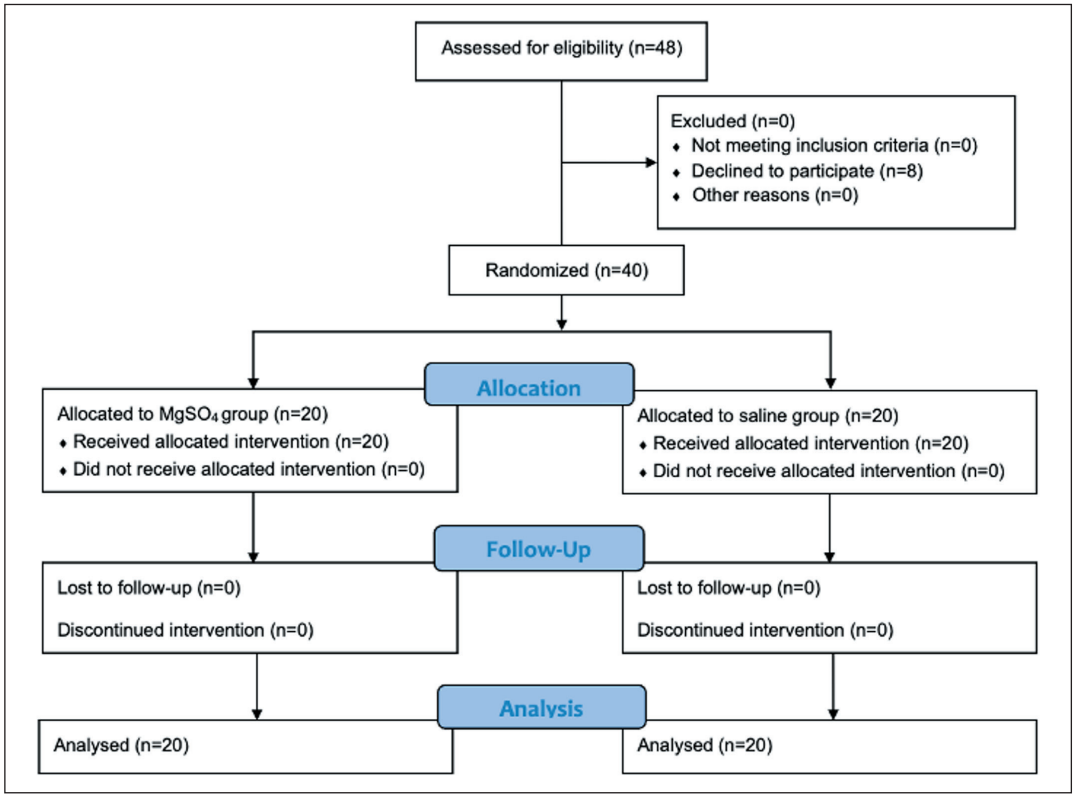


Figure 1: CONSORT flow diagram of patient enrollment, randomization, group allocation, follow-up and analysis.

search Ethics Committee (date: 23.02.2022; number: 2022-01) and registered at ClinicalTrials.gov (Registration No: NCT05826119) prior to the initiation of patient enrollment. Written informed consent was obtained from all participants.

Participants

The study was conducted at the Giresun Maternity and Child Health Training and Research Hospital between November 2023 and April 2024. Participants aged 20–70 years with an American Society of Anesthesiologists (ASA) physical status classification of I or II were included. The exclusion criteria were as follows: significant renal, cardiovascular, or hepatic dysfunction; neuromuscular disorders; use of calcium channel blockers; substance dependence; pregnancy or breastfeeding; morbid obesity (BMI >35 kg/m²); chronic opioid use or opioid dependence; and a history of hypersensitivity to the study drugs or $MgSO_4$.

Randomization

Prior to surgical intervention, patients were randomly assigned to one of two groups in a 1:1 ratio using opaque sealed envelopes. The randomization process utilized computer-generated codes created using SPSS v27.0 (IBM, New York, USA). Each patient selected an envelope, and the group assignment was determined by the information contained within. The study medications were prepared in the medication preparation room 30 min before surgery by a nurse not involved in the study. Throughout the study, all participants, including the patients, investigators, anesthesiologists responsible for anesthetic management, and surgeons, remained blinded to the group allocation. Study solutions were prepared in 100 mL syringes or infusion bags that were identical in volume and appearance and labeled solely with the patient's study number. The anesthesiologist managing the case, patient, surgeon, recovery room and ward nurses, and investigators collecting postoperative data were blinded to group assignment, ensuring a double-blind study design.

Anesthesia and Intraoperative Management

Midazolam (0.03 mg/kg) was administered intravenously for premedication. Standard monitoring included electrocardiography, non-invasive blood pressure, and peripheral oxygen saturation.

General anesthesia was induced with IV propofol 2–2.5 mg/kg and fentanyl 1–2 μg/kg, followed by rocuronium 0.6 mg/kg to facilitate tracheal intubation. Anesthesia was maintained with O_2/air (fraction of inspired oxygen 40%), age-adjusted sevoflurane at 0.8–1.2 minimum alveolar concentration, and an IV remifentanil infusion at 0.1–0.25 μg/kg/min. Heart rate and blood pressure were maintained within ± 20% of the preoperative values by titrating the remifentanil infusion rate.

During pneumoperitoneum, the intra-abdominal pressure was maintained at 12–14 mmHg. Hypotension was defined as a systolic blood pressure (SBP) <90 mmHg or a decrease >20% from the preoperative values. Bradycardia was defined as a heart rate (HR) <45 beats/min. Patients who developed these conditions were treated with IV ephedrine and atropine. At the end of the surgery, neuromuscular blockade was antagonized with sugammadex. After full recovery of neuromuscular function was confirmed based on clinical signs (head lift, handgrip strength, and SpO_2 >95%), the patients were extubated and transferred to the post-anesthesia care unit (PACU). Patients were transferred to the ward once the Modified Aldrete Score (MAS) exceeded 8, and the time taken to reach this threshold was recorded.

Heart rate (HR), mean arterial pressure (MAP) and SpO_2 were recorded at baseline (before $MgSO_4$ administration), 1 min after anesthesia induction, and subsequently at 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 90, 105, and 120 min; thereafter, measurements were obtained every 30 min. The surgical duration, extubation time, and total remifentanil consumption were recorded.

Study Interventions

Following anesthesia induction, patients in Group M received magnesium sulfate at a loading dose of 20 mg/kg, diluted in 0.9% NaCl and administered intravenously over 15 min, followed by a continuous infusion of 20 mg/kg/h until the end of surgery. Patients in Group K received an equal-volume 0.9% NaCl infusion (without magnesium) administered at the same rate and for the same duration. No other magnesium-containing medication was administered intraoperatively. Except for the study drug, anesthesia management and all other medications were standardized and administered identically in both the groups.

Pain Management and Postoperative Follow-Up

During the preanesthetic assessment, patients were instructed to use the Numeric Rating Scale (NRS), an 11-point scale ranging from 0 to 10 (0 = "no pain" and 10 = "the worst pain imaginable"). The patients were instructed to select the number that best represented their postoperative pain intensity. They were also trained in the use of the patient-controlled analgesia (PCA) device and informed that they could request opioids via the PCA if their resting NRS scores were ≥4.

All patients received IV ketoprofen 50 mg after anesthesia induction, IV paracetamol 1g approximately 30 min before the end of surgery, and tramadol 1 mg/kg. Postoperative analgesia consisted of IV paracetamol 1 g every 8 h, IV ketoprofen 50 mg every 12 h, and a standardized PCA device containing fentanyl 5 μg/mL (bolus dose 10 μg, lockout interval 5 min, no basal infusion, and an hourly limit of 50 μg). When the resting pain score was ≥4, rescue analgesia was provided with IV tramadol 0.5 mg/kg (maximum 400 mg/day).

Postoperative opioid consumption and total opioid use were calculated as morphine milligram equivalents (MME) using the GlobalRPH opioid-equivalency calculator. The following conversion factors were applied.

- 0.1 mg fentanyl = 10 mg morphine
- 100 mg tramadol = 10 mg morphine

Accordingly, IV tramadol administered for rescue analgesia and fentanyl consumed via PCA were converted to IV morphine milligram equivalents (MME) using the conversion factors.

NRS pain scores at rest and with movement were assessed at 30 min and 2, 4, 6, 12, and 24 h postoperatively. At the same time points, sedation was evaluated using the Ramsay Sedation Scale (1: awake, agitated, and/or crying; 2: awake, calm, and cooperative; 3: drowsy but responsive to verbal stimuli; 4: drowsy but responsive to glabellar tap; 5: drowsy, sluggish response to glabellar tap; 6: no response to stimuli), and postoperative nausea and vomiting (PONV) were assessed using the Verbal Descriptive Scale (VDS) (0: no nausea; 1: mild nausea; 2: moderate nausea; 3: vomiting; 4: multiple vomiting episodes). Administration of 10 mg IV metoclopramide was planned for patients with a nausea/vomiting score of 2 or for those who were unable to tolerate nausea. In addition, potential adverse events attributable to magnesium sulfate (respiratory depression, pruritus, arrhythmia, and hypotension requiring treatment) were recorded.

Outcome Measures

The primary outcome of this study was the cumulative opioid consumption during the first 24 postoperative hours. Cumulative opioid consumption was calculated by converting and summing the doses of fentanyl administered via PCA and tramadol administered as rescue analgesia into intravenous morphine MME.

Secondary outcomes included NRS pain scores at rest and with movement at 30 min and 2, 4, 6, 12, and 24 h postoperatively, nausea and vomiting scores, number of patients requiring rescue analgesia, time to first request for rescue analgesia, intraoperative remifentanil consumption, extubation time, intraoperative hemodynamic stability, incidence of hypotension and bradycardia and occurrence of potential magnesium sulfate–related adverse events (e.g., hypotension, bradycardia, prolonged neuromuscular blockade).

Sample Size

A priori power analysis was performed using an independent-samples t-test model. A reduction of at least 20% in 24-hour postoperative fentanyl consumption was pre-

defined as a clinically meaningful difference, consistent with previous regional anesthesia trials evaluating opioid-sparing effects (12). The statistical parameters required for the sample size calculation were derived from the study by Seyhan et al. (13), which reported postoperative opioid consumption following magnesium administration in a comparable anesthetic setting. Based on the reported mean and standard deviation values, a Cohen's d effect size of 1.318 was calculated. Using this effect size with 95% power and a two-sided alpha level of 0.05, the minimum required sample size was estimated as 16 patients per group. To compensate for potential dropouts or protocol exclusions, the target enrollment was increased to 20 patients per group, resulting in a total planned sample size of 40 participants.

Statistical Analysis

Descriptive statistics were presented as mean, standard deviation, median, minimum, maximum, frequency, and percentage, as appropriate. The distribution of variables was assessed using the Kolmogorov–Smirnov and Shapiro–Wilk tests. For comparisons of normally distributed independent continuous variables, the independent-samples t-test was used; for non-normally distributed independent continuous variables, the Mann–Whitney U test was applied. Categorical independent variables were analyzed using the chi-square test; when chi-square assumptions were not met, Fisher's exact test was used. All analyses were performed using SPSS version 27.0.

RESULTS

A total of 40 patients were included (Group K: n=20; Group M: n=20). The groups were comparable in terms of age, body mass index, surgical duration, and extubation time (all p>0.05). However, the proportion of ASA II patients was higher in the magnesium group than in the control group (95% vs 65%; p=0.018) (Table 1).

Intraoperatively, heart rate, mean arterial pressure, and SpO_2 values showed similar trends over time in both groups. Only at the 20th minute was the mean arterial pressure higher in the magnesium group (p=0.011); however, this difference was not associated with an increased incidence of clinically significant hypotension (Figure 2).

The incidence of hypotension and bradycardia did not differ between the groups (p=0.168 and p=0.311, respectively). Likewise, there were no significant between-group differences in total remifentanil consumption, remifentanil requirement per minute, or time to achieve a Modified Aldrete Score (MAS) > 8 (all p>0.05) (Table 2).

Postoperative NRS pain scores assessed at rest and during movement were similar between the groups at all evaluated time points (all p>0.05) (Table 3). The distributions of postoperative nausea and vomiting (PONV) scores and Ramsay Sedation Scale scores were also comparable between the groups at all time points (all comparisons p>0.05; Supplementary Tables S1 and S2).

Table 1: Baseline demographic and perioperative characteristics

Variable	Group K (n=20)	Group M (n=20)	p
Age (years)	47.5 (45.3–52.0)	48.5 (45.3–55.8)	0.714
BMI (kg/m²)	28.0 (24.5–31.5)	29.3 (27.6–31.4)	0.607
ASA physical status I/II, n (%)	7 (35.0) / 13 (65.0)	1 (5.0) / 19 (95.0)	**0.018**
Surgical duration (min)	130.0 (96.3–200.0)	110.0 (75.0–150.0)	0.178
Extubation time (min)	10.0 (5.0–10.0)	10.0 (6.3–10.0)	0.144

Data are presented as median (IQR) or n (%), t= independent samples t test, m= Mann–Whitney U test, χ^2= chi-square test, **BMI:** body mass index, **ASA:** American Society of Anesthesiologists, **IQR:** interquartile range, min:minute.

Table 2: Intraoperative hemodynamic and anesthetic outcomes

Variable	Group K (n=20)	Group M (n=20)	p
Hypotension, yes/no, n (%)	16 (80) / 4 (20)	12 (60) / 8 (40)	0.168
Bradycardia, yes/no, n (%)	12 (60) / 8 (40)	15 (75) / 5 (25)	0.311
Total remifentanil, μg	710.5 ± 356.3	636.5 ± 236.5	0.444
Remifentanil, μg/min	4.8 ± 1.7	5.4 ± 1.8	0.281
Time to MAS >8, min	8.9 ± 4.1	10.3 ± 4.1	0.286

Data are presented as mean ± SD or n (%), p values were calculated using the independent-samples t test (continuous variables) and the χ^2 test (categorical variables), **MAS:** Modified Aldrete Score, **SD:** standard deviation, **μg:** microgram, **min:** minute

Regarding rescue analgesia, the time to the first tramadol request and total tramadol consumption were similar between the groups (p=0.853 and p=0.640, respectively). In contrast, total fentanyl consumption was lower in the magnesium group (255 [182.5–392.5] μg vs 465.0 [262.5–500.0] μg; p=0.008). Similarly, total opioid consumption (MME) was significantly lower in the magnesium group (34.5 [26.9–42.9] mg vs 50.0 [31.3–69.1] mg; p=0.011) (Table 4).

DISCUSSION

In this study, we evaluated the effects of intraoperative magnesium sulfate infusion, added to a standard anesthetic technique and multimodal analgesia regimen, on early postoperative analgesia and opioid requirements in patients undergoing laparoscopic hysterectomy. The main finding of our study was that magnesium sulfate significantly reduced cumulative opioid consumption within the first 24 h postoperatively. In contrast, NRS pain scores at rest and with movement remained comparable between the two groups at all assessment timepoints. Likewise, no significant between-group differences were observed in the distribution of PONV or Ramsay Sedation Scale scores.

The potential contribution of magnesium to analgesia has been attributed to its anti-hyperalgesic effects through modulation of NMDA-mediated central sensitization (14,15). In the clinical literature, meta-analyses have indicated that perioperative intravenous magnesium can reduce opioid consumption, particularly during the first 24 h, whereas its effect on pain scores tends to be more variable and generally more modest (7,16). In gynecologic surgery, outcomes have been reported more inconsistently, likely due to heterogeneity in magnesium protocols (loading dose/infusion rate/duration) and in accompanying multimodal analgesia

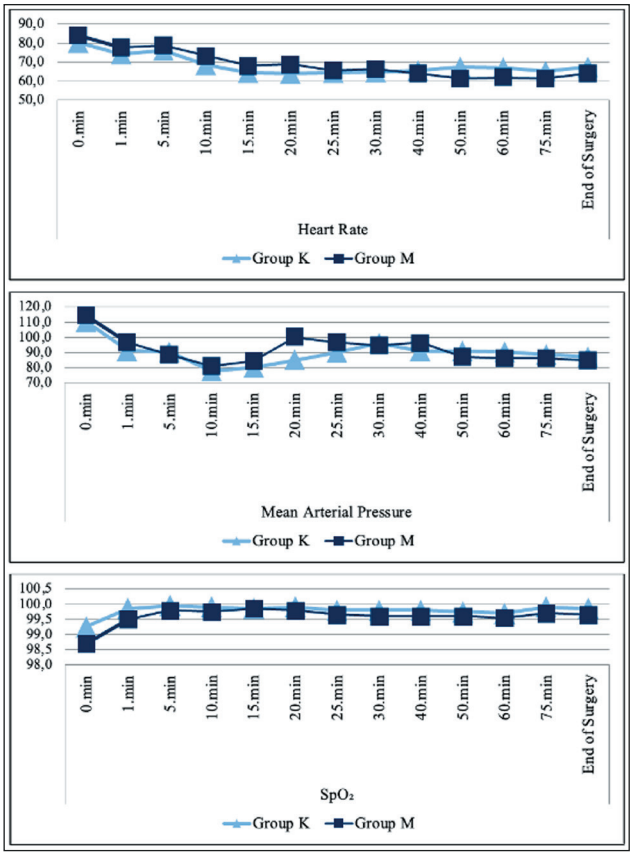


Figure 2: HR, MAP and SpO_2 according to group and time.

Table 3: Postoperative pain scores (NRS) at rest and with movement

Time point	Group K (n=20)	Group M (n=20)	p
NRS at rest			
30 min	4.0 (3.3–5.0)	5.0 (4.0–6.0)	0.152
2 h	4.0 (2.3–4.8)	4.0 (2.3–5.0)	0.562
4 h	2.0 (2.0–3.8)	3.0 (2.0–4.0)	0.161
6 h	2.0 (2.0–3.0)	2.5 (2.0–4.0)	0.214
12 h	2.0 (1.0–3.0)	2.0 (1.3–3.0)	0.503
24 h	2.0 (1.0–2.8)	2.0 (1.0–2.0)	0.697
NRS with movement			
30 min	6.0 (4.3–6.0)	6.0 (5.0–7.0)	0.216
2 h	5.0 (4.0–6.0)	5.0 (4.0–6.0)	0.628
4 h	4.0 (3.0–5.0)	4.0 (3.3–4.8)	0.577
6 h	3.0 (3.0–4.0)	3.0 (3.0–4.0)	0.725
12 h	3.0 (2.0–4.0)	3.0 (2.0–4.0)	0.362
24 h	3.0 (2.0–4.0)	3.0 (2.0–3.8)	0.311

Data are presented as median (IQR), p values were calculated using the Mann–Whitney U test, **NRS:** numeric rating scale, **min:** minute, **h:** hour.

Table 4: Opioid consumption and rescue analgesia

Outcome	Group K (n=20)	Group M (n=20)	p
Time to first rescue tramadol, min	20.0 (11.3–20.0)	20.0 (0.0–27.5)	0.853
Total tramadol consumption (mg)	75.0 (60–150)	72.5 (0.0–150.0)	0.640
Total fentanyl consumption (μg)	465.0 (262.5–500.0)	255.0(182.5–392.5)	**0.008**
Total opioid consumption, MME (mg)	50.0 (31.3–69.1)	34.5 (26.9–42.9)	**0.011**

Data are presented as median (IQR), p values were obtained using the Mann–Whitney U test, MME conversion: fentanyl 0.1 mg = morphine 10 mg; tramadol 100 mg = morphine 10 mg (as defined in the protocol), **IQR:** interquartile range, **MME:** morphine milligram equivalents.

regimens (17,18). In this respect, our focus on laparoscopic hysterectomy is noteworthy, as postoperative pain after laparoscopy is multifactorial (somatic trocar-site pain, visceral pain, and pneumoperitoneum-related shoulder pain). Under a robust multimodal analgesic background, detecting differences in one-dimensional measures, such as the NRS, may be challenging. Accordingly, in some cases the benefit of magnesium may manifest not as lower pain scores but rather as "less opioid required to achieve the same level of analgesia."

In our study, the coexistence of reduced opioid consumption with unchanged NRS scores is consistent with two complementary explanations: (i) magnesium exerted a true opioid-sparing effect by reducing opioid requirements without compromising analgesic adequacy, and (ii) the standardized multimodal analgesia protocol applied in accordance with the PROSPECT guideline may have lowered pain scores to a relatively similar baseline in both groups, making additional reductions in NRS from an adjuvant difficult to demonstrate (9). This pattern aligns broadly with previous evidence suggesting that, particularly when routine non-opioid analgesics are incorporated, the incremental impact of adjuvants may be more clearly reflected in cumulative opioid consumption than in pain intensity ratings (7,19).

Key safety concerns associated with magnesium in clinical practice include hypotension, bradycardia, and prolonged neuromuscular blockades. In our study, the absence of between-group differences in hypotension/bradycardia incidence, total remifentanil consumption, and time to extubation/achievement of MAS > 8 suggests that the administered dose and infusion scheme was hemodynamically tolerable. This observation is also consistent with the overall safety profile reported in broader syntheses, indicating that perioperative magnesium can be administered without major adverse events (20,21). The similar PONV distribution and lack of differences in Ramsay sedation scores further support that the reduction in opioid consumption was not achieved at the expense of increased sedation, thereby strengthening the potential clinical value of this approach.

This study had several limitations. First, serum magnesium levels were not monitored, which limits the interpretation of dose–response relationships and safety margins. Second, quantitative neuromuscular monitoring (e.g., train-of-four [TOF]) was not routinely used; although sugammadex reversal was administered and recovery was assessed clinically, the absence of objective monitoring may have reduced the precision of confirming full neuromuscular recovery and may allow residual neuromuscular blockade to go undetected. Third, the study was a single-center study with a relatively small sample size; therefore, inadequate power for secondary outcomes, such as pain scores and PONV, cannot be excluded. Finally, long-term and patient-reported outcomes, including chronic postoperative pain, functional recovery, time to discharge, and patient satisfaction, were not assessed. Future large-scale multicenter studies with standardized protocols are needed to clarify these aspects.

Conclusion

intraoperative magnesium sulfate infusion during laparoscopic hysterectomy in the context of multimodal analgesia reduced cumulative opioid consumption within the first 24 postoperative hours without producing a meaningful change in pain scores. These findings suggest that magnesium may be considered a low-cost and feasible adjuvant within opioid-sparing strategies; however, larger studies are required to define the optimal dosing/administration regimen and identify patient subgroups most likely to benefit.

Author Contributions

Study conception and design: **Ilke Tamdogan, Azime Bulut;** data collection: **Ilke Tamdogan, Dilek Yeniay;** analysis and interpretation of results: **Ilke Tamdogan, Tuna Albayrak;** draft manuscript preparation: **Fatma Alkan Bayburt.** The author(s) reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Financial Support

The authors declare that the study received no funding.

Ethical Approval

The study was approved by the Gümüşhane University Ethics Committee (date: 23.02.2022, number: 2022-01).

REFERENCES

1. ACOG Practice Bulletin No. 195. Clinical management guidelines for obstetrician-gynecologists: choosing the route of hysterectomy for benign disease. Obstet Gynecol. 2018;131(6).
2. Chou R, Gordon DB, de Leon-Casasola OA, Rosenberg JM, Bickler S, Brennan T, et al. Management of postoperative pain: a clinical practice guideline from the American Pain Society, the American Society of Regional Anesthesia and Pain Medicine, and the American Society of Anesthesiologists' Committee on Regional Anesthesia, Executive Committee, and Administrative Committee. J Pain. 2016;17(2):131-157.
3. Geng Z, Bi H, Zhang D, Xiao C, Song H, Feng Y, et al. The impact of multimodal analgesia based enhanced recovery protocol on quality of recovery after laparoscopic gynecological surgery: a randomized controlled trial. BMC Anesthesiol. 2021;21(1):179. https://doi.org/10.1186/s12871-021-01399-2
4. Raghvendra KP, Thapa D, Mitra S, Ahuja V, Gombar S, Huria A. Postoperative pain relief following hysterectomy: a randomized controlled trial. J Midlife Health. 2016;7(2):65-68. https://doi.org/10.4103/0976-7800.185327

5. Kido K, Katagiri N, Kawana H, Sugino S, Konno D, Suzuki J, et al. Effects of magnesium sulfate administration in attenuating chronic postsurgical pain in rats. Biochem Biophys Res Commun. 2021;534:395-400. https://doi.org/10.1016/j.bbrc.2020.11.069
6. Tsaousi G, Nikopoulou A, Pezikoglou I, Birba V, Grosomanidis V. Implementation of magnesium sulphate as an adjunct to multimodal analgesic approach for perioperative pain control in lumbar laminectomy surgery: a randomized placebo-controlled clinical trial. Clin Neurol Neurosurg. 2020;197:106091. https://doi.org/10.1016/j.clineuro.2020.106091
7. Albrecht E, Kirkham KR, Liu SS, Brull R. Peri-operative intravenous administration of magnesium sulphate and postoperative pain: a meta-analysis. Anaesthesia. 2013;68(1):79-90. https://doi.org/10.1111/j.1365-2044.2012.07335.x
8. Ferasatkish R, Dabbagh A, Alavi M, Mollasadeghi G, Hydarpur E, Moghadam AA, et al. Effect of magnesium sulfate on extubation time and acute pain in coronary artery bypass surgery. Acta Anaesthesiol Scand. 2008;52(10):1348-1352. https://doi.org/10.1111/j.1399-6576.2008.01783.x
9. Lirk P, Thiry J, Bonnet MP, Joshi GP, Bonnet F. Pain management after laparoscopic hysterectomy: systematic review of literature and PROSPECT recommendations. Reg Anesth Pain Med. 2019;44(4):425-436. https://doi.org/10.1136/rapm-2018-100024
10. Benevides ML, Fialho DC, Linck D, Oliveira AL, Ramalho DHV, Benevides MM. Intravenous magnesium sulfate for postoperative analgesia after abdominal hysterectomy under spinal anesthesia: a randomized, double-blind trial. Braz J Anesthesiol. 2021;71(5):498-504. https://doi.org/10.1016/j.bjane.2021.01.008
11. Tan W, Qian DC, Zheng MM, Lu X, Han Y, Qi DY. Effects of different doses of magnesium sulfate on pneumoperitoneum-related hemodynamic changes in patients undergoing gastrointestinal laparoscopy: a randomized, double-blind, controlled trial. BMC Anesthesiol. 2019;19(1):237. https://doi.org/10.1186/s12871-019-0886-4
12. Altıparmak B, Korkmaz Toker M, Uysal AI, Turan M, Gümüş Demirbilek S. Comparison of the effects of modified pectoral nerve block and erector spinae plane block on postoperative opioid consumption and pain scores of patients after radical mastectomy surgery: a prospective, randomized, controlled trial. J Clin Anesth. 2019;54:61-65. https://doi.org/10.1016/j.jclinane.2018.10.040
13. Seyhan T, Tugrul M, Sungur M, Kayacan S, Telci L, Pembeci K, et al. Effects of three different dose regimens of magnesium on propofol requirements, hemodynamic variables and postoperative pain relief in gynaecological surgery. Br J Anaesth. 2006;96(2):247-252. https://doi.org/10.1093/bja/aei291
14. Lee C, Song YK, Jeong HM, Park SN. The effects of magnesium sulfate infiltration on perioperative opioid consumption and opioid-induced hyperalgesia in patients undergoing robot-assisted laparoscopic prostatectomy with remifentanil-based anesthesia. Korean J Anesthesiol. 2011;61(3):244-250. https://doi.org/10.4097/kjae.2011.61.3.244
15. Guo BL, Lin Y, Hu W, Zhen CX, Zhang BC, et al. Effects of systemic magnesium on post-operative analgesia: is the current evidence strong enough? Pain Physician. 2015;18(5):405-418.
16. Avci Y, Rajarathinam M, Kalsekar N, Tawfic Q, Krause S, Nguyen D, et al. Unravelling the analgesic effects of perioperative magnesium in general abdominal surgery: a systematic review and meta-analysis of randomized controlled trials. Braz J Anesthesiol. 2024;74(4):844524. https://doi.org/10.1016/j.bjane.2024.844524
17. McKeown A, Seppi V, Hodgson R. Intravenous magnesium sulphate for analgesia after caesarean section: a systematic review. Anesthesiol Res Pract. 2017;2017:9186374. https://doi.org/10.1155/2017/9186374
18. Ryu JH, Kang MH, Park KS, Do SH. Effects of magnesium sulphate on intraoperative anaesthetic requirements and postoperative analgesia in gynaecology patients receiving total intravenous anaesthesia. Br J Anaesth. 2008;100(3):397-403. https://doi.org/10.1093/bja/aem407
19. Ng KT, Yap JLL, Izham IN, Teoh WY, Kwok PE, Koh WJ. The effect of intravenous magnesium on postoperative morphine consumption in noncardiac surgery: a systematic review and meta-analysis with trial sequential analysis. Eur J Anaesthesiol. 2020;37(3):212-223. https://doi.org/10.1097/EJA.0000000000001164
20. Fei S, Xia H, Chen X, Pang D, Xu X. Magnesium sulfate reduces the rocuronium dose needed for satisfactory double lumen tube placement conditions in patients with myasthenia gravis. BMC Anesthesiol. 2019;19(1):170. https://doi.org/10.1186/s12871-019-0841-4
21. Yazdi AP, Esmaeeli M, Gilani MT. Effect of intravenous magnesium on postoperative pain control for major abdominal surgery: a randomized double-blinded study. Anesth Pain Med. 2022;17(3):280-285. https://doi.org/10.17085/apm.22156

Supplementary Table S1: Postoperative nausea and vomiting scores

Time point	PONV category	Group K (n=20), n (%)	Group M (n=20), n (%)	p
30 min	None	10 (50.0)	11 (55.0)	0.752
	Mild	4 (20.0)	3 (15.0)	
	Moderate	4 (20.0)	5 (25.0)	
	Vomiting	2 (10.0)	1 (5.0)	
	Multiple vomiting episodes	0 (0.0)	0 (0.0)	
2 h	None	14 (70.0)	15 (75.0)	0.723
	Mild	5 (25.0)	0 (0.0)	
	Moderate	1 (5.0)	2 (10.0)	
	Vomiting	0 (0.0)	3 (15.0)	
	Multiple vomiting episodes	0 (0.0)	0 (0.0)	
4 h	None	15 (75.0)	17 (85.0)	0.429
	Mild	4 (20.0)	2 (10.0)	
	Moderate	1 (5.0)	0 (0.0)	
	Vomiting	0 (0.0)	1 (5.0)	
	Multiple vomiting episodes	0 (0.0)	0 (0.0)	
6 h	None	16 (80.0)	14 (70.0)	0.465
	Mild	2 (10.0)	1 (5.0)	
	Moderate	1 (5.0)	3 (15.0)	
	Vomiting	1 (5.0)	2 (10.0)	
	Multiple vomiting episodes	0 (0.0)	0 (0.0)	
12 h	None	19 (95.0)	18 (90.0)	1.000
	Mild	1 (5.0)	1 (5.0)	
	Moderate	0 (0.0)	0 (0.0)	
	Vomiting	0 (0.0)	1 (5.0)	
	Multiple vomiting episodes	0 (0.0)	0 (0.0)	
24 h	None	17 (85.0)	18 (90.0)	0.633
	Mild	2 (10.0)	1 (5.0)	
	Moderate	1 (5.0)	0 (0.0)	
	Vomiting	0 (0.0)	1 (5.0)	
	Multiple vomiting episodes	0 (0.0)	0 (0.0)	

Data are presented as n (%), PONV categories were defined as: 0:none, 1:mild nausea, 2: moderate nausea and 3: vomiting, 4:multiple emesis episodes.

Supplementary Table S2: Postoperative Ramsay Sedation Scale

Time point	Group K (n=20)	Group M (n=20)	p
30 min	2.1 ± 0.7 (2.0)	2.5 ± 0.6 (2.5)	0.062
2 h	2.0 ± 0.5 (2.0)	2.1 ± 0.6 (2.0)	0.756
4 h	2.0 ± 0.4 (2.0)	1.8 ± 0.4 (2.0)	0.251
6 h	1.9 ± 0.3 (2.0)	1.8 ± 0.4 (2.0)	0.382
12 h	1.9 ± 0.4 (2.0)	1.7 ± 0.5 (2.0)	0.149
24 h	2.0 ± 0.0 (2.0)	1.9 ± 0.3 (2.0)	0.152

Values are mean ± SD (median), Between-group comparisons were performed using the Mann–Whitney U test, RSS: Ramsay Sedation Scale, SD: standard deviation, min: minutes, h: hours.

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):142-150
DOI: 10.29058/mjwbs.1835988

Effects of preemptive analgesia and preoperative anxiety on emergence delirium after pediatric adenotonsillectomy

Pediatrik adenotonsillektomi sonrası uyanma deliryumu üzerine preemptif analjezi ve preoperatif anksiyetenin etkileri

Gamze Küçükosman[1], Bengü G. Köksal[2], Keziban Bollucuoğlu[2], Rahşan D. Okyay[2], Dilay İ. Alp[2], Özcan Pişkin[2], Hilal Ayoğlu[2]

[1]Department of Anesthesiology and Reanimation, University of Health Sciences Trabzon Kanuni Training and Research Hospital, Trabzon, Türkiye

[2]Department of Anesthesiology and Reanimation, Zonguldak Bülent Ecevit University, Faculty of Medicine, Zonguldak, Türkiye

Cite this article as: *Küçükosman G et al. Effects of preemptive analgesia and preoperative anxiety on emergence delirium after pediatric adenotonsillectomy. Med J West Black Sea. 2026;10(1):142-150.*

ABSTRACT

Aim: Postoperative emergence delirium (PED) is characterized by restlessness, disorientation, and agitation, particularly in children during emergence from anesthesia. Its etiology includes preoperative anxiety, head-neck surgery, inhalation agents, and pain. This study evaluated the effects of preemptive analgesia and preoperative anxiety on PED after pediatric adenotonsillectomy.

Material and Methods: After ethics committee approval and written parental consent were obtained, 64 pediatric patients were enrolled. In this prospective observational study, patients were categorized according to the timing of intravenous acetaminophen administration: 30 minutes before surgical incision(Group 1) or 15 minutes before the end of surgery (Group 2). Preoperative anxiety was assessed using the modified Yale Preoperative Anxiety Scale (m-YPAS), postoperative delirium using the Pediatric Anesthesia Emergence Delirium (PAED) scale, and pain using the Wong-Baker Faces Pain Scale (WBFPS). Measurements were recorded at 0, 5, 10, 15, 30, 45, and 60 minutes postoperatively. A p value <0.05 was considered statistically significant.

Results: Groups were comparable in demographics, anesthesia time, surgical history, parental satisfaction, and complications. Preoperative anxiety was high in both groups (m-YPAS ≥30). Group 1 had shorter extubation times(p<0.001). PAED scores ≥10 persisted for 10 min in Group 1 and 15 min in Group 2. PAED and WBFPS scores decreased over time, with a positive correlation between them in both groups (p<0.05).

Conclusion: Preemptive analgesia reduced postoperative pain but did not significantly affect PED within the first 10 minutes. These findings suggest that managing both anxiety and pain may be necessary to prevent PED in pediatric patients. Further studies are warranted.

Keywords: Preemptive analgesia, preoperative anxiety, emergence delirium, pediatric adenotonsillectomy

ÖZ

Amaç: Postoperatif uyanma deliryumu (PED), özellikle çocuklarda anesteziden uyanma sırasında huzursuzluk, yönelim bozukluğu ve ajitasyon ile karakterizedir. Etiyolojisinde preoperatif anksiyete, baş-boyun cerrahisi, inhalasyon ajanları ve ağrı rol oynar. Bu çalışmada, pediatrik adenotonsillektomi sonrası PED üzerine preemptif analjezi ve preoperatif anksiyetenin etkileri araştırılmıştır.

Gereç ve Yöntemler: Etik kurul onayı ve yazılı ebeveyn onamı alındıktan sonra 64 pediatrik hasta çalışmaya dahil edildi. Bu prospektif gözlemsel çalışmada hastalar, intravenöz asetaminofen uygulama zamanına göre; cerrahi insizyondan 30 dakika önce (Grup 1) veya cerrahinin bitiminden 15 dakika önce (Grup 2) olacak şekilde sınıflandırıldı. Preoperatif anksiyete modified Yale Preoperative Anxiety Scale (m-YPAS) ile, postoperatif deliryum Pediatric Anesthesia Emergence Delirium (PAED) ölçeği ile ve ağrı Wong-Baker Faces Pain Scale (WBFPS) ile değerlendirildi. Ölçümler postoperatif 0, 5, 10, 15, 30, 45 ve 60. dakikalarda kaydedildi. p<0,05 değeri anlamlı kabul edildi.

Bulgular: Gruplar demografik özellik, anestezi süresi, cerrahi geçmiş, ebeveyn memnuniyeti ve komplikasyonlar açısından benzer bulunmuştur. Her iki grupta preoperatif anksiyete yüksekti (m-YPAS ≥30). Grup 1'in ekstübasyon süresi anlamlı olarak daha kısaydı (p<0.001). PAED skorları Grup 1'de 10 dakika, Grup 2'de 15 dakika ≥10 olarak kaldı. PAED ve WBFPS skorları zamanla azalmış ve her iki grupta pozitif korelasyon göstermiştir (p<0.05).

Sonuç: Preemptif analjezi postoperatif ağrıyı azaltmış ancak ilk 10 dakikadaki PED skorlarını anlamlı şekilde etkilememiştir. Bulgular, PED'nin önlenmesinde hem anksiyete hem de ağrı yönetiminin birlikte ele alınmasının önemini göstermektedir.

Anahtar Kelimeler: Preemptif analjezi, Preoperatif anksiyete, uyanma deliryumu, pediatrik adenotonsillektomi

Corresponding Author: Gamze Küçükosman ✉ gamzebeu@gmail.com **Received:** 05.12.2025 **Revision:** 10.02.2026 **Accepted:** 06.03.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- The preemptive group (acetaminophen 30 minutes before surgery) demonstrated significantly shorter extubation times ($p < 0.001$).
- In Group 1, Pediatric Anesthesia Emergence Delirium scores remained ≥10 for the first 10 minutes postoperatively, whereas in Group 2, elevated PAED scores persisted through 15 minutes.
- PAED and WBFPS scores were initially positively correlated and declined over time. Although preemptive analgesia reduced postoperative pain, it was insufficient alone to meaningfully decrease PED in the early awakening period. These findings support the need for multimodal strategies, combining pain control with preoperative anxiety management, to more effectively prevent PED.

INTRODUCTION

Adenotonsillectomy is one of the most frequently performed otorhinolaryngology (i.e., ear-nose-throat, ENT) surgeries in children. During the period of emergence from anesthesia following pediatric adenotonsillectomy, unwanted behaviors and conditions such as pain, postoperative emergence delirium (PED), or both may occur (1-6). It has been shown that preoperative anxiety and pain play a role in the development of PED (5-10). It has been reported that preemptive analgesia reduces the severity of PED by lowering postoperative pain. However, results obtained on this topic, the timing of preemptive analgesia, and the effects of the method are still debated (1-12). Recent studies have demonstrated that preemptive acetaminophen administration alleviates postoperative pain in children undergoing ENT procedures (1-5,8,10-12).

This study was conducted to determine the effects of preemptive analgesia and preoperative anxiety on PED following pediatric adenotonsillectomy.

MATERIAL and METHODS

Ethics Committee Approval

This prospective observational study was conducted at Zonguldak Bülent Ecevit University Hospital between October 2019 and 2020 after obtaining local ethics committee approval (protocol No: 2019/14, ClinicalTrials.gov Identifier: NCT05802082) and written parental consent. All procedures in the study complied with the principles of the Declaration of Helsinki.

Study Design and Participants

The study sample included children aged 2-10 years who were scheduled for elective adenotonsillectomy and categorized in the American Society of Anesthesiologists (ASA) I-II risk groups. Patients who had allergies to the study drugs, a risk of difficult airway (e.g., obesity), neurological or developmental disorders (e.g., epilepsy), a recent or active upper respiratory tract infection, chronic medication use (e.g., anxiolytics, opioids, steroids), adverse events observed during emergence (hypoxia, hypercarbia, hypotension, hypoglycemia, or hypo-/hyperthermia), or lack of parental consent were excluded from the study.

Children and their parents were informed about the measurement instruments to be used in the study in the anesthesia outpatient clinics and preoperative waiting rooms. To make vascular access easier in the children who were not premedicated, when they were in the ENT inpatient clinic, a lidocaine-prilocaine cream (EMLA®) was applied to the back of the appropriate hand, the hand was covered, and the children were transferred to the operating room. The children were taken to the preoperative waiting room with one of their parents 30 min before surgery, and their age, sex, weight, ASA classification, and duration of fasting were recorded. The preoperative anxiety levels of the children were evaluated using m-YPAS (13) in the preoperative waiting room with their parents present.

Intervention and Grouping

After m-YPAS assessment, intravenous access was established on the dorsum of the hand where EMLA® cream had been applied, using a 22-gauge cannula, and fluid infusion was initiated.

In this prospective observational study, patients were categorized according to the timing of intravenous acetaminophen (10 mg/kg) administration: 30 minutes before surgical incision (Group 1) and 15 minutes before the end of surgery (Group 2). No random allocation was performed; the timing of acetaminophen administration was determined according to routine clinical practice. Patients in Group 1 were transferred to the operating room following acetaminophen infusion, whereas patients in Group 2 were transferred after completion of the m-YPAS assessment. Parents accompanied their children during this process.

Anesthesia Management

After admission to the operating room, routine monitoring was established in all patients, including electrocardiography (ECG), non-invasive blood pressure (NIBP), and peripheral oxygen saturation (SpO_2). To maintain intraopera-

tive normothermia, all patients were actively warmed using a warming blanket throughout the procedure. Anesthesia was administered by an anesthesiologist who was unaware of the study protocol. The administration of midazolam during anesthesia induction and the operating room temperature were recorded.

Anesthesia induction was performed using a standardized protocol in all patients. Intravenous induction was achieved with propofol (2–3 mg/kg) and fentanyl (1–2 μg/kg). Neuromuscular blockade for endotracheal intubation was established with rocuronium (0.6 mg/kg). Anesthesia was maintained with inhaled sevoflurane (2%) in a 50% oxygen–air mixture and remifentanil infusion.

Endotracheal intubation was performed by an experienced pediatric anesthesiologist using direct laryngoscopy. Endotracheal tube size was determined according to the age-based formula (age/4 + 4), and cuffed spiral endotracheal tubes were used in all patients. No cases of difficult intubation were encountered, and no repeated intubation attempts were required.

Following completion of the adenotonsillectomy and confirmation of adequate hemostasis by the surgeon, anesthetic gases and remifentanil infusion were discontinued. The time from induction to discontinuation of anesthetic agents was recorded as the anesthesia duration. The time from discontinuation of anesthetic agents to extubation was defined as the extubation time. The time from discontinuation of anesthetic agents to achieving an Aldrete score of 9 was defined as the recovery time.

At the end of surgery, neuromuscular blockade was reversed with sugammadex. Patients were extubated after full recovery of spontaneous respiration, protective airway reflexes, and adequate consciousness.

All intraoperative and postoperative medications were recorded. For postoperative nausea and vomiting management, metoclopramide (0.15 mg/kg) was administered to all patients; if nausea or vomiting persisted, ondansetron (0.1 mg/kg) was administered. Parents were allowed to accompany their children in the postoperative recovery unit.

Data Collection Instruments

Preoperative anxiety was assessed using the m-YPAS, PED was assessed using the Pediatric Anesthesia Emergence Delirium (PAED) score, and pain was assessed using the Wong-Baker Faces Pain Scale (WBFPS).

Modified Yale Preoperative Anxiety Scale

The Yale Preoperative Anxiety Scale (YPAS), developed to assess the anxiety levels of children before they undergo surgical interventions, was revised in 1997 and updated as m-YPAS (13). m-YPAS evaluates anxiety under five domains based on 22 forms of behavior. The raw score range of m-YPAS is 5-22, while evaluations are made by converting this range into 0-100 using the formula described below. The score for the form of behavior observed to the highest extent in each of the 5 m-YPAS domains is taken as the score in that domain. Because each domain of m-YPAS has different scores (4 or 6), and the child receives the score corresponding to their behavior in each domain, the scoring is made using the correction formula "(1/4 + 1/6 + 1/4 + 1/4 + 1/4) x 100/5", and a score of 23-100 is obtained. Children with scores equal to or smaller than 29 are considered "calm", while those with scores equal to or greater than 35 are considered "anxious"(13). The Turkish version of the modified Yale Preoperative Anxiety Scale demonstrated good internal consistency (Cronbach's alpha = 0.85 and 0.86 for two independent observers) and was reported to be a valid and reliable instrument for assessing preoperative anxiety in children (14).

Wong-Baker Faces Pain Scale

In the assessment of postoperative pain, scales with facial expression options are preferred in children, elderly people, and patients with communication problems (8). WBFPS, which was developed in 1988 by Wong et al., is a pain assessment scale that can be applied fast to and understood easily by children under 18 years of age. Each number on the scale in the range of 0-10 corresponds to a different facial expression, and these expressions indicate the level of pain felt by the patient. While the number 0 corresponds to a happy face indicating no pain at all, the number 10 corresponds to a crying face indicating the most severe pain. In assessments using WBFPS, the physician explains to the patient how the scale is used and asks them to express their pain by instructing them to 'mark the level of pain they are currently feeling'(8,15). In our study, we planned to administer oral ibuprofen at a dose of 10 mg/kg to the patients with WBFPS scores of ≥4.

Pediatric Anesthesia Emergence Delirium

PAED was defined in 2004 by Sikich et al. (16) to diagnose postoperative delirium in children over the age of 2. It has 5 psychometric items. Three of these items have significant power to distinguish delirium from pain, while the last 2 items have lower levels of sensitivity. PAED scores of ≥10 are interpreted as the 'presence of emergence delirium'. PAED has been acknowledged as an ideal, reliable, and valid scale to evaluate agitation in pediatric patients, demonstrating good internal consistency (Cronbach's alpha = 0.89) and interobserver reliability (0.84) in the original study by Sikich and Lerman (16). Although PAED has been used in several studies in Türkiye (17,18), no Turkish version with reported validity and reliability has been published to date. In our study, PAED scores of ≥10 were considered to be indicative of PED. In case of delirium, 1 mg/kg propofol was planned as a rescue treatment.

Parental satisfaction

In the postoperative period, the parents of the patients were asked to rate their levels of satisfaction with the process (1=bad, 2=poor, 3=good, or 4=excellent).

Each data collection instrument (m-YPAS, WBFPS, and PAED) was filled out by the same researcher who was not informed about group allocations, did not participate in the intraoperative procedures, but was trained about the scales used in the study. The measurements were made at the postoperative 0th, 5th, 10th, 15th, 30th, 45th, and 60th mins.

Each patient, their parent, the anesthesiologist, and the surgeon were blinded to the group allocations in the study. The time from admission to the postoperative reanimation room to transfer to the inpatient clinic (reanimation time) and complications, if any (nausea, vomiting, cough, head-neck pain, and increased secretion), were recorded.

Statistical Analysis

The sample size needed to conduct the study was calculated based on a 95% confidence interval (1-α), 95% testing power (1-β), and an effect size of d=0.838, and it was determined that at least 32 patients needed to be included in each group (19). The collected data were analyzed using the IBM Statistical Package for the Social Sciences (SPSS) V23. The Shapiro-Wilk test was used to assess normality. Categorical data were compared between groups using Pearson's chi-squared or Yates-corrected chi-squared tests. Two independent groups of normally distributed data were compared using independent-samples t-tests, while non-normally distributed data were compared using the Mann-Whitney U test. Intragroup comparisons across multiple time points were performed using the Friedman test with pairwise comparisons by the Dunn test. The relationships between m-YPAS scores and PAED scores at each time point were analyzed using Pearson's correlation coefficient for normally distributed data and Spearman's rho correlation coefficient for non-normally distributed data. Similarly, the relationships between pain scores and PAED scores were analyzed using Spearman's rho correlation coefficient. Quantitative variables are presented as mean ± standard deviation or median (minimum–maximum), and categorical variables as frequency (percentage). Statistical significance was set at $p<0.05$.

RESULTS

A total of 82 patients were initially assessed for eligibility. Of these, 18 were excluded for the following reasons: age >10 years (n = 2), allergy to study drugs (n = 2), risk of difficult airway (e.g., obesity) (n = 1), neurological or developmental disorders (n = 2), recent or active upper respiratory tract infection (n = 3), chronic medication use (e.g., anxiolytics, opioids, steroids) (n = 2), adverse events observed during emergence (hypoxia, hypercarbia, hypotension, hypoglycemia, or hypo-/hyperthermia) (n = 1), and lack of parental consent (n = 5).

There was no significant difference between the groups in terms of sex, age, weight, ASA risk classes, fasting times, operating room temperature, or anesthesia, emergence, and reanimation times. The patients in both groups had high m-YPAS scores, and there was no significant difference between the groups. It was observed that 82.8% (n=53) of the patients were anxious in the preoperative stage. Group 1 had a significantly shorter extubation time ($p<0.001$; Table 1).

The groups were similar in terms of their history of previous surgeries, status of premedication for anesthesia induction, intraoperative and postoperative additional drug use, postoperative complications, and parental satisfaction ($p>0.005$; Table 2).

Table 1: Comparison of quantitative data

	Group 1 (n=32)	Group 2 (n=32)	p
Gender (Female/Male)	14/18	15/17	1.000
Age (years)	5.47 ± 2.2	5.75 ± 2.4	0.618*
Body weight (kg)	23.53 ± 10.49	25.17 ± 13.05	0.554*
ASA Risk (I/II)	24/8	30/2	0.085
Room Temperature (^{0}C)	21.93 ± 1.51	21.85 ± 1.32	0.827**
Fasting Time (hours)	9.52 ± 1.63	9.84 ± 1.57	0.485*
Anesthetic Time (minutes)	58.13 ± 23.86	55.56 ± 23.96	0.614*
Extubation Time (minutes)	6.44 ± 2.41	9.06 ± 2.68	<0.001*
Recovery Time (minutes)	7.56 ± 3.45	9.25 ± 4.25	0.136*
Recovery Stay Time (minutes)	24.31 ± 6.39	22.91 ± 3.95	0.316*
m-YPAS score	45.76 ± 16.05	53.86 ± 22.15	0.160*

*Mann Whitney U test, **Independent two-sample t test, mean ± standard deviation;
ASA: American Society of Anesthesiologists, **m-YPAS:** Modified Yale Preoperative Anxiety Scale

Table 2: Comparison of categorical data

	Group 1 n=32 (%)	Group 2 n=32 (%)	p
Surgical History			
Yes	7 (21.9)	10 (31.3)	0.571**
No	25 (78.1)	22 (68.7)	
Premedication			
Yes	7 (21.9)	12 (37.5)	0.274**
No	25 (78.1)	20 (62.5)	
Intraoperative drug use			
No	8 (25)	4 (12.5)	0.664*
Prednol	9 (28)	12 (37.5)	
Jetocaine	6 (18.8)	4 (12.5)	
Prednol+Jetocaine	6 (18.8)	9 (28.1)	
Avil+Prednol+ Jetocaine	2 (6.3)	1 (3.1)	
Avil+Prednol	1 (3.1)	1 (3.1)	
Avil	0 (0)	1 (3.1)	
Postoperative complications			
No	13 (40.6)	8 (25)	0.196*
Cough	4 (12.5)	6 (18.8)	
Sore throat	4 (12.5)	3 (9.4)	
Cough+Sore throat	4 (12.5)	2 (6.3)	
Nausea+Cough	1 (3.1)	0 (0)	
Nausea+Cough+ Sore throat	2 (6.3)	0 (0)	
Headache	1 (3.1)	0 (0)	
Cough+Increased secretion	1 (3.1)	2 (6.3)	
Sore throat+Increased secretion	1 (3.1)	2 (6.3)	
Nausea+Vomiting+ Increased secretion	1 (3.1)	0 (0)	
Increased secretion	0 (0)	4 (12.5)	
Nausea+Vomiting	0 (0)	2 (6.3)	
Nausea+Vomiting+ Cough+Sore throat	0 (0)	1 (3.1)	
Cough+Increased secretion+Sore throat	0 (0)	2 (6.3)	
Parent Satisfaction			
Poor	1 (3.1)	0 (0)	0.601*
Acceptable	4 (12.5)	2 (6.3)	
Good	21 (65.6)	24 (75)	
Very Good	6 (18.8)	6 (18.8)	
Postoperative medication use			
None	25 (78.1)	26 (81.3)	0.550*
Analgesic	5 (15.6)	6 (18.7)	
Analgesic+ Antiemetic	1 (3.1)	0 (0)	
Analgesic+Nebul	1 (3.1)	0 (0)	

*Pearson's chi-square test, **Yates-corrected chi-square test, frequency (percentage);
Group 1: patients administered 10 mg/kg acetaminophen iv 30 min before surgery
Group 2: patients administered 10 mg/kg acetaminophen iv 15 min before surgery

While there was no intergroup difference in terms of the PAED and WBFPS scores of the patients, there were significant intragroup differences (p<0.001; Table 3). The median PAED scores of Group 1 were ≥10 within the postoperative first 10 min, while those of Group 2 were ≥10 up to the 15th min, and these scores decreased over time. None of the patients developed PED severe enough to require pharmacological treatment in the postoperative reanimation room.

Although WBFPS scores decreased over time in both groups, the presence of early pain showed a positive correlation between PAED and WBFPS (p<0.05; Table 4).

DISCUSSION

According to the results of our study, pediatric patients undergoing adenotonsillectomy surgery had generally high m-YPAS scores, and preoperative anxiety was common. The presence of parents alone was inadequate in reducing this anxiety. It was observed that preemptive acetaminophen administration reduced postoperative pain but did not have a significant effect on PED development. Our results demonstrated that preoperative anxiety and pain were significant factors in the development of PED following pediatric adenotonsillectomy, and pharmacological premedication was necessary for anxiety management.

Table 3: Comparison of PAED and WBFPS scores between and within groups

	Time (min)	Group 1	Group 2	p*
PAED				
	0	14 (10 - 20)^{a}	14.5 (6 - 20)^{ab}	0.925
	5	15 (8 - 20)^{a}	15 (3 - 19)^{a}	0.498
	10	12 (1 - 18)^{b}	12.5 (3 - 17)^{b}	0.850
	15	9.5 (0 - 17)^{c}	10 (0 - 16)^{c}	0.909
	30	6.5 (0 - 14)^{d}	5.5 (0 - 13)^{d}	0.407
	45	5 (0 - 10)^{e}	3.5 (0 - 12)^{e}	0.685
	60	3 (0 - 6)^{f}	3 (0 - 10)^{f}	0.521
	p**	<0.001	<0.001	
WBFPS				
	0	0 (0 - 8)^{d}	0 (0 - 8)^{cd}	0.371
	5	2 (0 - 8)^{bc}	2 (0 - 8)^{b}	0.231
	10	2 (0 - 10)^{a}	2 (0 -10)^{ab}	0.350
	15	2 (0 - 10)^{ab}	3 (0 -10)^{a}	0.108
	30	2 (0 - 8)^{bc}	2 (0 - 8)^{b}	0.600
	45	2 (0 - 4)^{c}	2 (0 - 8)^{c}	0.851
	60	0 (0 - 2)^{d}	0 (0 - 6)^{d}	0.351
	p**	<0.001	<0.001	

*Mann Whitney U test, mean ± standard deviation, median (minimum – maximum); ** Friedman test; a-f=There is no difference between times with the same letter.

Table 4: Relationship between m-YPAS-PAED and PAED-WBFPS scores of groups

	Time (min)	mYPAS-PAED	PAED -WBFPS
		r;p	r;p
Group 1	0	-0.085*; 0.645	-0.067**; 0.715
	5	0.084*; 0.649	0.302**; 0.093
	10	0.293*; 0.104	0.351**; 0.049
	15	0.211*; 0.246	0.420**; 0.017
	30	0.191*; 0.296	0.462**; 0.008
	45	0.087**; 0.637	0.243**; 0.181
	60	0.193**; 0.289	0.187**; 0.305
Group 2	0	0.178**; 0.330	0.003**; 0.985
	5	0.198**; 0.277	0.128**; 0.487
	10	0.276**; 0.126	0.401**; 0.023
	15	0.145**; 0.428	0.364**; 0.040
	30	0.400**; 0.023	0.378**; 0.033
	45	0.327**; 0.068	0.404**; 0.022
	60	0.340**; 0.057	0.466**; 0.007

*Pearson correlation coefficient, **r: Spearman's rho correlation coefficient; **m-YPAS:** Modified Yale Preoperative Anxiety Scale, **PAED:** Post Anesthetic Emergence Delirium, **WBYAS:** Wong Baker Facial Pain Scale

It was reported that 65% of children experienced intense anxiety in preoperative waiting areas and during anesthesia induction (20). It was emphasized that this anxiety led to the administration of sedation to children at high doses in the preoperative period and the development of agitation, pain, and behavioral problems in the postoperative period (21,22). In the literature, while it has been reported that there is no clear relationship between preoperative anxiety and PED, some studies have shown positive correlations between these two variables (4,6,18,23). It was reported that the presence of parents near their children before induction not only reduced preoperative anxiety and increased the cooperation of children but also lowered their agitation levels during emergence (23). In our study, high levels of anxiety persisted in the patients who were accompanied by their parents but not administered premedication, and the duration of PED presence in these patients was not significantly affected. This result suggested that the presence of parents alone was inadequate in preventing PED, and pharmacological premedication was necessary.

Post-tonsillectomy pain management in children is difficult (12). Preemptive analgesia has been reported to delay the onset of postoperative pain and reduce its severity and duration (1-12). Acetaminophen is prevalently used as a preemptive analgesic for post-tonsillectomy pain management in children and adults (24). In our study, we also used acetaminophen as a preemptive analgesic agent.

Pain is an important factor contributing to the development of PED, and the effective treatment of pain can reduce the prevalence and severity of PED (8). It was stated that preemptive analgesic administration could lower PED and pain levels. WBFPS is one of the measurement instruments that are based on the self-reports and behavioral analyses of children, and it is frequently employed (15). In this study, in which we used WBFPS scores to assess pain, while the pain scores of Group 1 were usually lower than those of Group 2, these differences were not statistically significant. On the other hand, there was a significant positive relationship between pain and PAED scores. In our opinion, preemptive acetaminophen administration in pediatric adenotonsillectomy cases is more effective in postoperative pain management. Considering the measurements made at all time points in this study, the patients in both groups had similar PAED and WBFPS scores, and these scores decreased over time.

PED is a clinical picture that usually spontaneously resolves after general anesthesia and lasts about 5-15 min (25). PED, which can be seen in up to 80% of children who are administered sevoflurane, is associated with rapid emergence and the anxiety status of the patient related to the ambient settings (26,27). In our study, the initial PAED scores of the patients in both groups were high. The duration of the presence of PED in our study was similar to those reported in other studies in the literature, and it was observed that as postoperative monitoring times increased, PAED scores decreased. The initially high PAED scores of the patients may be attributed to their rapid emergence from anesthesia related to sevoflurane usage. The shorter PED durations in Group 1 may also be explained by the lower rate of exposure to sevoflurane in these patients.

The limitations of this study included its small sample, short monitoring times, the absence of a placebo group, and the fact that individual differences were not investigated. Moreover, although we used scales with proven validity to collect data, the measurements and evaluations may have been influenced by a set of various factors and prevented us from a clearer understanding of the relationships between concepts. The combined usage of scales that can specifically distinguish preoperative anxiety, postoperative pain, and PED can reduce the likelihood of errors.

Conclusion

Preemptive analgesia reduced postoperative pain scores in pediatric patients. However, as it did not decrease PAED scores within the first 10 minutes postoperatively, it was concluded that it had no significant effect on PED. Further comprehensive studies focusing on both anxiety and pain management are warranted for the prevention of PED in pediatric patients.

Acknowledgments

We would like to thank Assoc. Prof. Naci Murat, a faculty member of the Department of Statistics, Faculty of Science and Literature, Ondokuz Mayıs University, for his contributions to the statistical analysis process in this study.

Author Contributions

Concept: **Gamze Küçükosman, Keziban Bollucuoğlu, Bengü Gülhan Köksal,** Design: **Gamze Küçükosman, Bengü Gülhan Köksal,** Data Collection or Processing: **Dilay İlke Alp, Özcan Pişkin, Rahşan Dilek Okyay,** Analysis or Interpretation: **Gamze Küçükosman, Bengü Gülhan Köksal, Hilal Ayoğlu, Naci Murat,** Literature search: **Dilay İlke Alp, Gamze Küçükosman,** Writing: **Gamze Küçükosman, Dilay İlke Alp, Hilal Ayoğlu,** Approval: **Gamze Küçükosman, Hilal Ayoğlu.**

Conflicts of Interest

The authors declare no conflict of interest.

Financial Support

There is no financial disclosure to declare in this study.

Ethical Approval

This study was approved by the Zonguldak Bülent Ecevit University Non-Interventional Clinical Research Ethics Committee, with an approval number of protocol No: 2019/145-18/09 and an approval date of 18.09.2019. ClinicalTrials.gov Identifier: NCT05802082)

REFERENCES

1. Aldamluji N, Burgess A, Pogatzki-Zahn E, Raeder J, Beloeil H; PROSPECT Working Group collaborators*. PROSPECT guideline for tonsillectomy: systematic review and procedure-specific postoperative pain management recommendations. Anaesthesia. 2021;76(7):947-961. https://doi.org/10.1111/anae.15491
2. Albornoz AE, Rana M, Hayes J, Englesakis M, Tsang M, Amin R, et al. Perioperative clinical practice recommendations for pediatric tonsillectomy: a systematic review. Can J Anaesth. 2024;71(2):187-200. https://doi.org/10.1007/s12630-023-02563-1
3. Redondo-Enríquez JM, Rivas-Medina M, Galán-Mateos MM. Updating Clinical Practice: Improving Perioperative Pain Management for Adeno-Tonsillectomy in Children. Children (Basel). 2024;11(10):1190. https://doi.org/10.3390/children11101190
4. Yang X, Lin C, Chen S, Huang Y, Cheng Q, Yao Y. Remimazolam for the prevention of emergence delirium in children following tonsillectomy and adenoidectomy under sevoflurane anesthesia: a randomized controlled study. Drug Des Devel Ther. 2022;16:3413-3420. https://doi.org/10.2147/DDDT.S379870
5. Zieliński J, Morawska-Kochman M, Dudek K, Czapla M, Zatoński T. The effect of pre-emptive analgesia on the postoperative pain in pediatric otolaryngology: a randomized, controlled trial. J Clin Med. 2022;11(10):2713. https://doi.org/10.3390/jcm11102713
6. Liu K, Liu C, Ulualp SO. Prevalence of emergence delirium in children undergoing tonsillectomy and adenoidectomy. Anesthesiol Res Pract. 2022;2022:1465999. https://doi.org/10.1155/2022/1465999
7. Ingelmo PM, Somaini M, Engelhardt T. Emergence delirium may not just be a recovery room problem. Paediatr Anaesth. 2021;31(10):1022-1024. https://doi.org/10.1111/pan.14238
8. Cettler M, Zielińska M, Rosada-Kurasińska J, Kubica-Cielińska A, Jarosz K, Bartkowska-Śniatkowska A. Guidelines for treatment of acute pain in children - the consensus statement of the Section of Paediatric Anaesthesiology and Intensive Therapy of the Polish Society of Anaesthesiology and Intensive Therapy. Anaesthesiol Intensive Ther. 2022;54(3):197-218. https://doi.org/10.5114/ait.2022.118693
9. Urits I, Peck J, Giacomazzi S, Patel R, Wolf J, Mathew D, et al. Emergence delirium in perioperative pediatric care: a review of current evidence and new directions. Adv Ther. 2020;37(5):1897-1909. https://doi.org/10.1007/s12325-020-01268-3
10. Zieliński J, Morawska-Kochman M, Zatoński T. Pain assessment and management in children in the postoperative period: A review of the most commonly used postoperative pain assessment tools, new diagnostic methods and the latest guidelines for postoperative pain therapy in children. Adv Clin Exp Med. 2020;29(3):365-374. https://doi.org/10.17219/acem/112600
11. Xuan C, Yan W, Wang D, Li C, Ma H, Mueller A, et al. Efficacy of preemptive analgesia treatments for the management of postoperative pain: a network meta-analysis. Br J Anaesth. 2022;129(6):946-958. https://doi.org/10.1016/j.bja.2022.08.015
12. Song IK, Park YH, Lee JH, Kim JT, Choi IH, Kim HS. Randomized controlled trial on preemptive analgesia for acute postoperative pain management in children. Paediatr Anaesth. 2016;26(4):438-443. https://doi.org/10.1111/pan.12854
13. Kain ZN, Mayes LC, Cicchetti DV, Bagnall AL, Finley JD, Hofstadter MB. The Yale Preoperative Anxiety Scale: how does it compare with a "gold standard"? Anesth Analg. 1997;85(4):783-788. https://doi.org/10.1097/00000539-199710000-00012
14. Hatipoğlu Z, Kırdök O, Özcengiz D. Validity and reliability of the Turkish version of the modified yale preoperative anxiety scale. Turk J Med Sci. 2019;49(3):730-737. doi: 10.3906/sag-1612-113.
15. Wong DL, Baker CM. Pain in children: Comparison of assessment scales. Pediatr Nurs. 1988;14(1):9-17.
16. Sikich N, Lerman J. Development and psychometric evalu ation of the pediatric anesthesia emergence delirium scale. Anesthesiology. 2004;100:1138-1145. https://doi.org/10.1097/00000542-200405000-00015
17. Karasu D, Karaca U, Özgünay SE, Yılmaz C, Yetik F, Özkaya G. The frequency of emergence delirium in children undergoing outpatient anaesthesia for magnetic resonance imaging. Int J Clin Pract. 2021;75(11):e14763. https://doi.org/10.1111/ijcp.14763
18. Yucedag F, Sezgin A, Bilge A, Basaran B. The effect on perioperative anxiety and emergence delirium of the child or parent's preference for parental accompaniment during anesthesia induction in children undergoing adenotonsillectomy surgery. Paediatr Anaesth. 2024;34(8):792-799. https://doi.org/10.1111/pan.14677
19. Polat R, Peker K, Baran I, Bumin Aydın G, Topçu Gülöksüz Ç, Dönmez A. Comparison between dexmedetomidine and remifentanil infusion in emergence agitation during recovery after nasal surgery: A randomized double-blind trial [retracted in: Anaesthesiologie. 2024;73(1):64. https://doi.org/10.1007/s00101-022-01177-4

20. Kain ZN, Mayes LC, Caldwell-Andrews AA, Karas DE, McClain BC. Preoperative anxiety, postoperative pain, and behavioral recovery in young children undergoing surgery. Pediatrics. 2006;118(2):651-658. https://doi.org/10.1542/peds.2005-2920
21. Chieng YJ, Chan WC, Liam JL, Klainin-Yobas P, Wang W, He HG. Exploring influencing factors of postoperative pain in school-age children undergoing elective surgery. J Spec Pediatr Nurs. 2013;18(3):243-252. https://doi.org/10.1111/jspn.12033
22. Jeizan P, Baharimehr K, Kamvar R, Abolghasemi Fard A, Zojaji S, Karimi A, et al. Dexmedetomidine Leads to Less Emergence Delirium Compared to Midazolam in Pediatric Tonsillectomy and/or Adenoidectomy: A Systematic Review and Meta-Analysis. Cureus. 2025;17(4):e81686. https://doi.org/10.7759/cureus.81686
23. Yao J, Gong H, Zhao X, Peng Q, Zhao H, Yu S. Parental presence and intranasal dexmedetomidine for the prevention of anxiety during anesthesia induction in children undergoing tonsillectomy and/or adenoidectomy surgery: A randomized controlled trial. Front Pharmacol. 2022;13:1015357. https://doi.org/10.3389/fphar.2022.1015357
24. Pierce CA, Voss B. Efficacy and safety of ibuprofen and acetaminophen in children and adults: a meta-analysis and qualitative review. Ann Pharmacother. 2010;44(3):489-506. https://doi.org/10.1345/aph.1M332
25. Prayunanto E, Widyastuti Y, Sari D. The association of modified Yale perioperative anxiety scale and pediatric anesthesia behavior on postoperative emergence delirium in children: A prospective cohort study. Bali J Anaesthesiol. 2023;7:88-93. https://doi.org/10.4103/bjoa.bjoa_24_23
26. Kain ZN, Caldwell-Andrews AA, Maranets I, McClain B, Gaal D, Mayes LC, et al. Preoperative anxiety and emergence delirium and postoperative maladaptive behaviors. Anesth Analg. 2004;99(6):1648-1654. https://doi.org/10.1213/01.ANE.0000136471.36680.97
27. Farag RS, Spicer AC, Iyer G, Stevens JP, King A, Bain PA, et al. Incidence of emergence agitation in children undergoing sevoflurane anesthesia compared to isoflurane anesthesia: An updated systematic review and meta-analysis. Paediatr Anaesth. 2024;34(4):304-317. https://doi.org/10.1111/pan.14596

APPENDIX 1. The mYPAS

A. Activity

1 = Looking around, curious, playing with toys, reading (or other age-appropriate behavior); moves around holding area/treatment room to get toys or go to parent;may move toward OR equipment.

2 = Not exploring or playing, may look down, may fidget with hands or suck thumb (blanket); may sit close to parent while waiting, or play has a definite manic quality.

3 = Moving from toy to parent in unfocused manner, nonactivity-derived movements; frenetic/frenzied movement or play; squirming, moving on table, may push mask away, or clinging to parent.

4 = Actively trying to get away, pushes with feet and arms, may move whole body; in waiting room, running around unfocused, not looking at toys or will not separate from parent, desperate clinging.

B. Vocalizations

1 = Reading (nonvocalizing appropriate to activity), asking questions, making comments, babbling, laughing, readily answers questions but may be generally quiet; child too young to talk in social situations or too engrossed in play to respond.

2 = Responding to adults but whispers, “baby talk,” only head nodding.

3 = Quiet, no sounds or responses to adults.

4 = Whimpering, moaning, groaning, silently crying.

5 = Crying or may be screaming “no.”

6 = Crying, screaming loudly, sustained (audible through mask).

C. Emotional expressivity

1 = Manifestly happy, smiling, or concentrating on play.

2 = Neutral, no visible expression on face.

3 = Worried (sad) to frightened, sad, worried, or tearful eyes.

4 = Distressed, crying, extremely upset, may have wide eyes.

D. State of apparent arousal

1 = Alert, looks around occasionally, notices or watches what anesthesiologist does with him/her (could be relaxed).

2 = Withdrawn, child sitting still and quiet, may be sucking on thumb or face turned into adult.

3 =Vigilant, looking quickly all around, may startle to sounds, eyes wide, body tensed.

4 = Panicked whimpering, may be crying or pushing others away, turns away.

E. Use of parents

1 = Busy playing, sitting idle, or engaged in age-appropriate behavior and does not need parent; may interact with parent if parent initiates the interaction.

2 = Reaches out to parent (approaches parent and speaks to otherwise silent parent), seeks and accepts comfort, may lean against parent.

3 = Looks at parents quietly, apparently watches actions, does not seek contact or comfort, accepts it if offered or clings to parent.

4 = Keeps parent at distance or may actively withdraw from parent, may push parent away or desperately clinging to parent and will not let parent go.

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):151-156
DOI: 10.29058/mjwbs.1795433

Relationship between demodex infestation and clinical parameters in polycystic ovary syndrome: A prospective observational study

Polikistik over sendromunda demodex enfestasyonu ile klinik parametreler arasındaki ilişki: Prospektif gözlemsel çalışma

Mehmet Bülbül[1], Mustafa Kaplanoğlu[2], Tuncay Çelik[3], İpek Çakılkaya[4], Emel Dündar[5]

[1]Department of Obstetrics and Gynecology, Zonguldak Bülent Ecevit University Faculty of Medicine, Zonguldak, Türkiye
[2]Department of Obstetrics and Gynecology, Medical Park Hospital, Adana, Türkiye
[3]Department of Parasitology, Adıyaman University Faculty of Medicine, Adıyaman, Türkiye
[4]Department of Obstetrics and Gynecology, Private Öztan Hospital, Uşak, Türkiye
[5]Family Health Center, Adıyaman Provincial Health Directorate, Adıyaman, Türkiye

Cite this article as: *Bülbül M et al. Relationship between demodex infestation and clinical parameters in polycystic ovary syndrome: a prospective observational study. Med J West Black Sea. 2026;10(1):151-156.*

ABSTRACT

Aim: Polycystic ovary syndrome (PCOS) is frequently accompanied by dermatologic findings and metabolic disturbances. Because Demodex mites thrive in sebaceous areas and have been linked to several inflammatory and metabolic conditions, we explored whether Demodex positivity is associated with clinical and laboratory features in women with PCOS.

Material and Methods: In this prospective observational study, 63 women aged 18–40 years with a confirmed diagnosis of PCOS based on the Rotterdam criteria were enrolled. The presence of Demodex spp. was assessed using the standardized superficial skin biopsy technique. Sociodemographic data, anthropometric measurements, ovarian volume, oral glucose tolerance test results, and a panel of hormonal parameters were collected.

Results: Demodex-positive patients had significantly higher body mass index, waist and abdominal circumference, ovarian volume, and fasting glucose levels, while their waist-to- abdominal ratio was lower compared to Demodex-negative patients ($p < 0.05$). In contrast, multivariate logistic regression analysis revealed that none of these parameters remained independently significant ($p > 0.05$).

Conclusion: The findings suggest that Demodex spp. infestation in women with PCOS may reflect both dermatological and metabolic alterations. Although univariate analyses indicated significant associations, these did not persist in multivariate models, underscoring the need for larger, prospective studies. Screening for Demodex may provide additional insight into the comprehensive management of PCOS, particularly in patients presenting with dermatological complaints.

Keywords: Polycystic ovary syndrome, demodex spp., fasting glucose, insulin resistance, hormonal profile, ovarian volume

ÖZ

Amaç: Polikistik over sendromu (PKOS), ovulatuvar disfonksiyon, hiperandrojenizm ve polikistik over morfolojisi ile karakterize, sıklıkla metabolik bozukluklarla birlikte görülen yaygın bir endokrin hastalıktır. PKOS'ta akne ve sebore gibi cilt belirtileri sık görülür ve cildin yağ bakımından zengin bölgelerinde yaşayan Demodex türleri çeşitli dermatolojik ve metabolik rahatsızlıklarla ilişkilendirilmiştir. Bu çalışmanın amacı, PKOS'lu kadınlarda Demodex türleri enfestasyonu yaygınlığını belirlemek ve antropometrik, hormonal ve glisemik parametrelerle ilişkisini değerlendirmektir.

Gereç ve Yöntemler: Bu prospektif gözlemsel çalışmaya, Rotterdam kriterlerine göre doğrulanmış PKOS tanısı almış 18-40 yaş arası 63 kadın dahil edildi. Demodex türlerinin varlığı, standart yüzeyel deri biyopsisi tekniği kullanılarak değerlendirildi. Sosyodemografik veriler, antropometrik ölçümler, over hacmi, oral glukoz tolerans testi sonuçları ve bir dizi hormonal parametre toplandı.

Bulgular: Demodex pozitif hastaların vücut kütle indeksi (VKİ), bel ve abdominal çevresi, over hacmi ve açlık kan şekeri düzeyleri anlamlı derecede daha yüksekken, bel-abdominal oranları Demodex negatif hastalara kıyasla daha düşüktü ($p<0,05$). Buna karşılık, çok değişkenli lojistik regresyon analizi bu parametrelerin hiçbirinin bağımsız olarak anlamlı olmadığını ortaya koydu ($p>0,05$).

Sonuç: Bulgular, PKOS'lu kadınlarda Demodex spp. enfestasyonunun hem dermatolojik hem de metabolik değişiklikleri yansıtabileceğini düşündürmektedir. Tek değişkenli analizler anlamlı ilişkiler gösterse de bunlar çok değişkenli modellerde kalıcı olmamıştır ve bu da daha büyük, prospektif çalışmalara duyulan ihtiyacı vurgulamaktadır. Demodex taraması, özellikle dermatolojik şikayetlerle başvuran hastalarda PKOS'un kapsamlı yönetimine ilişkin ek bilgi sağlayabilir.

Anahtar Kelimeler: Polikistik over sendromu, demodex türleri, açlık glikozu, insülin direnci, hormonal profil, over hacmi

Corresponding Author: Mehmet Bülbül ✉ mehmetbulbulmd@gmail.com **Received:** 02.10.2025 **Revision:** 26.01.2026 **Accepted:** 23.03.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- The prevalence of Demodex spp. was evaluated in women with polycystic ovary syndrome using a standardized superficial skin biopsy method.
- Demodex-positive patients had higher BMI, waist and abdominal circumference, ovarian volume, and fasting glucose levels.
- Waist-to-abdominal ratio was significantly lower in Demodex-positive individuals.
- No independent predictors of Demodex positivity were identified in multivariate logistic regression analysis.
- The findings suggest a possible association between metabolic alterations and Demodex infestation in PCOS.

INTRODUCTION

Polycystic ovary syndrome (PCOS) is a common endocrine disorder in reproductive-aged women, defined by ovulatory dysfunction (anovulation or oligo-ovulation), clinical and/or biochemical hyperandrogenism, and polycystic ovarian morphology. Its estimated prevalence in the general population is approximately 10%, and although its etiology is thought to be multifactorial, it has not yet been fully elucidated (1-5). PCOS presents with a wide clinical spectrum; hirsutism, obesity, acne, and alopecia are among the most common manifestations. Although acne is not included in the diagnostic criteria, it is a frequent complaint that can significantly impair quality of life. While the role of androgens in acne development has long been recognized, androgen levels may remain within the normal range in some PCOS patients with acne. This observation suggests that, in addition to systemic hormone levels, factors such as increased peripheral tissue sensitivity to androgens may contribute to acne pathogenesis (3,5).

Demodex species (particularly Demodex folliculorum and Demodex brevis) are saprophytic or opportunistic ectoparasites that reside in areas of the human face with high sebaceous gland density. While they may exist in limited numbers as part of the normal skin flora, they have also been associated with increased sebum production, aging, immunosuppression, rosacea, blepharitis, acne, and diabetes (6-14). Recent studies have suggested that disturbances in glucose metabolism may increase Demodex density, pointing to a possible association with insulin resistance or hyperglycemia (10-13). Since PCOS is frequently accompanied by metabolic disturbances such as glucose intolerance, hyperinsulinemia, and insulin resistance, the potential relationship between Demodex infestation and PCOS gains clinical significance. Therefore, in the etiopathogenesis of acne observed in PCOS patients, not only hormonal factors but also parasitic infestations should be considered.

Nevertheless, studies directly examining the relationship between Demodex infestation and impaired glucose metabolism in PCOS are quite limited, and current data remain insufficient (11,12). Given the complex endocrine and metabolic profile of PCOS, the role of parasitic factors such as Demodex has not been well defined. Accordingly, the present study aimed to determine the prevalence of Demodex infestation in women with PCOS and to investigate its associations with glucose regulation, hormonal profiles, and anthropometric parameters in detail.

MATERIAL and METHODS

The study was designed prospectively and carried out between November 1, 2020, and April 30, 2021, at the Department of Obstetrics and Gynecology of Adıyaman University Faculty of Medicine. Women aged 18–40 years presenting with menstrual irregularities and signs of hyperandrogenism, who were diagnosed with PCOS based on the Rotterdam criteria, were invited to participate. All participants provided written informed consent. The study protocol was approved by the Non-Interventional Clinical Research Ethics Committee of Adıyaman University (approval date: 17 November 2020; decision number: 2020/10-10). This study was conducted with the utmost ethical sensitivity in accordance with the Principles of the Declaration of Helsinki.

A total of 63 women diagnosed with PCOS and reporting acne complaints were included, provided they were not morbidly obese (body mass index [BMI]>40 kg/m^2), had no known systemic diseases, and met the inclusion criteria. For each participant, demographic data (age, BMI, obstetric history, smoking status, anthropometric measurements), laboratory results, and ultrasonographic findings were recorded. Abdominal circumference was preferred over hip circumference to better reflect central adiposity, which is more closely associated with metabolic activity in PCOS.

Exclusion criteria were: incomplete follow-up or insufficient data, systemic diseases associated with immunosuppression, a history of hormonal therapy within the past year, allergic diseases, previously diagnosed dermatological conditions (e.g., rosacea, seborrheic dermatitis, blepharitis, allergic dermatitis), and antibiotic use within the last month.

The presence of Demodex spp. was determined using the standard superficial skin biopsy (SSSB) method, performed on the nasolabial and chin regions as previously described. A 1 cm² area was marked on a microscope slide with a ruler, and the remaining surface was cleaned with ether. A drop of cyanoacrylate-based adhesive was applied to the marked area, and the slide was gently pressed onto the skin surface. After approximately one minute, the slide was carefully removed and examined within one hour. Two to three drops of immersion oil were applied to the sampled area, covered with a coverslip, and examined microscopically. An initial survey was performed under 4x magnification, followed by scanning at 10x and 40x magnification with a partially closed diaphragm. The density of Demodex spp. per cm² was evaluated by an experienced parasitologist (TÇ), who was blinded to the clinical and laboratory data of participants. A positive diagnosis was defined as the presence of five or more Demodex spp. per cm² (9,10).

Statistical Analysis

Statistical evaluation was performed using SPSS version 21 (IBM Corp., Armonk, NY, USA). A p value of less than 0.05 was considered significant. Continuous variables were reported as mean ± standard deviation, while categorical variables were expressed as percentages. Categorical comparisons were made with Pearson's chi-square test. The Shapiro-Wilk test was used to assess normality. According to the distribution pattern, continuous variables were compared using either the independent samples t-test or the Mann-Whitney U test.

To identify factors associated with Demodex positivity, a multivariate logistic regression analysis was performed using the backward stepwise (Wald) method. In the initial step, body mass index, waist circumference, abdominal circumference, total ovarian volume, fasting glucose level, and waist-to-abdominal ratio were entered into the model. At each subsequent step, variables with the highest p-values were sequentially removed according to the Wald statistic.

RESULTS

A total of 63 women diagnosed with PCOS were included in the study. Of these, 38 (60.3%) were identified as Demodex-positive, while 25 (39.7%) were Demodex-negative (Table 1).

When sociodemographic and anthropometric characteristics were compared, BMI (26.2 ± 5.9 vs. 23.3 ± 2.5; p = 0.026), waist circumference (79.3 ± 11.8 cm vs. 71.2 ± 6.9 cm; p = 0.003), and abdominal circumference (86.2 ± 12.9 cm vs. 75.3 ± 10.3 cm; p = 0.001) were significantly higher in the Demodex-positive group compared to the Demodex-negative group. However, the waist-to-abdominal ratio was lower in the Demodex positive group (0.92 ± 0.05 vs. 0.95 ± 0.05; p = 0.013). No significant differences were observed between the groups with respect to other demographic variables such as age, age at menarche, body weight, height, gravida, parity, or marital status (p > 0.05).

Laboratory data of the participants are presented in Table 2. Comparisons of laboratory findings showed that the mean fasting glucose level was significantly higher in the Demo-

Table 1: Sociodemographic data of participants according to Demodex positivity.

		Demodex Negative	Demodex Positive	p-value
	n	25	38	
Age (years, mean ± SD)		23.5±4.3	22.8±4.9	0.561*
Age of first menstruation (years, mean ± SD)		12.9±1.4	13.0±1.1	0.707*
Marital status (n (%))	Married	6 (24.0)	16 (42.1)	0.140
	Single	19 (76.0)	22 (57.9)	
Gravida (mean ± SD)		0.4±0.9	0.4±0.9	0.804**
Parita (mean ± SD)		0.4±0.9	0.3±0.8	0.932**
Abortus (mean ± SD)		0.00±0.00	0.03±0.16	0.422**
Height (cm, mean ± SD)		164.7±6.6	161.8±7.5	0.119*
Weight (kg, mean ± SD)		63.4±8.4	68.5±15.2	0.325**
BMI (kg/m², mean ± SD)		23.3±2.5	26.2±5.9	0.026*
Waist circumference (cm, mean ± SD)		71.2±6.9	79.3±11.8	0.003*
Abdominal circumference (cm, mean ± SD)		74.9±10.3	86.2±12.9	<0.001*
Waist-to-abdominal ratio (mean ± SD)		0.95±0.05	0.92±0.05	0.013*
Acanthosis (n, (%))		2 (8.0)	3 (7.9)	0.988

BMI: Body mass index, **SD:** Standard deviation. *Independent samples t-test, **Mann-Whitney U test

Table 2: Laboratory findings of participants according to Demodex positivity.

		Demodex Negative	Demodex Positive	p-value
	n	**25**	**38**	
HbA1c (%, mean ± SD)		5.3±0.4	5.3±0.5	0.719**
75 gr OGTT (mg/dl, mean ± SD)	Fasting Glucose	91.6+9.9	98.3±8.5	0.036**
	T1 (60 min)	123.0±34.3	135.0±47.1	0.428*
	T2 (120 min)	114.6±28.2	105.5±23.8	0.887**
Free T3 (pg/ml, mean ± SD)		3.7±1.1	3.4±0.4	0.486**
Free T4 (ng/dl, mean ± SD)		0.92±0.14	0.87±0.18	0.290**
TSH (IU/mL, mean ± SD)		1.86±1.15	2.17±1.87	0.459**
AMH (ng/ml, mean ± SD)		6.2±3.8	6.9±3.3	0.356*
DHEA-S (μg/dl, mean ± SD)		270.3±105.3	305.4±91.7	0.167**
FSH (IU/mL, mean ± SD)		7.6±2.5	8.4±5.1	0.447**
LH (IU/mL, mean ± SD)		8.0±3.7	8.2±3.5	0.905**
Vitamin D (ng/mL, mean ± SD)		12.0±5.1	12.7±5.5	0.621*
Hemoglobin (gr/dl, mean ± SD)		13.8±1.4	13.8±1.5	0.898*
Hematocrit (%, mean ± SD)		40.7±3.3	40.3±3.6	0.647*
Platelet (x10^3/μl, mean ± SD)		267.2±62.7	273.3±71.4	0.730*
Total antral follicle (mean ± SD)		20.3±13.6	22.8±12.2	0.455**
Total Ovarian Volume (cm^3, mean ± SD)		12.9±3.5	16.3±4.9	0.005*

SD: Standard deviation; **HbA1c:** Glycated hemoglobin; **TSH:** Thyroid-stimulating hormone; **AMH:** Anti-Müllerian hormone; **DHEA-S:** Dehydroepiandrosterone sulfate; **FSH:** Follicle-stimulating hormone; **LH:** Luteinizing hormone, *Independent samples t-test, **Mann-Whitney U test.

dex-positive group (98.3 ± 8.5 mg/dL vs. 91.6 ± 9.9 mg/dL; p = 0.036). In addition, the total ovarian volume was also significantly greater in this group (16.3 ± 4.9 mL vs. 12.9 ± 3.5 mL; p = 0.005). No statistically significant differences were observed between the groups regarding other hormonal, metabolic, or hematological parameters (p > 0.05).

In the multivariate logistic regression analysis using the backward stepwise (Wald) method, none of the included variables remained statistically significant in the final model (all p > 0.05), indicating that no parameter independently predicted Demodex positivity.

DISCUSSION

This study aimed to contribute to the limited literature by examining the association between Demodex spp. infestation and anthropometric, hormonal, and glycemic parameters in women with PCOS. Our findings demonstrated that body mass index, waist and abdominal circumference, ovarian volume, and fasting glucose levels were significantly higher in Demodex-positive individuals in univariate analyses. However, none of these parameters remained statistically significant as independent predictors in the multivariate logistic regression model. This finding may reflect the complex and interrelated metabolic-endocrine characteristics of PCOS, in which factors such as insulin resistance, obesity, and hyperandrogenism are closely intertwined. Consequently, the associations observed in univariate analyses may be attributable to collinearity among variables or the influence of shared underlying metabolic pathways rather than independent effects.

Previous studies have reported positive correlations between impaired glucose metabolism and increased Demodex density (11,12). In contrast, although fasting glucose levels were higher in the Demodex-positive group in our study, this variable did not remain an independent predictor. This finding indicates that the unique metabolic profile of PCOS—particularly hyperinsulinemia and insulin resistance, which are independent of glucose levels—may directly influence Demodex proliferation. Similarly, previous studies reported that even after glycemic control was achieved in women with gestational diabetes, elevated Demodex density persisted (10). These observations suggest that not only systemic glucose levels but also local skin microenvironmental factors and sebaceous gland activity play important roles (15,16).

From an anthropometric perspective, BMI, waist circumference, and abdominal circumference were significantly higher in the Demodex-positive group. However, these effects also disappeared in multivariate analyses. This may imply that increased adiposity indirectly promotes sebum produc-

tion and thereby creates a more favorable environment for Demodex, but the relationship is not necessarily direct.

Although no distinct hormonal differences were found, total ovarian volume was significantly greater in the Demodex-positive group. This common feature of PCOS may elevate local androgen levels and sebum secretion, thereby facilitating Demodex infestation. Nevertheless, the absence of independent significance in regression analysis suggests that this relationship is also likely indirect or mediated by glucose-related mechanisms.

The therapeutic implications of these findings are noteworthy. Metformin, commonly used in the management of PCOS, not only improves insulin sensitivity but may also reduce sebum production, potentially limiting Demodex proliferation. Likewise, combined oral contraceptives suppress androgen activity and reduce sebaceous gland secretion, which could lower parasite density. Therefore, integrating systemic metabolic approaches with dermatological interventions may be beneficial in managing PCOS patients (16). All these speculative findings deserve investigation through randomized controlled trials.

The discrepancies between our results and those of previous studies may stem from methodological differences, such as sample size, diagnostic methods for Demodex, sampling sites, and the homogeneity of the study population. The restriction of this study to women with PCOS limits the generalizability of the findings to the broader population. Nevertheless, considering the unique hormonal and metabolic characteristics of PCOS, such focused analyses provide valuable insights.

This study has some limitations. The relatively small sample size may have limited statistical power. As this was designed as a pilot study, no power analysis was conducted, and the results should be considered hypothesis-generating rather than confirmatory. This may be the reason for the discrepancy between univariate and multivariate analyses. Furthermore, Demodex assessment was restricted to two facial regions (nasal wings and chin), which may not reflect the overall infestation level across the entire facial surface. The fact that only PCOS patients with acne complaints were included in the study may create a selection bias. Hormonal and metabolic parameters were measured at a single time point, which may have overlooked circadian or cyclic variations. Lastly, given the cross-sectional design, the observed associations should not be interpreted as causal.

Conclusion

This study demonstrated that Demodex spp. infestation in women with PCOS may be associated with certain anthropometric and metabolic parameters. However, the absence of independent predictors in multivariate analyses suggests that these relationships may not be direct but rather mediated through indirect mechanisms. The results highlight a potential interaction between the complex hormonal and metabolic profile of PCOS and the cutaneous microecology.

The evaluation of Demodex infestation may contribute to a better understanding and management of dermatological symptoms in women with PCOS. In particular, in PCOS patients presenting with persistent acne or other dermatological manifestations, a simple Demodex screening could guide the inclusion of topical antiparasitic agents (e.g., metronidazole, ivermectin) in treatment plans. Such an approach may improve the dermatological management of PCOS and support the development of multidisciplinary treatment strategies.

Acknowledgments

The authors would like to thank all the patients who participated in this study and the staff of the Department of Obstetrics and Gynecology, Adıyaman University Faculty of Medicine, for their valuable support during data collection.

Author Contributions

Study conception and design: **Mehmet Bülbül, Mustafa Kaplanoğlu, Tuncay Çelik;** data collection: **Mehmet Bülbül, Mustafa Kaplanoğlu, İpek Çakılkaya, Emel Dündar;** analysis and interpretation of results: **Mehmet Bülbül, Mustafa Kaplanoğlu, Tuncay Çelik;** draft manuscript preparation: **Mehmet Bülbül, Mustafa Kaplanoğlu, Tuncay Çelik, İpek Çakılkaya, Emel Dündar.** The authors reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that they have no conflict of interest relevant to this study.

Financial Support

This research was not supported by any institution or organization.

Ethical Approval

This study was approved by the Non-Interventional Clinical Research Ethics Committee of Adıyaman University (Approval Date: 17.11.2020, Decision No: 2020/10-10). All participants provided written informed consent prior to enrollment.

Data Availability Statement

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES

1. Siddiqui S, Mateen S, Ahmad R, Moin S. A brief insight into the etiology, genetics, and immunology of polycystic ovarian syndrome (PCOS). J Assist Reprod Genet. 2022;39(11):2439-2473. https://doi.org/10.1007/s10815-022-02625-7
2. Gleicher N, Darmon S, Patrizio P, Barad DH. Reconsidering the polycystic ovary syndrome (PCOS). Biomedicines. 2022;10(7):1505. https://doi.org/10.3390/biomedicines10071505

3. Yang J, Chen C. Hormonal changes in PCOS. J Endocrinol. 2024;261(1):e230342. https://doi.org/10.1530/JOE-23-0342
4. Helvaci N, Yildiz BO. Polycystic ovary syndrome as a metabolic disease. Nat Rev Endocrinol. 2025;21(4):230-244. https://doi.org/10.1038/s41574-024-01057-w
5. Lie Fong S, Douma A, Verhaeghe J. Implementing the international evidence-based guideline of assessment and management of polycystic ovary syndrome (PCOS): how to achieve weight loss in overweight and obese women with PCOS? J Gynecol Obstet Hum Reprod. 2021;50(6):101894. https://doi.org/10.1016/j.jogoh.2020.101894
6. Chang YS, Huang YC. Role of Demodex mite infestation in rosacea: A systematic review and meta-analysis. J Am Acad Dermatol. 2017;77(3):441-447.e6. https://doi.org/10.1016/j.jaad.2017.03.040
7. Trave I, Salvi I, Canepa P, Parodi A, Cozzani E. Detection of demodex mites in papulopustular rosacea using microscopic examination and polymerase chain reaction: a comparative case-control study. Arch Dermatol Res. 2024;316(7):485. https://doi.org/10.1007/s00403-024-03228-1
8. Kim J, Choi S, Choi S, Lee YI, Kim J, Lee JH, et al. Contact hypersensitivity and Demodex mite infestation in patients with rosacea: a retrospective cohort analysis. Eur J Dermatol. 2022;32(6):716-723. https://doi.org/10.1684/ejd.2022.4358
9. Bakacak Z, Kaplanoğlu M, Bakacak M, Çelik T. Demodex folliculorum mite infestation in gynecological cancers: A case control study. Eur J Gynaecol Oncol. 2020;41(4):583-586. https://doi.org/10.31083/j.ejgo.2020.04.3687
10. Kaplanoğlu M, Bulbul M, Çelik T. High incidence of Demodex Folliculorum in women treated for gestational diabetes mellitus persists after treatment. Acta Med Eur. 2025;7(1):1-4. https://doi.org/10.5281/zenodo.14546805
11. Eroglu S, Cakmakliogullari M, Kal Cakmakliogullari E. Is the presence of Demodex folliculorum increased with impaired glucose regulation in polycystic ovary syndrome? J Obstet Gynaecol. 2020;40(4):546-550. https://doi.org/10.1080/01443615.2019.1634026
12. Toka Özer T, Akyürek Ö, Durmaz S. Association between Demodex folliculorum and Metabolic Syndrome. J Cosmet Dermatol. 2020;19(11):3145-3149. https://doi.org/10.1111/jocd.13721
13. Enginyurt O, Karaman U, Cetin F, Ozer A. The Prevalence of Demodex Species and Its Relationship With the Metabolic Syndrome in Women of Malatya Province, Turkey. Jundishapur J Microbiol. 2015;8(10):e24322. https://doi.org/10.5812/jjm.24322
14. Lacey N, Delaney S, Kavanagh K, Powell FC. Mite-related microbial antigens stimulate inflammatory cells in rosacea. Br J Dermatol. 2007;157(3).474-481. https://doi.org/10.1111/j.1365-2133.2007.08056.x
15. Forton F, De Maertelaer V. Rosacea, skin immunity and Demodex: Pityriasis folliculorum as a missing link. J Eur Acad Dermatol Venereol. 2017;31(11):1907-1916. https://doi.org/10.1111/jdv.14406
16. Eser A, Erpolat S, Kaygusuz I, Balci H, Kosus A. Investigation of Demodex folliculorum frequency in patients with polycystic ovary syndrome. An Bras Dermatol. 2017;92(6):807-810. https://doi.org/10.1590/abd1806-4841.20176043

ORIGINAL ARTICLE

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):157-163
DOI: 10.29058/mjwbs.1793536

Clinical and radiological outcome comparison of unilateral vs. bilateral percutaneous vertebroplasty for osteoporotic vertebral compression fractures

Osteoporotik omurga kompresyon kırıklarında tek taraflı ve çift taraflı perkütan vertebroplasti uygulamalarının klinik ve radyolojik sonuçlarının karşılaştırılması

Levent Aydın[1], Buse Sarıgül[2], Tufan Agah Kartum[3], Gonca Gül Öndüç[2], Saime Ayça Şahin[4], Ali Fatih Ramazanoğlu[2]

[1]Department of Neurosurgery, Medicana Beylikdüzü International Hospital, İstanbul, Türkiye
[2]Department of Neurosurgery, University of Health Science, Ümraniye Training and Research Hospital, İstanbul, Türkiye
[3]Department of Neurosurgery, İstanbul University-Cerrahpaşa,- Cerrahpaşa Faculty of Medicine, İstanbul, Türkiye
[4]Department of Neurosurgery, University of Health Science, Seyrantepe Hamidiye Etfal Training and Research Hospital, İstanbul, Türkiye

Cite this article as: *Aydın L et al. Clinical and radiological outcome comparison of unilateral vs. bilateral percutaneous vertebroplasty for osteoporotic vertebral compression fractures. Med J West Black Sea. 2026;10(1):157-163.*

ABSTRACT

Aim: Vertebral compression fractures are a common cause of pain and disability in the elderly population. Percutaneous vertebroplasty is an established minimally invasive intervention, yet the optimal approach (unilateral or bilateral) remains controversial. This study focused on comparing the clinical and radiological outcomes of unilateral versus bilateral percutaneous vertebroplasty in patients with painful osteoporotic vertebral compression fractures.

Material and Methods: A retrospective review was conducted on 203 patients (73 males, 130 females; mean age 65.3 ± 8.9 years) with 309 treated vertebral levels from 2014 to 2024. Patients underwent either unilateral (n = 104) or bilateral (n = 99) PVP. Pain and disability were assessed using the Visual Analog Scale (VAS) and Oswestry Disability Index (ODI), while radiological outcomes included anterior vertebral height restoration and segmental Cobb angle correction.

Results: Both approaches provided significant pain relief and functional improvement. Bilateral PVP achieved lower postoperative VAS (1.9 vs 2.1, p = 0.040) and ODI scores (23.1% vs 26.3%, p = 0.020) compared to unilateral PVP. Radiologically, bilateral procedures resulted in superior anterior height restoration (25.8% vs 19.2%, p = 0.001) and greater Cobb angle correction (9.1° vs 6.4°, p = 0.001). Cement leakage rates were similar in both groups (≈4%), with no symptomatic complications.

Conclusion: Unilateral and bilateral PVP are both safe and effective for osteoporotic VCFs. Bilateral PVP demonstrated superior radiological correction, whereas unilateral PVP remains advantageous due to reduced operative burden. Treatment decisions should be individualized based on patient anatomy and procedural goals.

Keywords: Bilateral percutaneous vertebroplasty, unilateral percutaneous vertebroplasty, vertebral compression fracture, vertebroplasty

ÖZ

Amaç: Vertebra kompresyon kırıkları (VKK) , yaşlı popülasyonda ağrı ve fonksiyon kaybının en yaygın nedenlerinden biridir. Perkütan vertebroplasti (PVP), minimal invaziv ve kanıtlanmış bir tedavi yöntemidir; ancak optimal yaklaşımın (tek taraflı ve ya çift taraflı) olması gerektiği halen tartışmalıdır. Bu çalışmanın amacı, ağrılı osteoporotik VKK'lerde tek taraflı ve çift taraflı PVP uygulamalarının klinik ve radyolojik sonuçlarını karşılaştırmaktır.

Gereç ve Yöntemler: 2014–2024 yılları arasında 203 hasta (73 erkek, 130 kadın; ort. yaş 65,3 ± 8,9 yıl) ve toplam 309 vertebra düzeyi retrospektif olarak incelendi. Hastalara tek taraflı (n = 104) veya çift taraflı (n = 99) PVP uygulandı. Klinik değerlendirmede VAS ve Oswestry Disabilite İndeksi (ODI) kullanılırken, radyolojik sonuçlarda anterior vertebra yüksekliği restorasyonu ve segmental Cobb açısı düzelmesi değerlendirildi.

Bulgular: Her iki yaklaşım da anlamlı ağrı azalması ve fonksiyonel iyileşme sağladı. Çift taraflı PVP, tek taraflı PVP'ye kıyasla daha düşük postoperatif VAS (1,9 vs 2,1; p = 0,040) ve ODI skorları (%23,1 vs %26,3; p = 0,02) ile sonuçlandı. Radyolojik olarak çift taraflı girişimler, daha yüksek anterior yükseklik restorasyonu (%25,8 vs %19,2; p = 0,001) ve daha fazla Cobb açısı düzelmesi (9,1° vs 6,4°; p = 0,001) sağladı. Sement kaçağı oranları her iki grupta benzerdi ve semptomatik komplikasyon görülmedi.

Sonuç: Hem tek taraflı hem de çift taraflı PVP, osteoporotik VKK tedavisinde güvenli ve etkili yöntemlerdir. Çift taraflı PVP daha üstün radyolojik düzelme sağlarken, tek taraflı PVP daha az cerrahi yük nedeniyle avantajlıdır. Tedavi seçimi, hasta anatomisi ve prosedürel hedefler göz önünde bulundurularak bireyselleştirilmelidir.

Anahtar Kelimeler: Çift taraflı perkütan vertebroplasti, osteoporotik kompresyon fraktürü, tek taraflı perkütan vertebroplasti, vertebroplasti

Corresponding Author: Levent Aydın ✉ mdleventaydin@gmail.com
Received: 30.09.2025 **Revision:** 15.02.2026 **Accepted:** 12.03.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Both unilateral and bilateral percutaneous vertebroplasty provided significant pain reduction and functional recovery in patients with osteoporotic vertebral compression fractures.
- Bilateral vertebroplasty achieved superior vertebral height restoration and greater segmental Cobb angle correction compared with the unilateral technique.
- Unilateral vertebroplasty remains an efficient alternative because of its lower procedural burden, supporting individualized technique selection according to anatomy and treatment goals.

INTRODUCTION

Vertebral compression fractures (VCFs) represent a major cause of pain, disability, and reduced quality of life in elderly populations, particularly among individuals with osteoporosis, trauma history, or metastatic spinal disease (1-3). These fractures may lead to chronic pain, progressive deformity, immobility, and increased morbidity. Conservative treatment often fails to provide adequate symptom relief and prolonged immobilization may further exacerbate comorbidities (4,5). Since its first use by Galibert et al. in 1987 for vertebral hemangioma, percutaneous vertebroplasty (PVP) has become a widely accepted minimally invasive procedure for osteoporotic and neoplastic VCFs (6). The technique involves transpedicular injection of polymethylmethacrylate (PMMA) cement into the collapsed vertebral body to achieve stabilization and pain reduction.

PVP can be performed via unilateral or bilateral transpedicular access; however, the optimal approach remains controversial. Bilateral injection allows for more symmetrical cement distribution and greater potential for kyphotic correction, whereas unilateral injection reduces operative time, radiation exposure, and procedural risks. The objective of this study was to evaluate and compare the clinical and radiological outcomes associated with unilateral and bilateral percutaneous vertebroplasty (PVP) in a clearly characterized group of patients presenting with symptomatic osteoporotic vertebral compression fractures.

MATERIAL and METHODS

Patient Selection and Data Collection

In this retrospective clinical observational study, the medical data of patients who received PVP treatment for osteoporotic VCF at our institution between 2014 and 2024 were analyzed. The inclusion criteria were; 1) severe axial pain that persisted despite four weeks of conservative treatment including bed rest, analgesics, orthoses, 2) Vertebral body height loss less than 50%, 3) absence of spinal canal compromise and neurological deficit. Exclusion criteria comprised individuals with known coagulopathies, active systemic or spinal infections, posterior vertebral wall destruction, severe and uncontrolled cardiopulmonary conditions that would contraindicate the procedure and malignancy. Ethics committee approval was obtained from University of Health Science, Umraniye Training and Research Hospital (Ethical ID: 284733070). Written informed consent form was obtained from all patients prior to intervention.

Radiological Evaluation

Preoperative and postoperative imaging assessments including computed tomography (CT) and magnetic resonance imaging (MRI) were performed for all patients. Measurements of the anterior vertebral body height were obtained from CT images for both the fractured vertebra and the adjacent levels. The rate of anterior height recovery was determined using the following formula: (postoperative anterior height – preoperative anterior height) divided by the mean anterior height of the adjacent upper and lower vertebrae, multiplied by 100.

The segmental Cobb angle was defined by the angle formed between the superior endplate of the vertebra above and the inferior endplate of the vertebra below the affected level. Additionally, standing radiographs were acquired at the 3rd, 6th, and 12th postoperative months for follow-up evaluation.

Clinical Evaluation

Pain intensity was quantified using the Visual Analog Scale (VAS), where 0 indicated no pain and 10 represented the most severe pain imaginable. Functional impairment was assessed using the Oswestry Disability Index (ODI), a validated tool for measuring disability related to spinal conditions. The clinical scores were recorded preoperatively and on the 10th postoperative day.

Surgical Technique

PVP was conducted in either the angiography suite or the operating room using a standard transpedicular approach under fluoroscopic guidance. The procedure was performed under local anesthesia and sedation. All patients were continuously monitored with ECG, pulse oximetry, and blood pressure tracking. An intravenous line was placed for administration of sedatives, analgesics, antibiotics, and anti-

emetics as needed. Moderate sedation was used in select cases to enhance patient comfort.

The patient was positioned prone and draped in a sterile manner. Local anesthetic was administered to the skin and underlying soft tissues. After local anesthesia and sterile preparation, a trocar was carefully advanced through the pedicle into the anterior third of the fractured vertebral body (Figure 1). Polymethylmethacrylate (PMMA) bone cement of medium viscosity was then injected under continuous biplanar fluoroscopy to ensure controlled filling and to minimize the risk of cement leakage. The injection was terminated once adequate vertebral body opacification and stabilization were achieved. After cement injection was completed, the trocar was carefully removed, bleeding was controlled with manual pressure and a sterile dressing was placed over the entry site. Given the minimally invasive technique, no suturing was necessary.

Statistical Analysis

All statistical evaluations were performed using SPSS software (version 25.0; IBM Corp., Armonk, NY, USA). The distribution of variables was tested for normality with the Shapiro-Wilk test. Intergroup comparisons between unilateral and bilateral PVP were carried out using either the independent samples t-test or the Mann Whitney U test, depending on distribution. For intragroup analyses, preoperative and postoperative differences in clinical (VAS, ODI) and radiological outcomes were assessed using paired t tests or Wilcoxon signed rank tests, as appropriate. Categorical data were analyzed using the Pearson chi square test or Fisher's exact test. P value below 0.05 was accepted as indicative of statistical significance.

RESULTS

A total of 203 patients were included in the study, comprising 130 females (64.0%) and 73 males (36.0%), with a mean age of 65.3±8.9 years (range: 45-82 years). A total of 309 vertebral levels were treated, with the most frequently affected site being the L1 vertebra, accounting for 31.8% of all interventions. The mean follow-up was 9 months (range:6-48) (Table 1). Among the cohort, 104 patients underwent unilateral PVP, while 99 patients received the pro-

Table 1: Baseline Demographic and Clinical Characteristics of Patients

Variable	Unilateral PVP (n = 104)	Bilateral PVP (n = 99)	Total (n = 203)	p value
Age (years), mean ± SD	65.1 ± 8.7	65.6 ± 9.1	65.3 ± 8.9	0.720 [a]
Sex (female), n (%)	67 (64.4)	63 (63.6)	130 (64.0)	0.910 [b]
Number of treated levels, mean ± SD	1.46 ± 0.7	1.52 ± 0.6	1.49 ± 0.7	0.530 [c]
Most frequent level (L1), n (%)	31 (29.8)	28 (28.3)	59 (29.1)	0.811 [b]

[a] Independent-samples t-test. [b] Pearson's chi-square test. [c] Mann–Whitney U test

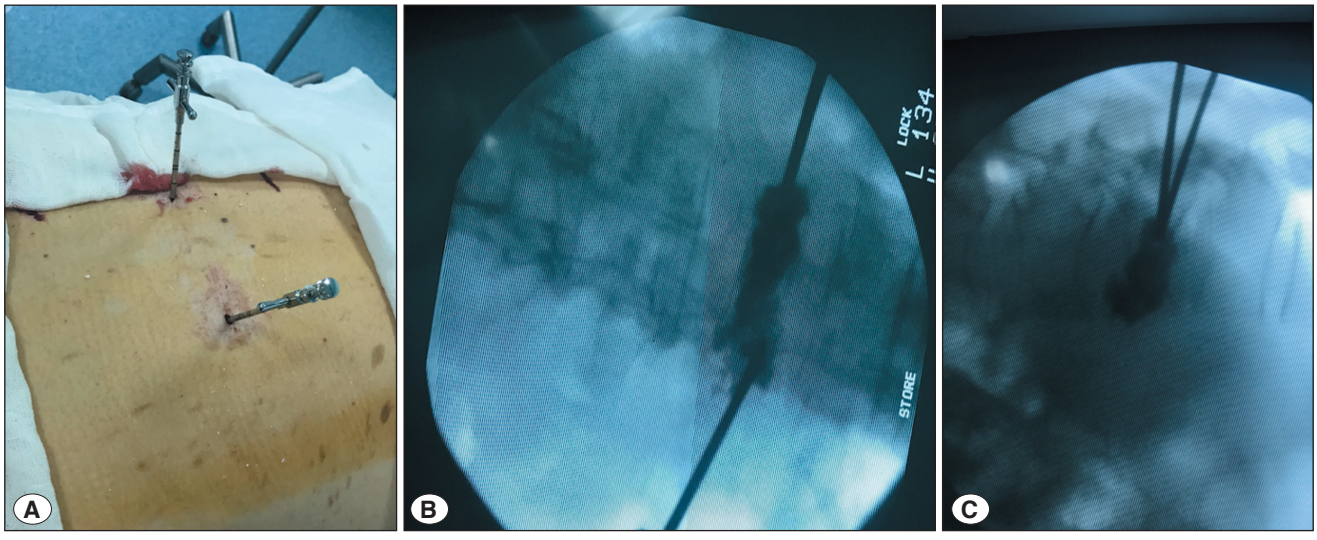


Figure 1: Bilateral percutaneous vertebroplasty for a symptomatic T11 compression fracture. **A)** Bilateral transpedicular cannulas positioned symmetrically for vertebral body access. **B)** Intraoperative anteroposterior fluoroscopic image showing stepwise PMMA injection under intermittent fluoroscopic control to ensure gradual filling and prevent cement migration. **C)** Lateral fluoroscopic view demonstrating homogeneous cement distribution within the vertebral body; serial imaging minimized the risk of leakage into venous structures, the spinal canal, or paravertebral tissues.

cedure bilaterally. The two groups did not differ significantly in terms of patient demographics or pre-treatment clinical status, including pain and disability indices (p=0.070, independent samples t-test). Both treatment modalities led to statistically significant reductions in pain and disability scores following the procedure (Figure 2).

Clinical Assessment

The unilateral group demonstrated a decline in mean VAS scores from 8.2±1.1 preoperatively to 2.1±0.9 postoperatively (p= 0.001, Wilcoxon Signed Rank test), while the bilateral group showed a reduction from 8.3±1.0 to 1.9±0.8 (p = 0.001, Wilcoxon Signed Rank test). The difference in postoperative pain scores was statistically significant in favor of bilateral intervention (p=0.04, Mann Whitney U test) (Table 2).

Similarly, the Oswestry Disability Index (ODI) improved from 68.5±11.3% to 26.3±9.2% in the unilateral group (p=0.001, Wilcoxon Signed Rank test) and from 69.2±10.8% to 23.1±8.4% in the bilateral group (p=0.001, Wilcoxon Signed Rank test), with a significant between-group difference postoperatively (p=0.020, Mann-Whitney U test). Pain relief, defined as a reduction of ≥4 points in VAS, was achieved in 89.4% of patients in the unilateral group and 91.2% in the

Table 2: Comparison of the treatment modalities and outcomes

	Unilateral (n = 104)	Bilateral (n = 99)	p-value
Pre-treatment clinical status			
Mean pre-op VAS score	8.2±1.1	8.3±1.0	0.700[c]
Mean pre-op ODI (%)	68.5±11.3	69.2±10.8	0.680[c]
Post-treatment outcomes			
Mean post-op VAS score	2.1±0.9	1.9±0.8	0.040[c]
Mean post-op ODI (%)	26.3±9.2	23.1±8.4	0.020[c]
Patients achieving ≥4-point VAS reduction, n (%)	93 (89.4)	90 (91.2)	0.723[b]
Radiographic outcomes			
Anterior vertebral height restoration (%)	19.2±4.7	25.8±5.1	0.001[a]
Cobb angle improvement (°)	6.4±2.3	9.1±2.6	0.001[a]
Complications			
Cement leakage, n (%)	4 (3.8)	4 (4.0)	0.943[b]
Symptomatic complications	None	None	

[a] Independent-samples t-test. [b] Pearson's chi-square test. [c] Mann–Whitney U test.

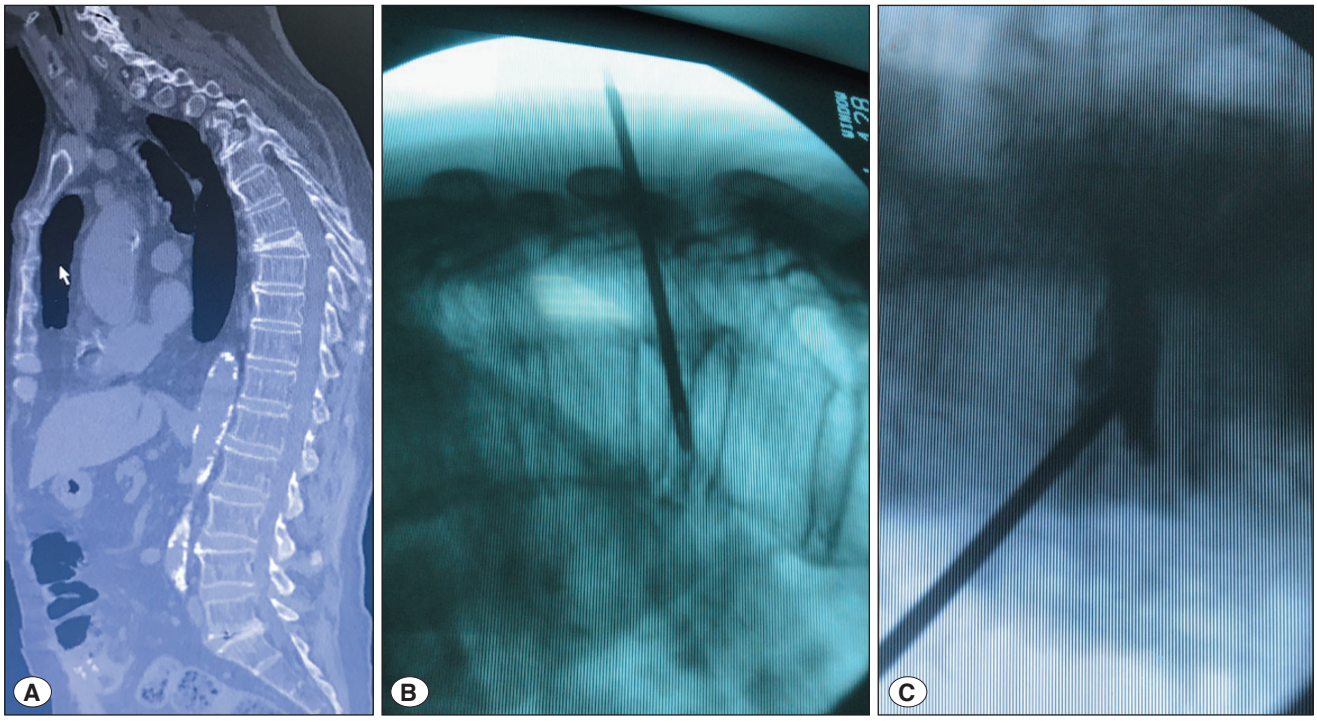


Figure 2: A) Lateral CT image showing a severe T8 compression (vertebra plana) with complete height loss and structural collapse, indicating instability and risk of kyphotic progression. **B)** Intraoperative lateral fluoroscopic view during unilateral percutaneous vertebroplasty. The transpedicular cannula is advanced stepwise under continuous imaging for safe positioning and controlled cement delivery. **C)** Post-procedural anteroposterior fluoroscopic image demonstrating satisfactory vertebral augmentation with partial height restoration and uniform PMMA distribution. Continuous imaging minimized the risk of cement leakage into adjacent structures.

bilateral group, with no statistically difference (p=0.66, chi square test).

Radiological Assessment

Similarly, local Cobb angle correction was greater in the bilateral group (9.1 ± 2.6° pre- vs. postoperative change, p = 0.001, Wilcoxon Signed Rank test) than in the unilateral group (6.4 ± 2.3°, p = 0.001, Wilcoxon Signed Rank test), with a significant intergroup difference (p = 0.001, Mann-Whitney U test).

Cement leakage was observed in eight cases, four from each group, corresponding to an overall incidence of 3.9%. Leakage rates were nearly identical between unilateral (3.8%) and bilateral (4.0%) procedures, with no statistically significant difference (p=0.93, Chi-square test). Importantly, none of the cases led to symptomatic or clinically significant complications.

DISCUSSION

Numerous studies have demonstrated that PVP is a reliable and minimally invasive intervention for patients with osteoporotic vertebral fractures who fail to achieve adequate symptom relief with conservative measures such as bed rest, immobilization, pharmacological therapy or orthotic bracing (7-9). In our study, we evaluated the radiological and clinical outcomes of patients who underwent either unilateral or bilateral PVP for osteoporotic VCF.

In comparing unilateral and bilateral approaches to PVP, it is essential to highlight the distinctive features of each technique. Bilateral PVP allows for more symmetrical cement distribution, providing biomechanically balanced load transmission across the vertebral body and potentially reducing the incidence of intravertebral microfractures. In contrast, unilateral PVP has been reported to achieve comparable radiological and clinical outcomes, while offering advantages such as shorter procedure time, lower radiation exposure, and greater cost-effectiveness (10).

One of the primary concerns in both unilateral and bilateral PVP is the risk of cement leakage. A meta analysis identified intervertebral clefts, cortical disruption, low cement viscosity and excessive cement volume as key contributors to leakage risk (11). Baroud and Bohner reported that cement viscosity is a critical factor with a transition from low to medium viscosity reducing leakage rates to below 10% during the procedure (12). Technological refinements such as remote cement delivery systems have also been introduced to improve injection control and minimize radiation exposure for operators (13). Dai et al., through a finite element analysis model, found that bilateral transpedicular vertebroplasty biomechanically reduces stress concentrations within the vertebral body and intervertebral discs, offering superior stability compared to the unilateral technique (14). Furthermore, in vitro studies suggest that a cement volume of approximately 3 cm^3 is sufficient for effective vertebral restoration, with bipedicular injection providing more uniform distribution (15).

From a radiological perspective, especially in multilevel osteoporotic vertebral compression fractures, alterations in sagittal and coronal spinal balance must be considered. A study found that patients undergoing bilateral PVP exhibited more significant improvements in the coronal Cobb angle compared to those treated unilaterally (16). However, another investigation employing voxel-based morphometry suggested that unilateral vertebroplasty, when performed with symmetrical cement distribution, may offer comparable or superior outcomes (17). Consistent with prior reports, our results showed that the bilateral approach yielded significantly superior anterior vertebral height restoration and greater improvement in Cobb angle compared to the unilateral technique, suggesting that bipedicular access may facilitate more effective structural realignment.

In terms of clinical outcomes, most studies utilizing Visual Analog Scale (VAS) and Oswestry Disability Index (ODI) scores have shown no statistically significant differences in long-term outcomes between unilateral and bilateral techniques (3,18-21). Our results are consistent with this evidence, showing that both approaches achieved significant and durable pain relief, with bilateral intervention providing slightly superior radiological correction. The advantages of unilateral PVP, namely reduced operative time, radiation exposure and cost, remain compelling and have been highlighted in several reports (19,22-28).

This study has several limitations. First, its retrospective and single-center design may introduce selection bias and limit the generalizability of the findings. Second, patients were not randomized to unilateral or bilateral procedures; thus, treatment decision may have been influenced by surgeon preference or anatomical considerations, potentially confounding the results. Third, although validated instruments such as VAS and ODI were used to assess pain and disability, these outcomes remain subjective and may be influenced by patient expectations. Another limitation is related to the assessment of segmental Cobb angle. In our study, measurements were obtained from pre- and postoperative CT scans, whereas the gold standard for evaluating sagittal and coronal alignment is standing whole-spine radiography. Therefore, our analysis reflects only the local angular correction at the treated level, and may not fully capture global spinal balance.

Conclusion

Although both unilateral and bilateral PVP are effective and safe interventions for managing osteoporotic VCF, current evidence suggests that neither technique offers a definitive

superiority in terms of long-term clinical outcomes. While bilateral PVP provides more symmetrical cement distribution and may offer certain biomechanical advantages, unilateral PVP remains a favorable alternative due to its shorter procedural duration, reduced radiation exposure, and cost-effectiveness. Ultimately, the choice between unilateral and bilateral approaches should be individualized, taking into account patient anatomy, fracture characteristics and procedural goals, while prioritizing both biomechanical integrity and clinical efficiency.

Author Contributions

Study conception and design: **Levent Aydın, Buse Sarıgül,** Data collection: **Levent Aydın, Buse Sarıgül, Gonca Gül Onduç,** Analysis and interpretation of results: **Levent Aydın, Saime Ayça Şahin,** Draft manuscript preparation: **Levent Aydın, Tufan Agah Kartum, Buse Sarıgül,** Critical revision of the article: **Ali Fatih Ramazanoglu, Buse Sarıgül,** Other (study supervision, fundings, materials, etc...): **Saime Ayça Şahin, Ali Fatih Ramazanoğlu.**

Conflicts of Interest

We declare that there are no conflicts of interest related to this study and no financial funding was received that could have affected the results. All authors have reviewed and approved the final version of the manuscript and the sequence of authorship has been agreed upon by all contributors.

Ethical Approval

Health Science University, Umraniye Training and Research Hospital (Ethical ID: 284733070).

REFERENCES

1. Ateş A, Gemalmaz HC, Deveci MA, Şimşek SA, Çetin E, Şenköylü A. Comparison of effectiveness of kyphoplasty and vertebroplasty in patients with osteoporotic vertebra fractures. Acta Orthop Traumatol Turc 2016;50(6):619-622. http://doi.org/10.1016/j.aott.2016.10.001
2. Qiao Y, Wang X, Liu Y, Hu J, Yuan FH, Zhao ZG. Comparison of unilateral and bilateral percutaneous kyphoplasty for osteoporotic vertebral compression fractures. J Pain Res 2023;16:1813-1823. http://doi.org/10.2147/JPR.S413048
3. Chen JF, Lee ST, Lui TN, Wu CT, Liao CC. Percutaneous vertebroplasty for the treatment of osteoporotic vertebral compression fractures: A preliminary report. Chang Gung Med J 2002;25(5):306-314
4. McKiernan F, Faciszewski T, Jensen R. Quality of life following vertebroplasty. J Bone Joint Surg Am 2004;86(12):2600-2606. http://doi.org/10.2106/00004623-200412000-00007
5. Zhang LG, Gu X, Zhang HL, Zhang QG, Cai XB, Tao K. Unilateral or Bilateral Percutaneous Vertebroplasty for Acute Osteoporotic Vertebral Fracture: A Prospective Study. J Spinal Disord Tech 2015;28(2):E85-E88. http://doi.org/10.1097/BSD.0000000000000064
6. Galibert P, Deramnd H, Rosat P, Le Gars D. Preliminary note on the treatment of vertebral angioma by percutaneous acrylic vertebroplasty. Neurochirurgie 1987;33:166-168
7. Kobayashi K, Shimoyama K, Nakamura K, Murata K. Percutaneous vertebroplasty immediately relieves pain of osteoporotic vertebral compression fractures and prevents prolonged immobilization of patients. Eur Radiol 2005;15(2):360-367. http://doi.org/10.1007/s00330-004-2536-8
8. Purkayastha S, Gupta AK, Kapilamoorthy TR, Kesavadas C, Thomas B, Krishnamoorthy T, et al. Percutaneous vertebroplasty in the management of vertebral lesions. Neurol India 2005;53(2):167-173
9. Hee HT. Percutaneous vertebroplasty: Current concepts and local experience. Neurol India 2005;53(4):475-482
10. Zhao X, Han X, Li C, Han B, Zhu T, Shi Y. Comparative study between unilateral and bilateral extrapedicular approaches for osteoporotic vertebral compression fractures: A retrospective study. Sci Prog 2025;108(1). http://doi.org/10.1177/00368504241302468
11. Zhan Y, Jiang J, Liao H, Tan H, Yang K. Risk Factors for Cement Leakage After Vertebroplasty or Kyphoplasty: A Meta-Analysis of Published Evidence. World Neurosurg 2017;101:633-642. http://doi.org/10.1016/j.wneu.2017.02.038
12. Baroud G, Bohner M. Biomechanical impact of vertebroplasty: Postoperative biomechanics of vertebroplasty. Joint Bone Spine 2006;73(2):144-150. http://doi.org/10.1016/j.jbspin.2005.07.003
13. Jin P, Liu X, Li M, Sun G. Clinical Experience Using a Remote Control Injection System in Vertebroplasty: Feasibility, Safety, and Cement Leakage of Osteoporotic and Malignant Compression Fractures. Clin Spine Surg 2017;30(3):E305-E309. http://doi.org/10.1097/BSD.0000000000000284
14. Dai H, Liu Y, Han Q, Zhang A, Chen H, Qu Y, et al. Biomechanical comparison between unilateral and bilateral percutaneous vertebroplasty for osteoporotic vertebral compression fractures: A finite element analysis. Front Bioeng Biotechnol 2022;10. http://doi.org/10.3389/fbioe.2022.950737
15. Kim JM, Shin DA, Byun DH, Kim HS, Kim S, Kim HI. Effect of bone cement volume and stiffness on occurrences of adjacent vertebral fractures after vertebroplasty. J Korean Neurosurg Soc 2012;52(5):435-440. http://doi.org/10.3340/jkns.2012.52.5.435
16. Tan Y, Liu J, Li X, Fang L, He D, Tan J, et al. Multilevel unilateral versus bilateral pedicular percutaneous vertebroplasty for osteoporotic vertebral compression fractures. Front Surg 2023;9. http://doi.org/10.3389/fsurg.2022.1049980
17. Lee HJ, Kim T, Park J, Cho J, Yi JS. Quantitative Comparison of Vertebral Structural Changes After Percutaneous Vertebroplasty Between Unilateral Extrapedicular Approach and Bilateral Transpedicular Approach Using Voxel-Based Morphometry. Neurospine 2023;20(4):1287-1302. http://doi.org/10.14245/ns.2346536.268
18. Fu X, Li Y, Tian P, Xu G jun, Li Z jun. Comparison between unilateral curved and bilateral straight percutaneous vertebral augmentation in the treatment of osteoporotic vertebral compression fractures: A meta-analysis. Jt Dis Relat Surg 2023;34(2):237-244. http://doi.org/10.52312/jdrs.2023.1034
19. Sun H, Li C. Comparison of unilateral and bilateral percutaneous vertebroplasty for osteoporotic vertebral compression fractures: A systematic review and meta-analysis. J Orthop Surg Res 2016;11(1). http://doi.org/10.1186/s13018-016-0424-7

20. Meral M, Orunoğlu M, Biçer E, Berkyürek E, Koç RK. Osteoporotic vertebral fracture; comparative analysis of unilateral and bilateral vertebroplasty results. J Turk Spinal Surg 2024;35(1):38-42
21. Dai C, Wang S, Zheng L, Wang K, Yu H, Chen Y. Unilateral and Bilateral Percutaneous Vertebroplasty for Thoracolumbar Osteoporotic Vertebral Compression Fracture with Distant Lumbosacral Pain: A Single-Center Retrospective Analysis. World Neurosurg 2025;195. http://doi.org/10.1016/j.wneu.2024.121457
22. Chen Y, Zhang H, Chen H, Ou Z, Fu Y, Zhang J. Comparison of the effectiveness and safety of unilateral and bilateral percutaneous vertebroplasty for osteoporotic vertebral compression fractures: A protocol for systematic review and meta-analysis. Medicine (Baltimore) 2021;100:e28453. http://doi.org/10.1097/MD.0000000000028453
23. Takura T, Yoshimatsu M, Sugimori H, Takizawa K, Furumatsu Y, Ikeda H. Cost-Effectiveness Analysis of Percutaneous Vertebroplasty for Osteoporotic Compression Fractures. Clin Spine Surg 2017;30(3):E205-E210. http://doi.org/10.1097/BSD.0000000000000302
24. Kallmes DF, Jensen ME. Percutaneous vertebroplasty. Radiology 2003;229:27-36. http://doi.org/10.1148/radiol.2291020722
25. Zhong R, Liu J, Wang R, Liu Y, Chen B, Jiang W, et al. Unilateral curved versus bipedicular vertebroplasty in the treatment of osteoporotic vertebral compression fractures. BMC Surg 2019;19(1). http://doi.org/10.1186/s12893-019-0647-0
26. Shi D, Li X, Huang F, Wei G, Lin Z. Unilateral percutaneous vertebroplasty in osteoporotic vertebral compression fractures: A clinical efficacy evaluation. Exp Ther Med 2024;27(4). http://doi.org/10.3892/etm.2024.12345
27. Liu Z, Zhang H, Ma X, Zhang W, Shi B. Comparison of the efficacy of unilateral versus bilateral pedicle approach vertebroplasty in the treatment of osteoporotic thoracic and lumbar vertebral compression fractures: A retrospective cohort study. Medicine (Baltimore) 2025;104(31):e43632. http://doi.org/10.1097/MD.0000000000043632
28. Kim T, Park J, Cho J, Yi JS, Lee HJ. Quantitative Comparison of Vertebral Structural Changes After Percutaneous Vertebroplasty Between Unilateral Extrapedicular Approach and Bilateral Transpedicular Approach Using Voxel-Based Morphometry. Neurospine 2023;20(4):1287-1302. http://doi.org/10.14245/ns.2346536.268

CASE REPORT

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):164-167
DOI: 10.29058/mjwbs.1572290

Hemichorea-hemiballismus developing due to non-ketotic hyperglycemia: A rare case report

Non-ketotik hiperglisemiye bağlı gelişen hemikore-hemiballismus: Nadir bir olgu sunumu

Merve İrem Atıcı[1], Baycan Kuş[1], Taner Şahin[1], İrem Kömürcü[2], Merva Tuna[2]

[1]Emergency Medicine Clinic, University of Health Sciences, Kayseri City Hospital, Kayseri, Türkiye
[2]Neurology Clinic, University of Health Sciences, Kayseri City Hospital, Kayseri, Türkiye

Cite this article as: *Atıcı Mİ et al. Hemichorea-hemiballismus developing due to non-ketotic hyperglycemia: a rare case report. Med J West Black Sea. 2026;10(1):164-167.*

ABSTRACT

Hemichorea-hemiballismus can be seen as a rare complication of non-ketotic hyperglycemia in patients with type 2 diabetes mellitus. We detected hemichorea-hemiballismus secondary to hyperglycemia in a patient who presented to the emergency department with speech disorder and involuntary movements in the left arm, as a result of our investigations. We decided to present this case because it is rare in the literature.

Keywords: Athetosis, hemiballismus, hemichorea, hyperglycemia

ÖZ

Hemikore-hemiballismus , Tip 2 diyabetes mellitus hastalarında non-ketotik hipergliseminin nadir ortaya çıkan bir komplikasyonu olarak görülebilir. Acil servise konuşma bozukluğu ve sol kolda istemsiz hareketler nedeniyle başvuran bir hastada yaptığımız tetkikler sonucunda hiperglisemiye sekonder gelişen hemikore-hemiballismus tespit ettik ve literatürde nadir rastlandığı için bu vakayı sunduk.

Anahtar Kelimeler: Ateatoz, hemiballismus, hemikore, hiperglisemi

Highlights

- Hemichorea-hemiballismus can manifest as a rare neurological complication of non-ketotic hyperglycemia, particularly in patients with Type 2 Diabetes Mellitus.
- The case reported describes an 81-year-old female presenting to the emergency department with a speech disorder and involuntary left-sided movements, with a blood glucose level of 726 mg/dl, leading to a diagnosis of hemichorea-hemiballismus secondary to hyperglycemia.
- Contrary to common findings, the patient's Cranial Computed Tomography (CT) and Diffusion Magnetic Resonance Imaging (DWI) did not reveal any acute or pathological findings.
- The management involved immediate blood glucose regulation, complemented by an initial valproic acid (1200 mg) infusion in the emergency department, followed by combination therapy with haloperidol during hospitalization.
- The patient's symptoms completely resolved by the third day of hospitalization following successful blood sugar regulation and combination pharmacological treatment.
- This case underscores the importance for emergency physicians to consider non-ketotic hyperglycemia as the etiology for gait disturbance and involuntary movements, even when initial cranial imaging is unremarkable.

Corresponding Author: Merve İrem Atıcı ✉ miremcimen@gmail.com **Received:** 27.02.2025 **Revision:** 22.07.2025 **Accepted:** 03.09.2025

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

INTRODUCTION

Chorea; refers to involuntary movements that are irregular, purposeless, brief, sudden, and rapid in nature, which can spread from one part of the body to another, such as the extremities, trunk, neck, and face. Ballismus, on the other hand, refers to higher amplitude, coarser involuntary movements than chorea, involving the proximal joints like the shoulder and hip, causing throwing or violent flailing movements. It is often unilateral and therefore called hemiballismus. Hemichorea can be seen in patients with diabetic nonketotic hyperglycemia. Symptoms usually resolve after blood sugar regulation, but in some cases, long-term medical treatments may be required (1-3). Although T1-weighted cranial magnetic resonance imaging (MRI) in most cases shows hyperintensity in the contralateral putamen and basal ganglia, no pathology was found in the brain computed tomography (CT), brain diffusion, and non-contrast MRI performed in our presented case. In this case, we examined a patient diagnosed with hemichorea-hemiballismus, one of the rare complications of Type 2 diabetes mellitus.

CASE

An 81-year-old female patient was brought to the emergency department due to speech disorder that started in the last 3 days and involuntary movements developing in the left upper and lower extremities. The patient's history included atrial fibrillation, diabetes mellitus type 2, hypertension, and a history of cerebrovascular disease. It was learned that she was irregularly using her medications (metformin, empagliflozin, candesartan, methimazole, insulin aspart, edoxaban). Physical examination revealed the patient was conscious, oriented, and cooperative, with no facial asymmetry, lateralizing motor deficit, or pathological reflexes. Cranial nerve examination was normal. Licking-like movements of the mouth and lips and choreiform movements in the left upper extremity were present. Laboratory tests showed blood sugar 726 mg/dl, serum sodium (Na^+) 129 mmol/l, serum osmolality 317 mOsm/l, venous blood gas pH: 7.327, urine ketone negative, glucose ++++. There were no pathological findings in the patient's other laboratory tests. No urgent, acute pathology was detected on the patient's cranial CT and brain diffusion MRI. (Figure 1, 2) Blood sugar monitoring was performed while the patient was in the emergency department,

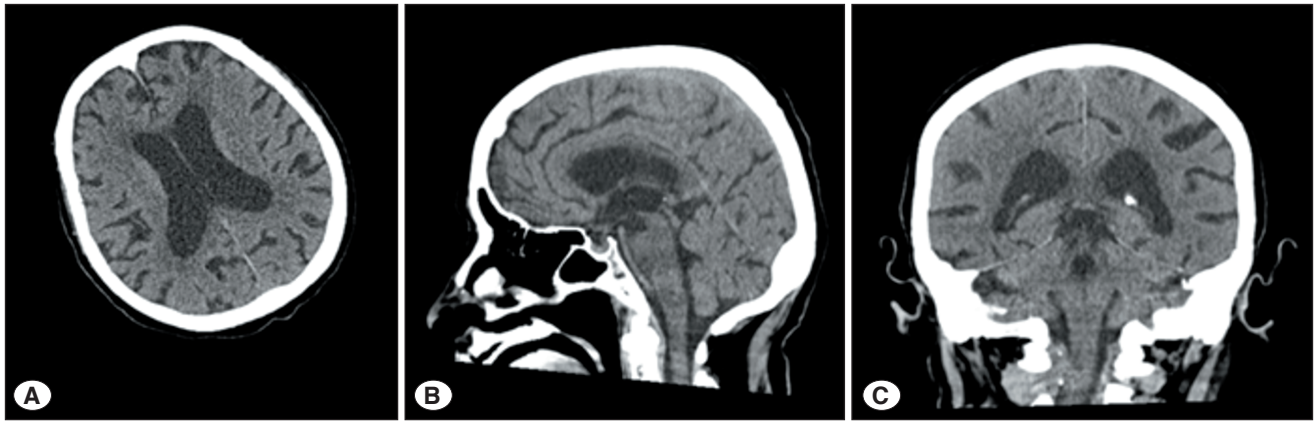


Figure 1: A) Axial view of patient's cranial CT, **B)** A Sagittal view of patient's cranial CT, **C)** Coronal view of patient's cranial CT.

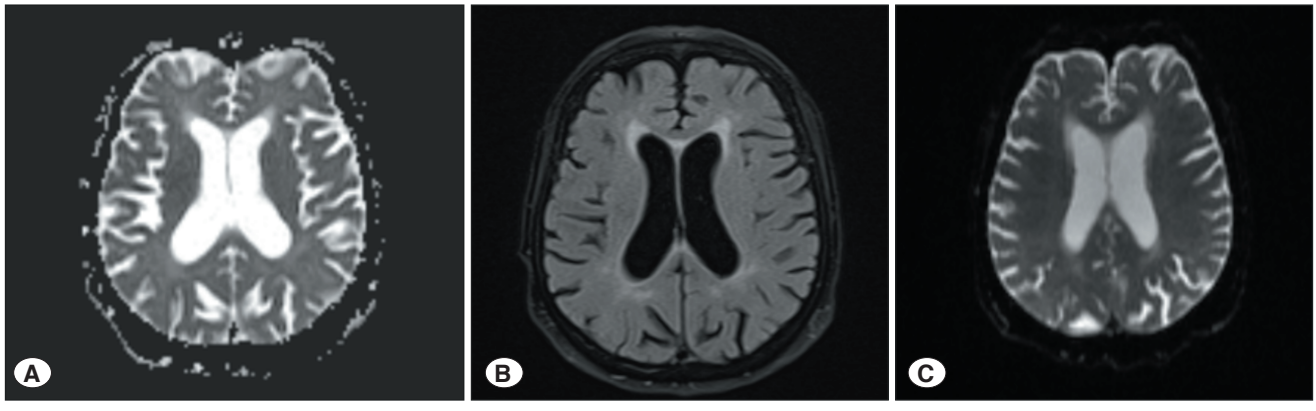


Figure 2: A) Axial DWI of patient's difussion MRI, **B)** T2 Flair of patient's difussion MRI , **C)** T1 images of the patient's difussion MRI

and upon the recommendation of the neurology consultant, valproic acid 1200 mg was administered as a 40-minute infusion. The patient was admitted to the neurology service with a diagnosis of hemichorea-hemiballismus secondary to non-ketotic hyperglycemia. It was learned that the patient's symptoms regressed on the third day of hospitalization in the neurology service with blood sugar regulation and haloperidol treatment, and she was discharged with recovery.

DISCUSSION

Chorea-ballismus can be seen in patients with diabetic non-ketotic hyperglycemia. Hemichorea can occur in elderly female patients, especially those with unregulated blood sugar and irregular medication use. Rarely, cases with bilateral chorea have also been reported. It has been reported that cranial MRI and CT imaging are normal in these patients, and abnormal findings regress with the correction of hyperglycemia (4). In our case, our patient also had Type 2 diabetes mellitus and was not using her medications regularly, and there were no pathological findings on the cranial imaging performed. Her complaints completely resolved after blood sugar regulation. In some cases, treatment may take longer and can even result in death due to increased hyperviscosity.

The pathophysiology of hyperglycemia-associated hemichorea-hemiballismus has not been fully elucidated. In their study of 18 patients, Chang et al., based on imaging findings, argued that hyperglycemia causes transient ischemia in patients and that this is the cause of the clinical presentation (5). According to the hypothesis developed by Shan et al. in their studies, dysfunction of GABAergic neurons in the caudate nucleus and putamen caused by hyperglycemia or cerebral ischemia leads to non-ketotic hyperglycemia-associated hemichorea-hemiballismus (6). Typical neuroleptics (haloperidol, pimozide, perphenazine, fluphenazine), atypical neuroleptics (olanzapine, quetiapine), sodium valproate, and levetiracetam are used in treatment (4). Valproic acid is chemically similar to GABA, and its mechanism of action is explained by increasing GABA levels as well as blocking voltage-dependent sodium and potassium channels in the brain (7). Therefore, in our presented case, valproic acid infusion was administered to our patient in the emergency department. Clinical improvement was achieved with combination therapy of haloperidol, a D2 receptor agonist, and valproic acid.

Although the pathophysiology is not clear, studies conducted so far have shown that symptoms resolve in most patients after blood sugar regulation. In our case, the findings also regressed and completely resolved on the third day of treatment following blood sugar regulation and combination therapy with haloperidol and valproic acid. However, it has been reported that hemiballismus persists for more than three months in 20% of patients, and medical treatment may be required for a longer duration (2).

Conclusion

In cases of non-ketotic hyperglycemia-associated hemichorea-hemiballismus reported in the literature, no specific findings were detected on cranial CT, but hyperintense foci may be present on T1-weighted MRI. In our case, cranial CT and T1-weighted MRI were normal; additionally, no pathology was detected on electroencephalography examination. Our patient's current clinical picture resolved with blood sugar regulation. This led to the evaluation of our case as non-ketotic hyperglycemia-associated hemichorea-hemiballismus. In patients presenting to the emergency department with gait disturbance and involuntary movements, even if cranial imaging is normal, it should not be forgotten that high and unregulated blood sugar can cause this clinical picture. Emergency physicians should keep in mind that non-ketotic hyperglycemia-associated hemichorea-hemiballismus can be a neurological complication of diabetes.

Author Contributions

Study design and conceptualization: **Taner Şahin;** data collection: **Merve İrem Atıcı and İrem Kömürcü;** analysis and interpretation of results: **Merve İrem Atıcı, Baycan Kuş, and Merva Tuna;** drafting the manuscript: **Merve İrem Atıcı.** The authors reviewed the results and approved the final version of the article.

Ethical Approval

This study is not an experimental or clinical research. As it is a case report, ethical approval was not required. Written informed consent was obtained from the patient for the publication of the case report.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Funding

The authors declare that the study received no funding.

Acknowledgements

We would like to thank Kayseri City Hospital, the Emergency Medicine Clinic, and our esteemed professors for providing us with the opportunity to work on this case.

REFERENCES

1. Kocaoğlu S. Hemichorea-hemiballismus due to non-ketotic hyperglycemia. Balikesir Med J. 2021;5(2):117-118. https://doi.org/10.33716/bmedj.858718
2. Kocasoy Orhan E, Atmaca MM, Atmaca M, Hanağasi HA. Chorea-ballismus associated with hyperglycemia. Noro Psikiyatr Ars. 2013;50(4):375-378. https://doi.org/10.4274/npa.y6468
3. Alagöz A, Koksal S, Acar B, Acar T, Bölük A. Hemiballismus that develops upon thalamic hemorrhage: a case report. Turk J Cerebrovasc Dis. 2016;22(1):21-24. https://doi.org/10.5505/tbdhd.2016.77598
4. Selçuk F, Sorgun M, Akpınar S. Hemiballismus and hemichorea due to cerebral vascular malformation triggered by hyperglycemia. Turk J Neurol. 2016;22:80-83. https://doi.org/10.4274/tnd.39205

5. Chang KH, Tsou JC, Chen ST, Ro LS, Lyu RK, Chang HS, et al. Temporal features of magnetic resonance imaging and spectroscopy in non-ketotic hyperglycemic chorea-ballism patients. Eur J Neurol. 2010;17:589-593. https://doi.org/10.1111/j.1468-1331.2009.02867.x
6. Shan DE, Ho DM, Chang C, Pan HC, Teng MM. Hemichorea-hemiballism: an explanation for MR signal changes. AJNR Am J Neuroradiol. 1998;19:863-870.
7. Özbek O, Işıldak Ö. The use of valproic acid in the treatment of epilepsy and determination of biological matrices. Hacettepe Univ J Fac Pharm. 2022;42(3):199-207. https://doi.org/10.52794/hujpharm.1062609

REVIEW

Medical Journal of Western Black Sea
Batı Karadeniz Tıp Dergisi

Med J West Black Sea 2026;10(1):168-178
DOI: 10.29058/mjwbs.1881313

Artificial intelligence in pain: A comprehensive review

Ağrıda yapay zekâ: Kapsamlı bir derleme

İlhan Celil Özbek[1], Erkan Özduran[2], Volkan Hancı[3]

[1]Department of Physical Medicine and Rehabilitation, Derince Training Research Hospital, Health Sciences University, Kocaeli, Türkiye
[2]Department of Physical Medicine and Rehabilitation, Division of Pain Medicine, Sivas Numune Hospital, Sivas, Türkiye
[3]Department of Anaesthesiology and Reanimation, Dokuz Eylül University, School of Medicine, İzmir, Türkiye

Cite this article as: *Özbek İC et al. Artificial intelligence in pain: A comprehensive review. Med J West Black Sea. 2026;10(1):168-178.*

ABSTRACT

Pain is a major public health problem worldwide due to its high prevalence and substantial negative impact on quality of life. Artificial intelligence (AI)-based chatbots are increasingly used to access health-related information; however, evidence regarding the readability, quality, reliability, and alignment of pain-related information provided by these systems with clinical practice guidelines remains limited and heterogeneous. This review aimed to comprehensively examine the existing literature evaluating responses generated by AI-based chatbots in the field of pain with respect to readability, information quality, reliability, and adherence to clinical practice guidelines. Studies published between 2024 and 2025 that assessed AI chatbot responses to pain-related questions were analyzed. ChatGPT, Gemini, Perplexity, DeepSeek, and other large language models were evaluated. Readability was assessed using the Flesch Reading Ease, Flesch-Kincaid Grade Level, and SMOG indices, while information quality and reliability were evaluated using DISCERN, the JAMA Benchmark Criteria, EQIP, and the Global Quality Score. Adherence to clinical guidelines was examined through comparisons with relevant national and international recommendations. The reviewed studies demonstrated that AI models generally provide accurate basic information about pain. However, most responses exceeded the recommended readability levels for patient education materials. Information quality and reliability were typically rated as moderate, with reported deficiencies in the discussion of treatment risks, alternative options, and source transparency. Although adherence to clinical guidelines was acceptable at the level of general principles, inconsistencies were identified in diagnostic details and treatment sequencing. While AI-based chatbots show potential as supportive tools for pain-related information, their current use as independent sources for clinical decision-making or primary patient education is limited. Human oversight and the generation of health literacy-appropriate content are essential for the safe use of these systems.

Keywords: Pain, artificial intelligence, chatbots, health literacy, readability, quality assessment

ÖZ

Ağrı, yüksek yaygınlığı ve yaşam kalitesi üzerindeki önemli olumsuz etkisi nedeniyle dünya çapında büyük bir halk sağlığı sorunudur. Yapay zekâ (YZ) tabanlı sohbet robotları, sağlıkla ilgili bilgilere erişmek için giderek daha fazla kullanılmaktadır; ancak, bu sistemler tarafından sağlanan ağrı ile ilgili bilgilerin okunabilirliği, kalitesi, güvenilirliği ve klinik uygulama kılavuzlarıyla uyumu hakkındaki kanıtlar sınırlı ve heterojendir. Bu derleme, ağrı alanında YZ tabanlı sohbet robotları tarafından üretilen yanıtları okunabilirlik, bilgi kalitesi, güvenilirlik ve klinik uygulama kılavuzlarına uyum açısından değerlendiren mevcut literatürü kapsamlı bir şekilde incelemeyi amaçlamıştır. 2024 ve 2025 yılları arasında yayınlanan ve YZ sohbet robotlarının ağrı ile ilgili sorulara verdiği yanıtları değerlendiren çalışmalar analiz edildi. ChatGPT, Gemini, Perplexity, DeepSeek ve diğer büyük dil modelleri değerlendirildi. Okunabilirlik, Flesch Okuma Kolaylığı, Flesch-Kincaid Sınıf Düzeyi ve SMOG endeksleri kullanılarak değerlendirilirken, bilgi kalitesi ve güvenilirliği DISCERN, JAMA Kriterleri, EQIP ve Küresel Kalite Puanı kullanılarak değerlendirildi. Klinik kılavuzlara uyum, ilgili ulusal ve uluslararası önerilerle karşılaştırmalar yoluyla incelendi. İncelenen çalışmalar, yapay zekâ modellerinin genel olarak ağrı hakkında doğru temel bilgiler sağladığını göstermiştir. Bununla birlikte, çoğu yanıt, hasta eğitim materyalleri için önerilen okunabilirlik seviyelerini aşmıştır. Bilgi kalitesi ve güvenilirliği genellikle orta düzeyde olarak değerlendirilmiş olup, tedavi riskleri, alternatif seçenekler ve kaynak şeffaflığı tartışmalarında eksiklikler bildirilmiştir. Klinik kılavuzlara uyum genel prensipler düzeyinde kabul edilebilir olsa da, tanısal ayrıntılarda ve tedavi sıralamasında tutarsızlıklar tespit edilmiştir. Yapay zekâ tabanlı sohbet robotları, ağrı ile ilgili bilgiler için destekleyici araçlar olarak potansiyel gösterse de, klinik karar verme veya birincil hasta eğitimi için bağımsız kaynaklar olarak mevcut kullanımları sınırlıdır. Bu sistemlerin güvenli kullanımı için insan gözetimi ve sağlık okuryazarlığına uygun içerik üretimi şarttır.

Anahtar Kelimeler: Ağrı, yapay zekâ, sohbet botları, sağlık okuryazarlığı, okunabilirlik, kalite değerlendirmesi

Corresponding Author: Erkan Özduran ✉ erkanozduran@gmail.com **Received:** 03.02.2026 **Revision:** 30.03.2026 **Accepted:** 14.04.2026

Copyright © 2026 The Author(s). Published by Zonguldak Bülent Ecevit University. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Highlights

- Artificial intelligence-based chatbots generally provide accurate basic information on pain; however, the readability of their responses frequently exceeds the recommended level for patient education materials.
- Across the included studies, the quality and reliability of AI-generated pain-related content were predominantly moderate, with notable deficiencies in source transparency and referencing.
- Evaluations of clinical guideline adherence demonstrated that AI chatbots align with general management principles but perform poorly in identifying red flags, prioritizing treatments, and delivering individualized recommendations.
- Comparative analyses revealed variability among ChatGPT, Gemini, Perplexity, and DeepSeek, with more advanced models generating more consistent responses that remain insufficiently patient-centered.
- Current evidence suggests that AI chatbots hold promise as supportive tools in pain management; however, their clinical application requires human oversight and standardized evaluation frameworks.

INTRODUCTION

Pain affects the lives of billions of people worldwide on a daily basis. Considering its high prevalence together with the associated loss of workforce productivity, disability, and mortality, it is evident that pain management constitutes a critical global health issue and a public health priority (1-3). Analysis of data from the chronic pain module of the 2019 National Health Interview Survey (NHIS) revealed that approximately 20.5% of the adult population in the United States—corresponding to 50.2 million individuals—reported experiencing pain on most days or continuously (4).

Data from a global study on pain prevalence encompassing 52 different countries demonstrated substantial variation in pain prevalence across nations, ranging from a minimum of 9.9% in China to a maximum of 50.3% in Morocco. The authors attributed this variability primarily to social and economic factors, emphasizing the prominent roles of income inequality and gender inequality in shaping these differences (1). In a study conducted in Türkiye, the prevalence of pain in the adult population was found to be remarkably high, reaching 80.84%. The authors suggested that this prevalence, which exceeds the global average, may be influenced by several factors, including social, cultural, and economic developments, changes in individuals' perception and interpretation of pain, and increased life expectancy (5).

As a result of the integration of technology into nearly every aspect of daily life, accessing health-related information via the internet has become an indispensable component of modern living (6). Available data indicate that nine out of ten individuals in the United States actively use the internet, and approximately 75% of these users consider it a primary tool for obtaining health-related information (7). Artificial Intelligence (AI), a relatively recent paradigm centered on minimizing human intervention by leveraging information technologies, is now widely employed across a broad spectrum of fields, including science, technology, and healthcare. At its core, AI refers to the development of computer systems capable of performing tasks that traditionally require human intelligence, such as decision-making, learning from experience, natural language understanding, image recognition, and problem-solving. With its transformative potential in healthcare delivery, AI has demonstrated substantial advancements within the medical domain (8-10).

Large Language Models (LLMs) are advanced AI systems trained using deep learning techniques on vast textual datasets and built upon transformer-based neural networks, enabling them to generate contextually appropriate, personalized, and human-like conversational responses. Owing to their exceptional predictive capacity derived from large-scale computation, these models are widely deployed as popular conversational agents such as ChatGPT, Gemini, Perplexity, and DeepSeek. They are continuously refined through iterative processes involving human feedback to align their outputs with ethical and pragmatic human preferences (11-13). AI-driven chatbots and virtual assistants are capable of interacting with patients, responding to queries, providing basic medical information, and even supporting appointment scheduling processes (14-16).

Despite these advantages, the literature has raised concerns that the responses generated by AI chatbots may be misleading and potentially detrimental to public health. Evaluations of ChatGPT's performance on the United States Medical Licensing Examination (USMLE) have shown accuracy rates ranging from only 44% to 64.4% in analyses based on multiple-choice questions. Similarly, in the 2023 Taiwan Pharmacist Licensing Examination, the model achieved correct response rates between 54.5% and 67.6%. These findings heighten concerns regarding the suitability of ChatGPT as a primary reference source

or study aid in healthcare-related contexts, particularly with respect to response quality and reliability (17).

In addition to concerns about reliability and quality, another emerging issue is whether the responses provided by AI chatbots are presented at an appropriate readability level that can be fully understood by individuals. Health information can only be accurately comprehended by patients in accordance with their health literacy levels. Acknowledging this, authoritative institutions such as the National Institutes of Health (NIH), the American Medical Association (AMA), and the United States Department of Health and Human Services recommend that internet-based patient education materials be written at or below a sixth-grade reading level. Texts prepared above this threshold are generally considered difficult for patients to read and interpret (18). AI has been utilized to assess and enhance the readability of online health resources and offers considerable potential to improve health literacy. However, this potential has not yet been fully realized, and significant concerns remain regarding accuracy, incompleteness, bias, and the risk of misinformation. Consequently, although the incorporation of human oversight and supervision is frequently recommended to ensure reliable information delivery, the practical implementation of such oversight remains unclear (19,20).

In the field of pain and pain management, persistent concerns regarding the quality and reliability of AI systems, ambiguities surrounding the concept of health literacy, and the limited availability of evidence-based data supporting these emerging applications at the intersection of AI and health literacy highlight the need for a comprehensive and systematic review of the existing literature. Findings derived from such an analysis would substantially contribute to the understanding the current state of the scientific landscape and to establishing a more structured framework for advancing future research. , this review specifically address the following research questions:

Which specific instruments have been used to assess health literacy in AI-related studies within the field of pain, and what are the readability evaluation outcomes of responses generated by AI chatbots?

From the perspective of quality and reliability, what do the findings of AI-related pain studies indicate?

METHODS

Study Design

This study was conducted as a systematic scoping review with the aim of comprehensively mapping the existing literature on AI applications in the field of pain. The review focused on outcomes related to readability, quality, reliability, and health literacy. The methodology was planned in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) guidelines. Given that the evidence base in this area is still emerging, encompasses diverse study designs, and exhibits substantial heterogeneity in outcome measures, a scoping review approach was deemed the most appropriate.

Literature Search Strategy

A comprehensive literature search was conducted across the PubMed/MEDLINE, Scopus, Web of Science, and Embase databases. All databases were searched from inception to February 2, 2026. In addition, the reference lists of relevant articles were manually screened, and supplementary searches were performed using Google Scholar to identify additional eligible studies.

The search strategy was designed to combine three main conceptual domains: pain, artificial intelligence, and readability/quality assessment. The primary keywords used included:

- Pain-related terms: pain, chronic pain, pain management, analgesia
- AI-related terms: artificial intelligence, AI, chatbot, large language model, LLM, ChatGPT, Gemini, conversational agent
- Evaluation-related terms: readability, health literacy, quality, accuracy, reliability

These terms were combined using Boolean operators (AND/OR) and adapted as appropriate for each database.

Although the literature search was conducted from database inception to February 2, 2026, the studies ultimately included in this review were published between 2024 and 2025, reflecting the period during which eligible studies meeting all inclusion criteria were available.

An example of the search strategy used in PubMed is presented below:

("pain" OR "chronic pain" OR "pain management" OR analgesia) AND ("artificial intelligence" OR AI OR chatbot OR "large language model" OR LLM OR ChatGPT OR Gemini OR "conversational agent") AND (readability OR "health literacy" OR quality OR accuracy OR reliability)

Eligibility Criteria

Studies were selected based on predefined inclusion and exclusion criteria.

Inclusion criteria

- Studies evaluating AI systems in the context of pain or pain management

- Research examining the readability, quality, reliability, accuracy, or health literacy of medical information generated by AI
- Studies involving chatbots, conversational agents, or large language models
- Original research articles published in peer-reviewed journals

Exclusion criteria

- Editorials, commentaries, opinion pieces, and letters lacking primary data
- Technical AI studies without clinical or patient-oriented pain-related content
- Duplicate publications

Study Selection Process

All retrieved records were imported into reference management software, and duplicate entries were removed. Titles and abstracts were screened independently by two reviewers. Full texts of potentially eligible studies were subsequently assessed according to the inclusion criteria. Any disagreements were resolved through discussion and consensus. The study selection process is presented in a PRISMA flow diagram (Figure 1).

Data Extraction

A standardized data extraction form was developed. The following information was extracted from each included study:

- Author(s) and year of publication
- Country
- Type of AI system evaluated
- Pain-related topic or clinical context
- Study design and methodology
- Readability assessment tools used
- Quality and/or reliability evaluation methods
- Key findings
- Reported limitations

Data extraction was performed independently by two reviewers, and accuracy was ensured through cross-check.

Quality Appraisal

Due to marked methodological heterogeneity among the included studies in terms of design characteristics, assessment tools, and reported outcome measures, a narrative appraisal approach was adopted. Each study was systematically examined with respect to its methodological design, the validity of the readability and quality assessment instruments employed, analytical approach, and transparency of reporting. Methodological strengths and limitations were identified and taken into account during the synthesis and interpretation of the findings.

Data Synthesis

The extracted data were synthesized descriptively. Studies were grouped into thematic categories, including readability, health literacy instruments, accuracy, and reliability assessments. Findings were presented in both tabular and narrative formats to highlight prevailing trends, methodological approaches, and existing research gaps.

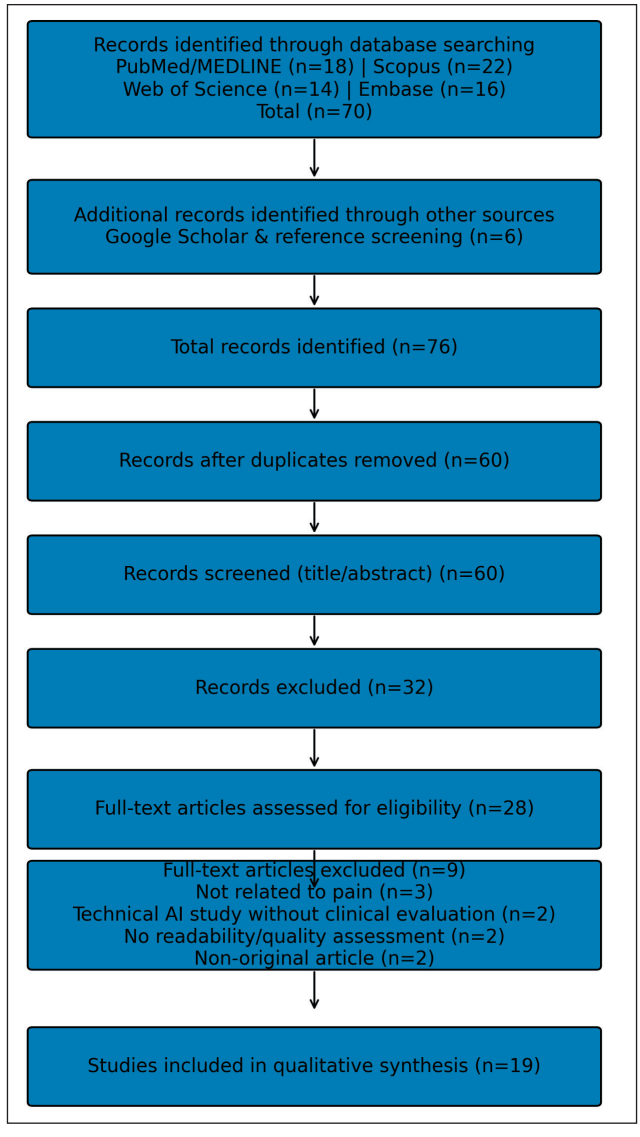


Figure 1: PRISMA flow diagram of the study selection process. PRISMA flow diagram illustrating the identification, screening, eligibility assessment, and final inclusion of studies evaluating artificial intelligence-based chatbots in the field of pain.

Table 1: Characteristics and key findings of studies evaluating artificial intelligence chatbots for pain-related information

Author (Year)	Country	Pain Domain	Evaluated AI Models	Study Design	Readability Tools	Quality / Reliability Tools	Key Findings
Ozduran et al. (22)	Türkiye	General pain	ChatGPT, Gemini, Perplexity	Cross-sectional	FRES, FKGL, SMOG, GFI, CLI	DISCERN, JAMA, GQS	All models exceeded the 6th-grade reading level; overall quality was moderate; source transparency was insufficient
Ozduran et al. (22)	Türkiye	Low back pain	ChatGPT, Gemini, Perplexity	Cross-sectional	FRES, FKGL, SMOG	DISCERN, JAMA	Readability was inadequate; DISCERN scores were mostly "fair-moderate"
Scaff et al. (23)	Australia	Low back pain	ChatGPT	Cross-sectional	-	Accuracy, clinical appropriateness	Overall accuracy was acceptable; personalization was limited
Tabanli & Demirkiran (24)	Türkiye	Low back pain	ChatGPT-3.5 vs ChatGPT-4.0	Comparative	FKGL	Content analysis	GPT-4 produced more comprehensive and empathetic responses; psychosocial aspects were partially lacking
Rossettini et al. (25)	Italy	Lumbar radicular pain	ChatGPT, Gemini, Perplexity, Copilot, Claude	Cross-sectional	-	Guideline adherence	General principles were aligned; red flags and treatment prioritization were insufficient
Gianola et al. (26)	Italy	Lumbar radicular pain	ChatGPT	Cross-sectional	-	Guideline comparison	Clinical decision-making was superficial compared with guidelines
Deng et al. (27)	China	PDPH	ChatGPT, DeepSeek	Cross-sectional	-	Guideline adherence	Basic information was accurate; deficiencies noted in invasive interventions
Lo Bianco et al. (28)	Italy	Interventional pain	ChatGPT	Cross-sectional	FKGL	DISCERN, GQS	Technical language predominated; patient-centeredness was low
Yücel et al. (29)	Türkiye	Cancer pain	ChatGPT	Cross-sectional	-	ESMO adherence	General adherence was good; adverse effects and personalization were insufficient
García-Rudolph et al. (30)	Spain	Neuropathic pain	GPT-3.5, GPT-4	Cross-sectional	-	Consistency, accuracy	GPT-4 was more consistent; educational value was limited
Ateş Demiroğlu et al. (31)	Türkiye	Myofascial pain	ChatGPT	Cross-sectional	FKGL	DISCERN	Academic tone dominated; not suitable for patient education
Brant-Zawadzki et al. (32)	USA	Acute chest pain	ChatGPT-4, Gemini	Simulation-based	-	Quality assurance	Should be used cautiously in emergency triage
Aljamani et al. (33)	Jordan	Endodontic pain	ChatGPT, Gemini	Cross-sectional	FRES, FKGL	JAMA	Accuracy was acceptable; references were insufficient
Mohammad-Rahimi et al. (34)	Iran	Endodontics	ChatGPT	Cross-sectional	FKGL	Accuracy, reliability	Moderate reliability
Degirmencioglu & Temel (35)	Türkiye	Toothache	ChatGPT	Cross-sectional	FRES, SMOG	DISCERN	Poor readability; moderate quality
Gondode et al. (36)	India	Chronic pain medications	Chatbot vs brochure	Comparative	-	Patient education	Chatbot was more accessible; content depth was limited
Basharat et al. (37)	USA	Low back pain	ChatGPT	Cross-sectional	FKGL	Clinical accuracy	Errors were mainly due to missing information
Dong et al. (38)	China	MSK pain - topical NSAIDs	ChatGPT, Gemini, DeepSeek, Grok	Comparative	FRES, FKGL	Accuracy, PEMAT-P	High reading level; low actionability
Ah-Yan et al. (39)	Canada	Low back pain	ChatGPT vs Copilot	Qualitative	-	Clinical Likert scale	High validity and safety

RESULTS

General Characteristics of the Included Studies

The studies included in this review were published between 2024 and 2025. Although the literature search was conducted from database inception until February 2, 2026, all studies meeting the inclusion criteria fell within this publication range. The included studies aimed to evaluate the performance of AI-based chatbots in the field of pain with respect to readability, information quality, reliability, and adherence to clinical guidelines. The vast majority of studies employed a cross-sectional design in which AI model responses were analyzed using queries derived from frequently asked patient-centered questions (21-24,28-31,37-39).

The most commonly addressed pain subtopics included general pain (21), low back pain and lumbar radicular pain (22-26,37,39), chronic pain interventions such as radiofrequency ablation and cryoanalgesia (28), cancer-related pain (29), neuropathic pain (30), myofascial pain syndrome (31), dental and endodontic pain (33-35), acute chest pain triage (32), and topical NSAID use in musculoskeletal pain (38).

The AI models evaluated across the studies included ChatGPT (versions 3.5, 4.0, and 4o), Gemini, Perplexity, DeepSeek, Copilot, and Claude. While some studies focused on a single model (29-31,35,37), the majority conducted comparative analyses involving multiple AI models (21-25,27,32,33,38,39).

Readability Outcomes

Readability assessments were most commonly conducted using the Flesch Reading Ease Score (FRES), Flesch-Kincaid Grade Level (FKGL), SMOG Index, Gunning Fog Index, and Coleman-Liau Index (21,22,28,31,35,37,38).

Overall, AI-generated responses were consistently reported to exceed the recommended sixth-grade reading level and were classified within the moderate-to-difficult readability range (21,22,35,37,38). Ozduran et al. demonstrated that responses generated by ChatGPT, Gemini, and Perplexity to general pain-related questions all exceeded the recommended threshold for patient education materials according to SMOG and FKGL scores (21). Similarly, studies focusing on low back pain reported that readability levels across all evaluated models were clinically inadequate (22,37).

Dong et al. reported that responses regarding topical NSAID use in musculoskeletal pain corresponded to a ninth- to tenth-grade reading level across all models, and that none met acceptable thresholds for patient actionability (38).

Comparable trends were observed across specific pain subtypes. In queries related to radiofrequency ablation and cryoanalgesia, ChatGPT responses were found to contain technical language and demonstrate poor readability (28). Studies examining myofascial pain syndrome, dental pain, and low back pain emphasized that AI-generated responses were close to an academic reading level and lacked simplified, patient-centered explanations (31,35,37).

Quality and Reliability Assessments

The most frequently used tools for evaluating information quality and reliability included DISCERN, the JAMA Benchmark Criteria, EQIP, the Global Quality Score (GQS), and clinician-rated Likert scales (21-23,28,31,33-39).

Across most studies, AI-generated responses were reported to provide moderate-quality information. Ozduran et al. found that DISCERN scores in studies on general pain and low back pain predominantly fell within the “fair-moderate” range, with notable deficiencies in areas such as treatment risks, alternative options, and source transparency (21,22).

Basharat et al. demonstrated that ChatGPT responses related to low back pain were generally acceptable in terms of overall accuracy, with most errors attributable to information omission rather than factual inaccuracies (37). The same study highlighted that response length was not consistently associated with higher accuracy.

Dong et al. reported significant differences among LLMs in terms of accuracy and over-conclusiveness, noting that Gemini and Grok exhibited relatively more balanced performance. Nevertheless, all evaluated models were found to be insufficient with regard to full guideline adherence and clinical actionability (38).

In contrast, a qualitative study by Ah-Yan et al. showed that responses generated by ChatGPT and Copilot for low back pain were rated highly by physiatrists in terms of validity, safety, and clinical usefulness (39).

Adherence to Clinical Guidelines

Adherence to clinical guidelines was primarily assessed in the contexts of lumbar radicular pain, post-dural puncture headache, cancer pain, and pharmacological management of musculoskeletal pain (25-27,29,38).

Rossettini et al. reported that ChatGPT, Gemini, and Perplexity were generally aligned with clinical guidelines at the level of broad principles; however, inconsistencies were identified in areas such as red flag recognition, diagnostic guidance, and treatment sequencing (25). Similarly, Gianola et al. demonstrated that ChatGPT offered more superficial recommendations compared with guideline-based clinical decision-making processes (26).

Dong et al. found that responses regarding topical NSAID use were largely consistent with guideline recommendations; however, notable gaps were observed concerning contraindications, patient subgroups, and practical application details (38).

Inter-Model Comparisons (ChatGPT vs Gemini vs Perplexity vs DeepSeek vs Copilot)

Studies conducting multi-model comparisons indicated that ChatGPT tended to produce longer and more detailed responses, Gemini offered more structured outputs, Perplexity emphasized source-based information, and Copilot provided comparatively balanced responses with regard to clinical safety (21-25,32,38,39).

No single model demonstrated a clear advantage in terms of readability; all models exceeded the recommended level for patient education materials (21,22,37,38). In quality assessments, ChatGPT, Gemini, and Copilot achieved similar clinical safety scores, while Perplexity attained relatively higher JAMA Benchmark scores due to greater source transparency (21,25).

DISCUSSION

This comprehensive review delineates the current landscape of information provided by AI-based chatbots in the field of pain with respect to readability, quality, reliability, and adherence to clinical guidelines. The reviewed studies collectively indicate that AI systems are generally capable of delivering accurate foundational information related to pain; however, they also reveal substantial limitations in the contexts of patient education, clinical decision support, and health literacy (21-39).

Interpretation of Findings from the Perspective of Readability and Health Literacy

One of the most consistent findings of this review is that the majority of responses generated by different AI models exceed the recommended readability level for patient education materials (21,22,28,31,35). This observation suggests that current AI applications in the field of pain domain are not yet fully aligned with the goals of patient-centered communication.

Despite national and international recommendations advocating for the use of language at or below a sixth-grade reading level, AI-generated responses frequently incorporate technical terminology. This may increase the risk of misunderstanding and inappropriate decision-making among individuals with limited health literacy. The clinical implications of this issue become particularly pronounced in conditions requiring long-term management, such as chronic pain and low back pain (21-23).

Information Quality and Reliability: Moderate Overall Performance

Assessments of information quality and reliability demonstrate that AI chatbots most commonly achieve moderate DISCERN, GQS, and JAMA scores (21-23,28,31,33-36). These findings suggest that while AI-generated information cannot be considered inherently unreliable, its use as a standalone primary information source may be problematic.

Common deficiencies reported across studies include inadequate source citations, limited presentation of treatment alternatives, and superficial discussion of potential risks. In studies focusing on cancer-related pain and interventional pain management, AI systems were found to provide guideline-consistent general information but lacked the capacity to support individualized clinical decision-making (28,29).

Limitations in Adherence to Clinical Guidelines

Guideline-based comparisons indicate that AI models demonstrate an acceptable performance with respect to general principles of pain management; however, inconsistencies persist in critical areas such as diagnostic specificity, recognition of red flags, and treatment sequencing (25-27).

These limitations raise concerns that, particularly in common clinical scenarios such as low back pain and lumbar radicular pain, the unsupervised use of AI systems may result in misleading guidance (25,26). In more specific clinical contexts, such as post-dural puncture headache, studies have reported high levels of factual accuracy but insufficient practical detail to guide clinical application (27).

Clinical Implications of Inter-Model Differences

Comparative analyses among ChatGPT, Gemini, Perplexity, and DeepSeek suggest that rather than a single superior model, each system exhibits distinct strengths and limitations (21-25,27,32). ChatGPT's comprehensive and explanatory responses, Gemini's structured output, and Perplexity's source-oriented approach offer varying advantages depending on the clinical or educational context.

Nevertheless, the consistent failure of all models to achieve optimal performance in terms of readability and guideline adherence supports the notion that these systems should be positioned as adjunctive tools, rather than replacements for clinician judgment, in pain management.

In clinical practice, AI-based chatbots may serve as supportive tools across several scenarios. For patient education, chatbots can provide accessible explanations of pain conditions, treatment options, and self-management strategies, enhancing patient engagement and health literacy. In triage contexts, AI systems could help prioritize patient needs by identifying potential red flags or urgent symptoms, thereby assisting clinicians in resource allocation. Furthermore, AI chatbots may act as adjunctive decision support tools for healthcare providers, offering guideline-based summaries or reminders that complement clinical judgment. While these applications are promising, human oversight remains essential to ensure accuracy, contextual appropriateness, and patient safety.

Comparison with Country-Specific and Non-English Evidence

While most studies included in this review evaluated English-language outputs, emerging research in other languages provides additional insights into the performance and limitations of AI-generated health information across different linguistic contexts. For example, a study conducted in Türkiye by Ozduran et al. evaluated Turkish internet-based patient education materials related to low back pain and reported moderate readability but low reliability and poor information quality based on JAMA and DISCERN criteria (40). These findings highlight persistent challenges in ensuring accessible and trustworthy health information in non-English digital environments. Similar language-related disparities have also been observed in AI-generated responses. For instance, Gonzalez Fiol et al. demonstrated that ChatGPT-generated responses to labor epidural questions were significantly less accurate in Spanish compared with English, raising concerns about potential inequities in information quality among non-English-speaking users (41). In another context, Wang et al. evaluated large language models using Chinese anesthesiology residency examination questions and found strong performance in knowledge recall but reduced accuracy in complex clinical reasoning tasks, suggesting that language and contextual factors may influence model reliability (42). Furthermore, linguistic complexity and the limited availability of structured digital resources—issues highlighted in studies on Arabic-language AI applications such as those discussed by Madwi et al.—may further affect the development and evaluation of AI-driven health information systems (43). Taken together, these findings suggest that the limitations observed in English-language studies may not only persist but may be amplified in other languages. Therefore, future research should prioritize multilingual evaluations and culturally adapted assessment frameworks to ensure that AI-generated medical information remains accurate, reliable, and accessible across diverse linguistic and regional contexts. Furthermore, as has been shown many times in the literature, future studies focusing on different cultures and languages that examine bibliometric and altmetric data related to pain in the field of artificial intelligence will guide the literature (44).

Summary of the Evidence

Overall, this review demonstrates that AI-based chatbots in the field of pain:

- Generally provide accurate foundational information about pain,
- Fall short of recommended patient education standards in terms of readability,
- Exhibit predominantly moderate performance with respect to information quality and reliability, and
- Align with clinical guidelines at the level of general principles but remain insufficient for detailed clinical decision-making (21-39).

The available evidence suggests that AI-based chatbots may be used to support patient information and self-management, but their independent use in clinical decision-making processes without human oversight is not appropriate.

Strengths and Limitations

Strengths

- This review represents one of the first comprehensive syntheses addressing the use of AI in pain management through an integrated evaluation of readability, quality, reliability, and guideline adherence.
- The inclusion of a wide range of pain subtypes and multiple AI models enhances the generalizability of the findings.
- The frequent use of standardized and validated assessment tools (e.g., DISCERN, JAMA, FRES, SMOG) across included studies strengthens the methodological rigor.

Limitations

- All included studies employed cross-sectional designs, preventing assessment of causal relationships or the impact of AI-generated information on real-world clinical outcomes. Future research should prioritize prospective studies to evaluate how AI chatbots influence clinical decision-making, patient behavior, and health outcomes in real-world pain management scenarios.
- Considerable heterogeneity existed in the queries, AI models, and assessment methods used across studies, which may limit comparability and generalizability.
- Most studies focused on English-language content, restricting the evaluation of AI performance across different languages and cultural contexts.
- Continuous updates and version changes in AI models mean that reported findings may evolve over time.
- Potential sources of bias should be acknowledged, including selection bias (e.g., focus on English-language studies), reporting bias (e.g., incomplete disclosure of AI model versions), and assessment bias (subjectivity in readability or quality evaluations). Recognizing these limitations is essential for interpreting current evidence and guiding more standardized, methodologically robust future research.

Conclusion

Despite the rapid proliferation of AI-based chatbots in the field of pain, this comprehensive review demonstrates that these systems continue to exhibit significant limitations with respect to readability, information quality, reliability, and adherence to clinical guidelines. The reviewed studies indicate that while AI models are generally capable of providing accurate and consistent foundational information about pain, they remain inadequate in meeting recommended health literacy standards for patient education.

Current evidence suggests that AI holds promise as a supportive tool for facilitating access to information in pain management; however, its independent use as a primary source of patient information or within clinical decision-making processes is not appropriate. In critical domains such as risk-benefit assessment, individualized recommendations, and recognition of red flags, AI-generated responses tend to remain superficial when compared with established clinical guidelines. Consequently, the unsupervised use of AI applications in pain management poses potential risks to patient safety. Future prospective studies are urgently needed to evaluate the actual clinical impact of AI-supported interventions on patient outcomes, adherence to treatment plans, and overall pain management effectiveness.

Nevertheless, improvements in generating content at appropriate readability levels, tighter integration with clinical practice guidelines, and systematic evaluation of AI-supported information in real-world clinical scenarios may contribute to positioning these technologies more safely and effectively within pain management frameworks.

Author Contributions

Study conception and design: **Ilhan Celil Ozbek, Erkan Ozduran, Volkan Hanci;** data collection: **Ilhan Celil Ozbek, Erkan Ozduran, Volkan Hanci;** analysis and interpretation of results: **Ilhan Celil Ozbek, Erkan Ozduran, Volkan Hanci;** draft manuscript preparation: **Ilhan Celil Ozbek, Erkan Ozduran, Volkan Hanci.** The author(s) reviewed the results and approved the final version of the article.

Conflicts of Interest

The authors declare that there is no conflict of interest to disclose.

Ethical Approval

This study does not require ethics committee approval as it does not involve human, animal or sensitive data.

Use of Artificial Intelligence

During the preparation of this study, we used OpenAI's ChatGPT (version 5-turbo) between January and February 2026 to improve the English language, enhance clarity, and strengthen the academic tone of the manuscript. Following the use of this tool, the authors carefully reviewed and edited the content as necessary and take full responsibility for the content of the publication.

REFERENCES

1. Zimmer Z, Fraser K, Grol-Prokopczyk H, Zajacova A. A global study of pain prevalence across 52 countries: examining the role of country-level contextual factors. Pain. 2022;163(9):1740-1750. https://doi.org/10.1097/j.pain.0000000000002557
2. Breivik H, Eisenberg E, O'Brien T; OPENMinds. The individual and societal burden of chronic pain in Europe: the case for strategic prioritisation and action to improve knowledge and availability of appropriate care. BMC Public Health. 2013 Dec 24;13:1229. https://doi.org/10.1186/1471-2458-13-1229.
3. Rice ASC, Smith BH, Blyth FM. Pain and the global burden of disease. Pain. 2016;157(4):791-6. https://doi.org/10.1097/j.pain.0000000000000454.
4. Yong RJ, Mullins PM, Bhattacharyya N. Prevalence of chronic pain among adults in the United States. Pain. 2022;163(2):e328-e332. https://doi.org/10.1097/j.pain.0000000000002291
5. Eti Aslan F, Çınar F. Prevalence of pain in adult population in Türkiye. Agri. 2023;35(2):83-95. https://doi.org/10.14744/agri.2022.26086
6. Özbek İC. Evaluation of the Quality, Reliability, and Popularity of YouTube Videos on Thoracic Outlet Syndrome: A Critical Analysis. Journal of Physical Medicine and Rehabilitation Sciences. 2025;28(3):207-17
7. Ozduran E, Hanci V, Erkin Y. Evaluating the readability, quality and reliability of online patient education materials on chronic low back pain. Natl Med J India. 2024;37(3):124-130. https://doi.org/10.25259/NMJI_327_2022
8. Özbek İC. Evaluation of Artificial Intelligence-Supported Osteoarthritis Information Texts: Content Quality and Readability Analysis. Journal of Physical Medicine and Rehabilitation Sciences. 2025;28(1):21-9.
9. Ensari E, Onder ENA, Ertan P. A Comparative Assessment of Large Language Models in Pediatric Dialysis: Reliability, Quality and Readability. Ther Apher Dial. 2025 Oct;29(5):739-746. https://doi.org/10.1111/1744-9987.70058.
10. Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism. 2017;69S:S36-S40. https://doi.org/10.1016/j.metabol.2017.01.011
11. Bashah A, Salem A, Al-Waqeerah A, Ghaleb E, Wahan N, Awad A,, et al. Evaluation of deepseek, gemini, ChatGPT-4o, and perplexity in responding to salivary gland cancer. BMC Oral Health. 2025;25(1):1358. https://doi.org/10.1186/s12903-025-06726-4
12. Özbek İC, Hancı V, Özduran E. Digital Guidance: Quality and Readability Analysis of Artificial Intelligence-Generated Spondyloarthropathy Texts. Turk J Osteoporos. 2025;31(1):12-18. https://doi.org/10.4274/tod.galenos.2024.76743
13. Anıl H, Kayra MV. The digital dialogue on premature ejaculation: evaluating the efficacy of artificial intelligence-driven responses. Int Urol Nephrol. 2025 Sep;57(9):2829-2836. https://doi.org/10.1007/s11255-025-04461-x.
14. Hopkins AM, Logan JM, Kichenadasse G, Sorich MJ. Artificial intelligence chatbots will revolutionize how cancer patients access information: ChatGPT represents a paradigm-shift. JNCI Cancer Spectr. 2023;7(2):pkad010. https://doi.org/10.1093/jncics/pkad010

15. Özbek İC, Özduran E. Digital Rehabilitation in Parkinson's Disease: The Role of Artificial Intelligence-Assisted Exercise Training. Turk J Osteoporos. Published online Sep 12, 2025. https://doi.org/10.4274/tod.galenos.2025.66664

16. Topan H, Baydoğan GM, Sürme Y. Revolutionizing patient care: assessing ai-created discharge education for total hip replacement patients. BMC Med Educ. 2025 Nov 24;25(1):1636. https://doi.org/10.1186/s12909-025-08186-4.

17. Shiferaw MW, Zheng T, Winter A, Mike LA, Chan LN. Assessing the accuracy and quality of artificial intelligence (AI) chatbot-generated responses in making patient-specific drug-therapy and healthcare-related decisions. BMC Med Inform Decis Mak. 2024;24(1):404. https://doi.org/10.1186/s12911-024-02824-5

18. Kara M, Ozduran E, Kara MM, Özbek İC, Hancı V. Evaluating the readability, quality, and reliability of responses generated by ChatGPT, Gemini, and Perplexity on the most commonly asked questions about Ankylosing spondylitis. PLoS One. 2025 Jun 18;20(6):e0326351. https://doi.org/10.1371/journal.pone.0326351.

19. Abeo ANA, Armstrong S, Scriney M, Goss H. Artificial Intelligence Techniques and Health Literacy: A Systematic Review. Mayo Clin Proc Digit Health. 2025;3(4):100269. https://doi.org/10.1016/j.mcpdig.2025.100269

20. Gokalp MG, Yucel SC. Comparative analysis of nursing care plans produced by artificial intelligence models (ChatGPT, Gemini, and DeepSeek) in terms of readability, reliability, and quality. BMC Nurs. 2026 Jan 12. https://doi.org/10.1186/s12912-026-04295-7.

21. Ozduran E, Akkoc I, Büyükçoban S, Erkin Y, Hanci V. Readability, reliability and quality of responses generated by ChatGPT, Gemini, and Perplexity for the most frequently asked questions about pain. Medicine (Baltimore). 2025;104(11):e41780. https://doi.org/10.1097/MD.0000000000041780.

22. Ozduran E, Hancı V, Erkin Y, Özbek İC, Abdulkerimov V. Assessing the readability, quality and reliability of responses produced by ChatGPT, Gemini, and Perplexity regarding most frequently asked keywords about low back pain. PeerJ. 2025;13:e18847. https://doi.org/10.7717/peerj.18847.

23. Scaff SPS, Reis FJJ, Ferreira GE, Jacob MF, Saragiotto BT. Assessing the performance of AI chatbots in answering patients' common questions about low back pain. Ann Rheum Dis. 2025;84(1):143-149. https://doi.org/10.1136/ard-2024-226202.

24. Tabanli A, Demirkiran ND. Comparing ChatGPT 3.5 and 4.0 in low back pain patient education: addressing strengths, limitations, and psychosocial challenges. World Neurosurg. 2025;196:123755. https://doi.org/10.1016/j.wneu.2025.123755.

25. Rossettini G, Bargeri S, Cook C, Guida S, Palese A, Rodeghiero L, et al. Accuracy of ChatGPT-3.5, ChatGPT-4o, Copilot, Gemini, Claude, and Perplexity in advising on lumbosacral radicular pain against clinical practice guidelines. Front Digit Health. 2025;7:1574287. https://doi.org/10.3389/fdgth.2025.1574287.

26. Gianola S, Bargeri S, Castellini G, Cook C, Palese A, Pillastrini P, et al. Performance of ChatGPT compared to clinical practice guidelines in making informed decisions for lumbosacral radicular pain. J Orthop Sports Phys Ther. 2024;54(3):222-228. https://doi.org/10.2519/jospt.2024.12151.

27. Deng J, Qiu X, Dong C, Xu L, Dong X, Yang S, et al. Evaluating ChatGPT and DeepSeek in postdural puncture headache management: a comparative study with international consensus guidelines. BMC Neurol. 2025;25(1):264. https://doi.org/10.1186/s12883-025-04280-8.

28. Lo Bianco G, Cascella M, Natoli S, D'Angelo FP, Sinagra E, Marchesini M, et al. Assessing artificial intelligence-powered responses to common patient questions on radiofrequency ablation and cryoanalgesia for chronic pain. J Clin Med. 2025;14(19):6814. https://doi.org/10.3390/jcm14196814.

29. Yücel K, Sutcuoglu O, Yazıcı O, Ozet A, Ozdemir N. Can artificial intelligence provide accurate and reliable answers to cancer patients' questions about cancer pain? Comparison of chatbots based on ESMO cancer pain guideline. memo. 2024;17(4):302-306.

30. García-Rudolph A, Sanchez-Pinsach D, Opisso E, Soler MD. Exploring new educational approaches in neuropathic pain: assessing accuracy and consistency of artificial intelligence responses from GPT-3.5 and GPT-4. Pain Med. 2025;26(1):48-50. https://doi.org/10.1093/pm/pnae094.

31. Ateş Demiroglu SB, Özpolat Bulut Ö, Bağcier F. AI-powered insights: analyzing ChatGPT's responses on myofascial pain syndrome. J Bodyw Mov Ther. 2025;44:558-563. https://doi.org/10.1016/j.jbmt.2025.05.043. PMID:40954629.

32. Brant-Zawadzki G, Klapthor B, Ryba C, Youngquist DC, Burton B, Palatinus H, et al. The performance of ChatGPT-4 and Gemini Ultra 1.0 for quality assurance review in emergency medical services chest pain calls. Prehosp Emerg Care. 2025;29(3):210-217. https://doi.org/10.1080/10903127.2024.2376757.

33. Aljamani S, Hassona Y, Fansa HA, M Saadeh H, Dafi Jamani K.. Evaluating large language models in addressing patient questions on endodontic pain. J Endod. 2025;51(11):1617-1624. https://doi.org/10.1016/j.joen.2025.04.015.

34. Mohammad-Rahimi H, Ourang SA, Pourhoseingholi MA, Dianat O, Dummer PMH, Nosrat A. Validity and reliability of artificial intelligence chatbots as public sources of information on endodontics. Int Endod J. 2024;57(3):305-314. https://doi.org/10.1111/iej.14014.

35. Degirmencioglu D, Temel AN. Evaluation of the quality and readability of ChatGPT responses to toothache queries. Cureus. 2025;17(11):e96436. https://doi.org/10.7759/cureus.96436.

36. Gondode P, Duggal S, Garg N, Sethupathy S, Asai O, Lohakare P. Comparing patient education tools for chronic pain medications: artificial intelligence chatbot versus traditional patient information leaflets. Indian J Anaesth. 2024;68(7):631-636.

37. Basharat A, Shah R, Wilcox N, Tur G, Tripati S, Kansal P, et al. ChatGPT and low back pain - Evaluating AI-driven patient education in the context of interventional pain medicine. Interv Pain Med. 2025 Sep 2;4(3):100636. https://doi.org/10.1016/j.inpm.2025.100636.

38. Dong C, Qiu X, Deng J, Xu L, Dong X, Chen S, et al. Comparative evaluation of large language models in delivering guideline-compliant recommendations for topical NSAID use in musculoskeletal pain: a multidimensional analysis. Clin Rheumatol. 2025 Nov;44(11):4703-4710. https://doi.org/10.1007/s10067-025-07640-4.

39. Ah-Yan C, Boissonnault È, Boudier-Revéret M, Mares C. Impact of artificial intelligence in managing musculoskeletal pathologies in physiatry: a qualitative observational study evaluating the potential use of ChatGPT versus Copilot for patient information and clinical advice on low back pain. J Yeungnam Med Sci. 2025;42:11. https://doi.org/10.12701/jyms.2024.01151.

40. Özduran E. “Bel Ağrısı” ile İlgili Türkçe İnternet Kaynaklı Hasta Eğitim Materyallerinin Okunabilirliklerinin Değerlendirilmesi. DEU Tıp Derg. 01 Ağustos 2022;36(2):135-50. https://doi.org/10.18614/deutip.1174522

41. Gonzalez Fiol A, Mootz AA, He Z, Delgado C, Ortiz V, Reale SC. Accuracy of Spanish and English-generated ChatGPT responses to commonly asked patient questions about labor epidurals: a survey-based study among bilingual obstetric anesthesia experts. Int J Obstet Anesth. 2025;61:104290. https://doi.org/10.1016/j.ijoa.2024.104290. Epub 2024 Nov 6.

42. Wang S, Chi X, Hao Q, Wang H, Tao H, Xiao J, et al. Large language models in Chinese anesthesiology residency examinations: a comparative analysis of performance, reliability and clinical reasoning. BMC Med Educ. 2026 Jan 31;26(1):348. https://doi.org/10.1186/s12909-026-08704-y.

43. Madwi FHM. Integrating artificial intelligence in Arabic language education: challenges and opportunities. Dzil Majaz J Arab Lit. 2025;3(1):45-55.

44. Bagcier F, Yurdakul OV, Ozduran E. Top 100 cited articles on ankylosing spondylitis. Reumatismo. 2021;72(4):218-227. https://doi.org/10.4081/reumatismo.2020.1325

Özgün Araştırmalar / Original Researchs

Investigation of factors associated with the level of temporomandibular dysfunction in university students
Merve Bulguroğlu, Halil İbrahim Bulguroğlu, Gökçe Kılıç, Selcan Suiçmez, Serenay Zorlu, Cansu Gevrek Aslan

In-silico engineering of an antimicrobial growth factor fusion protein to prevent post-laser infection in hemorrhoidal wounds
Özlem Karaca Ocak, Zoha Naeem

The correlation between serum YKL-40, proteinuria and inflammation in patients with systemic lupus erythematosus
İlyas Öztürk, Fatma Betül Güzel, Serhat Kura, Enes Süleyman Gün, Burak Okyar, Sinan Kazan, Muhammed Seyithanoğlu, Gözde Yıldırım Çetin, Orçun Altunören, Özkan Güngör

Predictors of early readmission in patients awaiting interval cholecystectomy after conservative treatment of acute cholecystitis: A retrospective cohort study
Kazım Duman

Investigation of the effects of odontoid bone anatomical location and structure on cervical lordosis angle in cervical CT examinations: A clinical study
Alemiddin Özdemir, Ahmet Melih Erdoğan, Özge Sevimoğlu, Buket Oğuz, Buse Kaymakcı, Selma Çalışkan, Bülent Bakar

Clinical and radiological response assessment after whole brain radiotherapy for patients with non-small cell lung cancer with brain metastases
Funda Çukurçayır, Mert Saynak, Canan Çelebi, Alaattin Özen, Görkem Türkkan, Kamuran İbiş, Mustafa Akın, Zafer Koçak

Evaluation of viral respiratory pathogens detected in a tertiary university hospital between 2021 and 2025
Canberk Çınar, Sümeyye Özkaya, İlknur Bıyık, Yeliz Tanrıverdi Çaycı, Demet Gür Vural, Kemal Bilgin, Asuman Birinci

Comparative analysis of pH measurements: Urine dipstick and pH meter
Murat Can, Berrak Güven, Osman Emre Ünlü

Applications and research trends of large language models in healthcare: A bibliometric analysis
Hakan Öztürk, Elvan Hayat

Clinical characteristics and hematological inflammatory markers in pediatric peripheral facial paralysis: A retrospective study
Meral Karadağ, Melek Yaman Ortaköylü

Evaluation of practical application of Tell-Hashomer, Livneh and Yalçınkaya-Özen criteria for diagnosis of Familial Mediterranean Fever in children
Tuğba Kazlı, Müferet Ergüven

The effect of a surgical nursing course structured with role-playing and flipped learning methods on students' motivation and attitudes toward clinical practice
Sibel Altıntaş, Sevim Çelik, Elif Karahan, Özge Uçar, Meryem Yücel

Postoperative frequency of normocalcemic elevated parathormone in patients surgically treated for primary hyperparathyroidism: A single-center experience
Gülşah Elbüken, Ogün İrem Bilen, Beyza Olcay Öztürk, Hünkar Ağgül, Sibel Özkan-Gürdal, Burcu Tokuç, Sayid Shafi Zuhur

Adneksiyal kitlesi olan kadınlarda siberkondri düzeyini etkileyen faktörler: Tek merkezli kesitsel çalışma
Hande Nur Öncü, Gökçen Ege, Neslihan Öztürk, Candost Hanedan, Vakkas Korkmaz

Sensory perception profiles in children with epilepsy: A prospective controlled study
Seda Ayaz Taş, Merve Başol Göksülük, Özge Karanlık Özcan, Seda Yakıt Yeşilyurt, Sedef Kestellioğlu, Sezen Tezcan, Ayşegül Danış

Effect of intraoperative magnesium sulfate infusion on early postoperative opioid consumption in laparoscopic hysterectomy: A randomized, double-blind, controlled trial
İlke Tamdoğan, Dilek Yeniay, Fatma Alkan Bayburt, Tuna Albayrak, Azime Bulut

Effects of preemptive analgesia and preoperative anxiety on emergence delirium after pediatric adenotonsillectomy
Gamze Küçükosman, Bengü G. Köksal, Keziban Bollucuoğlu, Rahşan D. Okyay, Dilay İ. Alp, Özcan Pişkin, Hilal Ayoğlu

Relationship between demodex infestation and clinical parameters in polycystic ovary syndrome: A prospective observational study
Mehmet Bülbül, Mustafa Kaplanoğlu, Tuncay Çelik, İpek Çakılkaya, Emel Dündar

Clinical and radiological outcome comparison of unilateral vs. bilateral percutaneous vertebroplasty for osteoporotic vertebral compression fractures
Levent Aydın, Buse Sarıgül, Tufan Agah Kartum, Gonca Gül Öndüç, Saime Ayça Şahin, Ali Fatih Ramazanoğlu

Olgu Sunumu / Case Report

Hemichorea-hemiballismus developing due to non-ketotic hyperglycemia: A rare case report
Merve İrem Atıcı, Baycan Kuş, Taner Şahin, İrem Kömürcü, Merva Tuna

Derleme / Review

Artificial intelligence in pain: A comprehensive review
İlhan Celil Özbek, Erkan Özduran, Volkan Hancı