FORBES JOURNAL OF MEDICINE

F O R B E S T I P D E R G i S i

Cilt / Volume: 6 Sayı / Issue: 3 Aralık / December 2025

www.forbestip.org
e-ISSN: 2757-5241

Sağlıklı Yarınlar için Bilimin İzindeyiz

FORBES
JOURNAL OF MEDICINE - TIP DERGİSİ

e-ISSN: 2757-5241

Aralık / December 2025
Cilt / Volume: 6
Sayı / Issue: 3

Forbes Tıp Dergisi, İzmir Buca Seyfi Demirsoy Eğitim ve Araştırma Hastanesi'nin yayın organıdır.

Dergi dört ayda bir yılda 3 sayı (Nisan, Ağustos, Aralık) olarak yayımlanan açık erişim, ücretsiz ve hakemli bir dergidir.

"Amaç ve Kapsam", "Yayın Etiği", "Yazarlar için Kontrol Listesi" ve "Makale Hazırlama" hakkında daha fazla bilgi için lütfen derginin web sayfasını **(https://forbestip.org)** ziyaret edin.

Derginin editöryal ve yayın süreçleri ile etik kuralları ICMJE, WAME, CSE, COPE, EASE ve NISO gibi uluslararası kuruluşların kurallarına uygun olarak şekillenmektedir.

Forbes Tıp Dergisi, **Ulakbim TR Dizin, EmCare, EmBiology, EMBASE, EBSCO, Türk Medline, J-Gate, Türkiye Atıf Dizini, Gale** ve **DOAJ** tarafından indekslenmektedir.

Dergi elektronik ortamda yayımlanmaktadır.

Kurucu: Buca Seyfi Demirsoy Eğitim ve Araştırma Hastanesi Başhekimi Prof. Dr. M. Yekta Öncel

Baş Editör: Prof. Dr. M. Yekta Öncel

Forbes Journal of Medicine is the publication of İzmir Buca Seyfi Demirsoy Training and Research Hospital.

It is published three times a year (April, August, December) FJM is an open access, free and peer-reviewed Journal

Please refer to the journal's webpage **(https://forbestip.org)** for "Aims and Scope", "Ethical Policy", "Checklist for Authors" and "Manuscript Preparation".

The editorial and publication process of Forbes Journal of Medicine are shaped in accordance with the guidelines of the ICMJE, WAME, CSE, COPE, EASE, and NISO.

Forbes Journal of Medicine is indexed in **Ulakbim TR Dizin, Turkey Citation Index, EmCare, EmBiology, EMBASE, EBSCO, Türk Medline, J-Gate, Gale** and **DOAJ**.

The journal is published electronically.

Owner: M. Yekta Öncel, M.D. Prof. on Behalf of Buca Seyfi Demirsoy Training and Research Hospital Head Physician

Editor in Chief: M. Yekta Öncel, M.D. Prof.

Yönetim Yeri/Administrative Office
İzmir Buca Seyfi Demirsoy EAH, Başhekimlik, Buca/İzmir
Tel: +90 (232) 452 52 52 /4293 **E-mail:** forbestipdergisi@gmail.com

Yayın Türü/Publication Type
Yaygın Süreli/Scholarly Periodical

Yayınevi/Publisher
GALENOS YAYINEVİ
Molla Gürani Mah. Kaçamak Sk. No: 21/1 34093 Fındıkzade, Fatih, İstanbul, Türkiye
Tel: +90 (530) 177 30 97 **E-mail:** info@galenos.com.tr/yayin@galenos.com.tr **Web:** www.galenos.com.tr **Yayıncı Sertifika No:** 14521

Yayınlanma Tarihi/Online Publication Date
Aralık 2025/December 2025

FORBES
JOURNAL OF MEDICINE - TIP DERGİSİ

Yayın Kurulu

FORBES
JOURNAL OF MEDICINE - TIP DERGİSİ

Editorial Board

FORBES
JOURNAL OF MEDICINE - TIP DERGİSİ

Editöryal Danışma Kurulu

FORBES
JOURNAL OF MEDICINE - TIP DERGİSİ

Editorial Advisory Board

FORBES
JOURNAL OF MEDICINE - TIP DERGİSİ

İçindekiler/Contents

Derleme / Review

The Dual Faces of Vesicant Agents: From Chemical Weapon to Chemotherapeutic Agent
Vezikan Ajanların İki Yüzü: Kimyasal Silahtan Kemoterapötik Ajanlara
Ahu PAKDEMİRLİ **200**

Orijinal Araştırmalar / Original Articles

Metastatic Lymph Node Ratio as a Prognostic Factor in Gastric Adenocarcinoma
Mide Adenokarsinomunda Prognostik Faktör: Metastatik Lenf Oranı
Can UÇ, Tufan GÜMÜŞ, Volkan SAYUR, Recep TEMEL, Başak DOĞANAVŞARGİL, Taylan Özgür SEZER, Özgür FIRAT, Sinan ERSİN **207**

Toplum Kökenli Pnömonili Çocuklarda Yoğun Bakım Gereksinimini Öngörmede Hemogram Türevli Enflamatuvar İndekslerin Tanısal Değeri
Diagnostic Value of Hemogram-Derived Inflammatory Indices in Predicting Intensive Care Requirement in Children with Community-Acquired Pneumonia
Leyla ŞERO, Muhammet ASENA **213**

Medial Patellofemoral Ligament Reconstruction with and without Tibial Tubercle Osteotomy in Patellofemoral Instability Surgery: Clinical Outcomes
Patellofemoral İnstabilite Cerrahisinde Tibial Tüberkül Osteotomisi ile ve Osteotomisiz Medial Patellofemoral Ligament Rekonstrüksiyonu: Klinik Sonuçlar
Yılmaz ÖNDER, Mahmut TUNÇEZ, Tuğrul BULUT **221**

Anti-Müllerian Hormone and Insulin-Like Peptide 3 Levels in Adolescent Polycystic Ovary Syndrome: Correlation with Androgen Levels and Ultrasonographic Features
Polikistik Over Sendromlu Adolesanlarda Anti-Müllerian Hormon ve İnsülin Like Peptit Düzeyleri: Androjen Düzeyleri ve Ultrasonografik Bulgularla İlişkisi
Ayça KÖMÜRLÜOĞLU TAN, E. Nazlı GÖNÇ, Z. Alev ÖZÖN, Nurgün KANDEMİR, H. Nursun ÖZCAN, Ayfer ALİKAŞİFOĞLU **227**

Artificial Intelligence Generated Questions in Medical Education: How Prompt Design in Different Chatbots Shapes Assessment in Obstetrics and Gynecology?
Tıp Eğitiminde Yapay Zeka Tarafından Oluşturulan Sorular: Farklı Chatbotlarda Prompt Tasarımı Kadın Hastalıkları ve Doğum Alanında Değerlendirmeyi Nasıl Şekillendiriyor?
Zuhal YAPICI COŞKUN, Yavuz Selim KIYAK, Özlem COŞKUN, Mehmet KOCA, İrem BUDAKOĞLU, Özhan ÖZDEMİR **237**

Türkiye'de Sağlık Çalışanlarında Beyin Göçü
Brain Drain Among Health Workers in Türkiye
Orhan ÇAKIR, Metehan AKKOÇ, Çağrı Işıl ÇAKIR, Ayşe Şimal GENÇER **242**

Effects of White Noise and Swaddling on Pain and Physiological Parameters During Eye Examination in Healthy Term Infants: A Randomized Controlled Trial
Sağlıklı Term Bebeklerde Göz Muayenesi Sırasında Beyaz Gürültü ve Kundaklamanın Ağrı ve Fizyolojik Parametreler Üzerindeki Etkileri: Randomize Kontrollü Bir Çalışma
Sultan GÜNER BAŞARA, Hüsniye ÇALIŞIR, Emel BAHADIR ARSU, Özdemir ÖZDEMİR **249**

FORBES
JOURNAL OF MEDICINE - TIP DERGİSİ

İçindekiler/Contents

Comparison of Nurses' Digital Literacy Levels: A Study in Digital and Non-Digital Hospitals
Hemşirelerin Dijital Okuryazarlık Düzeylerinin Karşılaştırılması: Dijital ve Dijital Olmayan Hastanelerde Bir Çalışma
Mustafa ORHAN, Süleyman MERTOĞLU **259**

The Relationship of Clinical and Molecular Factors with Prognosis in Metastatic Colorectal Cancer
Metastatik Kolorektal Kanserde Klinik ve Moleküler Faktörlerin Prognoz ile İlişkisi
Evrim ATACA, Merve KESKİNKILIÇ, Şulen SARIOĞLU, Yasemin BAŞPINAR, Tuğba YAVUZŞEN **266**

The Role of Diffusion Tensor Magnetic Resonance Imaging in Understanding Neurological and Ocular Outcomes in Preterm Infants with Periventricular Leukomalacia
Periventriküler Lökomalazili prematüre infantlarda nörolojik ve oküler sonuçların değerlendirilmesinde Diffüzyon Tensor Manyetik Rezonans Görüntülemenin rolü
Yıldız EKEMEN KELEŞ, Elif DEMİRKILINÇ BİLER, Filiz AFRASHİ, Mehmet Cem ÇALLI, Mehmet YALAZ4, Sema TANRIVERDİ, Sanem YILMAZ, Gül AKTAN, Sarenur GÖKBEN **273**

Retrospective Evaluation of the Relationship Between Serum Inflammatory Markers and the Efficacy of Methotrexate in the Treatment of Ectopic Pregnancy
Serum Enflamasyon Belirteçleri ile Ektopik Gebelik Tedavisinde Kullanılan Metotreksat'ın Etkinlik İlişkisinin Retrospektif Olarak Değerlendirilmesi
Aslı AKDÖNER, Onur YAVUZ, Alper MANKAN, Elif Aleyna KÜÇÜK, Erkan ÇAĞLIYAN **283**

Konjenital Diyafragma Hernisinde Mortalite Risk Faktörleri: Tek Merkez Deneyimi
Mortality Risk Factors in Congenital Diaphragmatic Hernia: A Single Centre Experience
Duygu TUNÇEL, Leyla ŞERO, Suat ÇAL, Nilüfer OKUR **293**

Olgu Sunum / Case Report

Horizontal Gaze Palsy with Progressive Scoliosis (HGPPS) in a Consanguineous Family: A Case Report
Akraba Evliliği Olan Bir Ailede Horizontal Bakış Paralizisi ve Progresif Skolyoz: Bir Olgu Sunumu
Dilara Ece TOPRAK, Gözde Merve TEKEL, Yusuf Can DOĞAN, Esra IŞIK, Atilla Hikmet ÇİLENGİR **299**

Dizin / Index

2025 Hakem Dizini / *2025 Referee Index*

2025 Yazar Dizini / *2025 Author Index*

2025 Konu Dizini / *2025 Subject Index*

DOI: 10.4274/forbes.galenos.2025.94547
Forbes J Med 2025;6(3):200-206

Review / Derleme

The Dual Faces of Vesicant Agents: From Chemical Weapon to Chemotherapeutic Agent

Vezikan Ajanların İki Yüzü: Kimyasal Silahtan Kemoterapötik Ajanlara

Ahu PAKDEMİRLİ

University of Health Sciences, İzmir Faculty of Medicine, Department of Physiology, İzmir, Türkiye

Cite as: Pakdemirli A. The dual faces of vesicant agents: from chemical weapon to chemotherapeutic agent. 2025;6(3):200-206

ABSTRACT

Vesicant agents, particularly analogs of sulfur mustard (SM) and nitrogen mustard (NM), exhibit a striking dichotomy between devastating chemical toxicity and therapeutic potential in oncology. The cytotoxic facet of SM causes rapid cellular damage through bifunctional alkylation, depletion of glutathione and ATP, and PARP hyperactivation. This process manifests as oncosis or necrosis, acute skin vesiculation, eye damage, and respiratory tract injuries. Over decades, genotoxicity mediated by persistent DNA interstrand cross-links, oxidative stress, chronic inflammation, and field cancerization increases the risk of squamous cell carcinoma of the lung and skin. Epidemiological studies, particularly among Iran-Iraq War gas survivors, have confirmed these long-term cancer risks. Mechanistic studies indicate that reactive oxygen species-mediated mutagenesis, stem cell niche disruption, and clonal proliferation play central roles in vesicant-induced carcinogenesis. Paradoxically, these same alkylating mechanisms fueled the therapeutic development of NMs, formed the basis of modern chemotherapy, and raised concerns about secondary malignancies and biosafety. Future research should focus on early-detection biomarkers, PARP and receptor-interacting protein kinase 1 inhibitors that reduce acute toxicity, and advanced carrier systems that minimize genomic damage while maintaining antitumor efficacy. The history of vesicants demonstrates how understanding the mechanisms of a toxic chemical can transform it into a therapeutic tool and guide the evolution of cancer treatment. This review comprised a comprehensive examination of English- and Turkish-language literature, using PubMed-indexed keywords, including 'vesicant agents', 'SM', 'NM', 'chemotherapy', and 'carcinogenesis'.

Keywords: Vesicant agents, sulfur mustard, nitrogen mustard, chemotherapy, carcinogenesis, alkylating agent, chemical warfarev

ÖZ

Vesikan ajanlar, özellikle sülfür mustard (SM) ve nitrojen mustard (NM) analogları, yıkıcı kimyasal toksisite ile onkolojide terapötik potansiyel arasında çarpıcı bir ikilik sergiler. SM'in sitotoksik yüzü, bifonksiyonel alkilleme, glutatyon/ATP tükenmesi ve PARP hiperaktivasyonu yoluyla hızlı hücresel hasara yol açar. Bu süreç onkozis/nekroz, akut deri vezikülasyonu, göz hasarı ve solunum yolu yaralanmaları ile kendini gösterir. On yıllar içinde, genotoksik yüz, kalıcı DNA zincirler arası çapraz bağlar, oksidatif stres, kronik inflamasyon ve alan kanserleşmesi yoluyla ortaya çıkarak akciğer ve deri skuamöz hücreli kanser riskini artırır. Özellikle İran-Irak Savaşı gaz mağdurlarına yapılan epidemiyolojik çalışmalar bu uzun vadeli kanser risklerini doğrulamıştır. Mekanistik araştırmalar, reaktif oksijen türleri aracılı mutagenez, kök hücre nişi bozulması ve klonal çoğalmanın vesikan kaynaklı karsinogenezde merkezi rol oynadığını göstermektedir. Çelişkili olarak, aynı alkilleyici mekanizmalar NM'ların terapötik gelişimini tetiklemiş, modern kemoterapinin temelini oluşturmuş ve ikincil malignite ile biyogüvenlik kaygılarını ortaya koymuştur. Gelecek araştırmalar erken teşhis biyobelirteçleri, PARP ve reseptör etkileşimli protein kinaz 1 inhibitörleri ile akut toksisitenin azaltılması ve antitümör etkinliği korurken genomik hasarı minimize eden ileri taşıyıcı sistemler üzerine odaklanmalıdır. Vesikan ajanların tarihi, toksik bir kimyasalın mekanizmalarının anlaşılmasıyla tedavi aracına dönüştürülebileceğini ve kanser tedavisinin evrimini

Received/Geliş: 01.11.2025
Accepted/Kabul: 28.11.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Ahu PAKDEMİRLİ, Prof., MD, PhD
University of Health Sciences, İzmir Faculty of Medicine, Department of Physiology, İzmir, Türkiye
ahu@pakdemirli.com
ORCID: 0000-0001-9224-3007

nasıl yönlendirebileceğini göstermektedir. Bu derleme, PubMed indeksli 'vezikan ajanlar', 'sülfür mustard', 'nitrojen mustard', 'kemoterapi' ve 'karsinogenez' gibi anahtar kelimelerin taranarak İngilizce ve Türkçe literatürün incelenmesiyle derlenmiştir.

Anahtar Kelimeler: Vesikan ajanlar, sülfür mustard, nitrojen mustard, kemoterapi, karsinogenez, alkilleyici ajan, kimyasal savaş

INTRODUCTION

This review examines the dual-use paradigm of vesicant agents-substances engineered for chemical warfare yet repurposed as chemotherapeutic tools-focusing on sulfur mustard (SM) and nitrogen mustard (NM) analogs. By synthesizing recent evidence, we highlight the molecular, clinical, and ethical dimensions of this duality, from acute toxicity to oncogenic risks and oncology applications, while addressing regulatory and biosecurity implications in the post-CW Convention era. This review was conducted through a comprehensive examination of English and Turkish literature, using PubMed-indexed keywords including 'vesicant agents', 'sulfur mustard', 'nitrogen mustard', 'chemotherapy' and 'carcinogenesis'.

The story of vesicant agents is a stark narrative of destruction and discovery. One of the best-known vesicant agents, SM, was first synthesized in the 19th century; its debut as a chemical weapon on the battlefields of Ypres during World War I heralded a new era of warfare, earning it the grim sobriquet "King of Battle Gases".[1,2] Its military efficacy lies not in high lethality, but in its ability to inflict debilitating, resource-intensive injuries that overwhelm medical systems and erode morale.[3] However, beyond its notorious history as a vesicant, SM has emerged as an indispensable tool in molecular toxicology, offering profound insights into cellular stress responses, death pathways, and carcinogenesis.[4] The central theme of SM's biology is its intrinsic duality. A single exposure can trigger two divergent pathological timelines. In the immediate cytotoxic phase, within hours, SM induces massive cellular injury, leading to necrotic cell death (oncosis) in highly exposed tissues, followed by an intense inflammatory response that characterizes its acute blistering and damaging effects on the skin, eyes, and respiratory tract.[5] The delayed genotoxic face: over decades, sublethal DNA damage can evade repair, leading to genomic instability, clonal expansion of mutated cells, and a significantly elevated risk of cancer, particularly in the lungs and skin.[6,7] Yet this architect of suffering catalyzed three scientific revolutions:

1. In molecular toxicology: SM revealed that bifunctional alkylation was the ur-mechanism of genotoxicity, a finding that predates modern DNA sequencing by 80 years.[4]
2. Cell death biology: Its dose-dependent activation of oncosis, apoptosis, necroptosis defined ATP thresholds for programmed vs. unprogrammed demise.[5]
3. Oncology: SM's genotoxic legacy [International Agency for Research on Cancer (IARC) Group 1] provided the first human model of inflammation-driven field cancerization.[6,7]

This review synthesizes current knowledge on both fronts. We will explore the chemistry defining its reactivity, dissect the cellular pathways it disrupts, and examine long-term epidemiological data linking exposure to cancer. Furthermore, we trace the remarkable journey of its chemical analogs from the battlefield to the clinic, highlighting how the same destructive principles were harnessed to create the first chemotherapeutic agents.

The Janus-Faced Timeline: Cytotoxic vs. Oncogenic Divergence

A single exposure to SM initiates a cascade of complex pathological events that unfold along two major, yet interrelated, trajectories: an acute cytotoxic phase and a delayed genotoxic phase. These two axes represent the immediate and long-term consequences of SM exposure, and together they define the compound's notorious duality as both a devastating chemical weapon and a model for studying chronic toxic injury.

In the acute phase, SM acts as a highly reactive alkylating agent targeting essential cellular macromolecules such as DNA, proteins, and low-molecular-weight antioxidants, including glutathione (GSH).[8] The rapid formation of alkylated adducts triggers widespread biochemical disruption: DNA strand breaks, oxidative stress, and enzymatic inhibition converge to induce cell death via necrosis, apoptosis, or, in some cases, mixed regulated cell-death pathways. Clinically, this manifests within hours as erythema and edema at the exposure site, followed by the formation of large, painful blisters and ulcerations, which are characteristic of vesicant injury (Table 1).[9] The eyes and respiratory tract are particularly vulnerable and often develop conjunctivitis, corneal erosions, bronchial inflammation, and necrosis.[10] Systemically, bone marrow suppression can lead to transient leukocytosis followed by lymphopenia and then pancytopenia, underscoring the hematopoietic toxicity of SM. The underlying cellular mechanisms are characterized by oxidative stress and rapid depletion of cellular

Table 1. A single SM exposure bifurcates into two orthogonal pathological axes

Phase	Timeline	Dominant mechanism and pathway	Key lesion	Clinical manifestation	Molecular signature
Acute cytotoxic	0-72 h	DNA/protein alkylation, oxidative stress, cell death (necrosis/ apoptosis)	ICLs (4%) + ROS (10x↑)	Blistering, inflammation, marrow suppression	HMGB1↑50x, IL-1β↑10x
Chronic oncogenic/ delayed genotoxicity	10-40 y	Persistent DNA damage, chronic inflammation, immune dysregulation	O^6-alkylG (40% mut)	Fibrosis, chronic lung/skin/eye disease, cancer	G:C→A:T (sig. mut)

ICL: Interstrand cross-link, ROS: Reactive oxygen specy

energy stores, which together precipitate mitochondrial dysfunction, tissue necrosis, and a vigorous inflammatory response that exacerbates tissue damage.[11]

The delayed genotoxic phase emerges long after the acute symptoms have subsided and represents the insidious, progressive dimension of SM toxicity. Sublethal doses that do not cause immediate cell death often leave behind unrepaired or misrepaired DNA lesions, setting the stage for persistent genomic instability and mutagenesis.[12] Over time, this molecular damage manifests clinically as chronic respiratory diseases-including bronchiolitis obliterans and pulmonary fibrosis-alongside skin atrophy, pigmentary changes, and ocular complications such as chronic keratitis and limbal stem cell deficiency. Epidemiological studies of exposed populations, particularly among Iranian war veterans, have also revealed a markedly increased incidence of malignancies, especially squamous cell carcinoma (SCC) of the skin and various forms of lung cancer. These chronic pathologies are perpetuated by ongoing inflammation, immune dysregulation, and the clonal expansion of mutated cells, which collectively drive progressive organ dysfunction and carcinogenesis.[13]

Thus, SM exposure represents a biphasic toxicological paradigm: a rapidly destructive acute insult followed by a protracted genotoxic aftermath. Understanding these intertwined processes not only elucidates the mechanisms of SM-induced injury but also provides a broader framework for studying inflammation-driven carcinogenesis and long-term sequelae of alkylating agent exposure.[10,13]

Chemical Identity and Biochemical Warfare at the Molecular Level

SM's potency is a direct consequence of its structure. As a lipophilic molecule, it readily penetrates cellular membranes. Its core mechanism involves intramolecular cyclization to form a highly reactive cyclic ethylene-sulfonium ion, a powerful electrophile that readily attacks nucleophilic sites of vital cellular components.[14,15]

Primary Molecular Targets

Among the molecular events triggered by SM, DNA alkylation and cross-link formation represent the most critical and damaging mechanisms. SM demonstrates a strong preference for alkylating the N7 position of guanine bases within DNA, producing monoadducts that distort the double helix and interfere with normal base pairing. More importantly, due to its bifunctional chemical structure, SM can generate interstrand cross-links (ICLs)-highly lethal DNA lesions that prevent the separation of complementary strands during replication and transcription, thereby halting essential cellular processes.[16,17] When the cell attempts to repair these ICLs through excision or recombination pathways, the process frequently results in double-strand breaks (DSBs), which are among the most mutagenic and clastogenic forms of DNA damage. These DSBs are a critical initiating factor for genomic instability, laying the molecular foundation for cytotoxicity and carcinogenesis.[18]

In addition to its genotoxicity, SM exerts broad inactivation of proteins and enzymes by alkylating key amino acid residues, such as cysteine, histidine, and glutamate. This modification alters the structural integrity and catalytic function of essential enzymes, many of which are central to energy metabolism and antioxidant defense. Notably, inhibition of glucose-6-phosphate dehydrogenase and other sulfhydryl-dependent enzymes disrupts the hexose monophosphate shunt, diminishing cellular NADPH levels and impairing the regeneration of reduced GSH. The consequent loss of redox homeostasis sensitizes the cell to oxidative injury and metabolic collapse.[19]

A pivotal early biochemical hallmark of SM exposure is the rapid depletion of GSH, the cell's primary antioxidant and detoxifying molecule.[20] As GSH conjugates with SM and its reactive intermediates, intracellular GSH stores are rapidly depleted, undermining the cell's ability to neutralize reactive oxygen species (ROS). The resulting oxidative stress amplifies cellular injury through lipid peroxidation, oxidation of nucleic acids and proteins, and further membrane destabilization. This self-propagating

cycle of oxidative and alkylation damage perpetuates mitochondrial dysfunction and DNA oxidation, intensifying both acute cytotoxicity and long-term genotoxic effects.[21]

Mechanisms of Acute Toxicity and Cellular Demise

The cellular response to SM is a cascade of metabolic failure, culminating in a dose-dependent decision between different modes of cell death.

The Metabolic Crisis

The hyperactivation of the DNA repair enzyme poly (ADP-ribose) polymerase (PARP) in response to excessive DNA damage consumes vast amounts of NAD+, leading to a catastrophic drop in ATP levels.[22] This energy crisis is compounded by mitochondrial dysfunction: SM and ROS induce the mitochondrial permeability transition, which collapses the proton gradient and halts ATP synthesis.[23]

The Spectrum of Cell Death

The mode of death-oncosis/necrosis, apoptosis, and necroptosis- is dictated by the severity of the insult and the remaining energy reserves (Table 2). In oncosis/necrosis, at high, militarily relevant doses, rapid ATP depletion prevents the energy-dependent execution of programmed cell death. Cells undergo oncosis-characterized by swelling, organelle disintegration, and plasma membrane rüptüre-leading to the release of damage-associated molecular patterns that fuel a robust and damaging inflammatory response.[5,24] At lower doses, where some ATP remains, cells can initiate apoptosis (programmed cell death). This involves caspase activation, chromatin condensation, and formation of apoptotic bodies for efficient clearance, minimizing inflammation.[25] Necroptosis, emerging evidence indicates SM can also trigger this programmed form of necrosis, via receptor-interacting protein kinase (RIPK) activation, further contributing to the inflammatory pathology.[26]

The Oncogenic Face: From DNA Adduct to Malignancy

The IARC classifies SM as a Group 1 agent, "carcinogenic to humans".[27] The journey from exposure to cancer is a multi-decade process driven by persistent genomic lesions.

Molecular Pathogenesis of Sulfur Mustard-Induced Carcinogenesis

Mutagenesis: while ICLs are cytotoxic, misrepaired monoadducts are mutagenic. The persistence of O6-alkylguanine adducts can lead to G:C to A:T transition mutations during DNA replication, a mutational signature common in carcinogenesis induced by alkylating agents.[28] Genomic instability: error-prone repair of SM-induced DSBs and the collapse of stalled replication forks at ICLs lead to chromosomal aberrations, micronuclei formation, and large-scale genomic rearrangements, providing a fertile ground for oncogene activation and tumor-suppressor loss.[29] Chronic inflammation and clonal selection: repeated cycles of tissue damage, ulceration, and repair in SM-exposed skin and respiratory epithelium create a pro-tumorigenic microenvironment. The presence of inflammatory cytokines, growth factors, and ROS promotes the proliferation and survival of initiated (mutated) cell clones.[30]

Site-Specific Cancers: Epidemiological and Clinical Evidence

Long-term clinical observations of individuals exposed to SM during the Iran-Iraq War have yielded the most conclusive human data linking this agent to cancer development. In particular, respiratory malignancies have emerged as a major delayed outcome of exposure. Epidemiological analyses have demonstrated that individuals exposed to SM exhibit a markedly increased incidence of lung cancer, estimated to be more than twice that observed in unexposed populations, with the risk rising proportionally with the extent and duration of exposure.[31] Histopathological evaluations of these cases frequently identify SCC and adenocarcinoma as the predominant tumor types, typically occurring in a background of chronic airway injury characterized by bronchitis, bronchiolitis obliterans, and pulmonary fibrosis—conditions that create a pro-carcinogenic microenvironment conducive to malignant transformation.[32]

Table 2. Key cell death pathways induced by sulfur mustard

Pathway	Characteristics	Energy (ATP) dependent?	Inflammatory outcome	Primary SM dose
Oncosis/necrosis	Cellular swelling, membrane rupture, release of intracellular contents	No	Highly inflammatory	High
Apoptosis	Cell shrinkage, nuclear fragmentation, formation of apoptotic bodies	Yes	Anti-inflammatory / tolerogenic	Low to moderate
Necroptosis	Regulated necrosis, RIPK1/RIPK3/ MLKL pathway activation	No	Highly inflammatory	Moderate to high

SM: Sulfur mustard, RIPK1: Receptor-interacting protein kinase 1, MLKL: Mixed lineage kinase domain-like pseudokinase

Similarly, cutaneous malignancies represent another critical long-term complication among SM survivors. Persistent or recurrent skin lesions at sites of previous vesication have been recognized as pre-neoplastic foci that may progress to SCC or, less commonly, basal cell carcinoma. A systematic review of affected populations revealed that the risk of developing such skin cancers is substantially elevated compared with that in non-exposed individuals.[33] The underlying mechanisms appear to involve a combination of field cancerization, resulting from widespread alkylation-induced DNA damage across large skin areas, and chronic inflammation associated with repeated cycles of ulceration and wound healing. Together, these processes create an environment that favors clonal expansion of mutated keratinocytes and the gradual emergence of malignant lesions.

The Therapeutic Legacy: From Mustard Gas to Chemotherapy

Paradoxically, the study of SM's systemic effects led to a medical revolution. The observation of profound leukopenia and lymphoid atrophy in victims of the 1943 Bari harbor incident prompted pharmacologists Louis S. Goodman and Alfred Gilman to investigate NMs as potential treatments for lymphoproliferative cancers.[34]

Nitrogen Mustards as Chemotherapeutic Agents

NMs, such as mechlorethamine, cyclophosphamide, and melphalan, share SM's core mechanism of action: forming DNA ICLs that are fatal to rapidly dividing cells. Their development marked the birth of modern cancer chemotherapy (Table 3).[35]

The Enduring Paradox in Therapy

The use of NM chemotherapy mirrors the dual nature of SM. While they are curative for many cancers, they are potent human carcinogens. Patients treated with alkylating agents have a markedly increased risk of developing secondary myeloid neoplasms and other cancers, often with a characteristic latency of 5-10 years.[36] This underscores the fundamental principle that the very mechanism that kills cancer cells-genomic disruption-can also initiate it, a risk that must be carefully managed in clinical practice.

CONCLUSION

Vesicant agents occupy a distinctive place in toxicology and oncology, embodying a dual nature that is simultaneously destructive and instructive. Haber's seminal 1986 analysis first articulated how these agents push the limits of cellular resilience, producing rapid energy collapse, overwhelming oxidative stress, and necrotic inflammatory cascades. These early mechanistic insights became the foundation for the conceptual framework that Szinicz further refined nearly two decades later, in 2005-namely, the continuity between acute cytotoxicity and delayed oncogenic outcomes following vesicant exposure.

The biphasic biological trajectory was originally delineated by the biochemical and histopathological studies of Papirmeister and colleagues, who documented the early acute phase characterized by bifunctional alkylation, depletion of GSH and ATP, PARP hyperactivation, and oncosis/necrosis. This acute phase manifests in severe dermal, ocular, and respiratory tissue injury. Subsequent mechanistic advances, especially those contributed by Kehe and Smith, clarified that unrepaired DNA ICLs and a persistently inflamed tissue microenvironment shape the delayed genotoxic phase—culminating in field cancerization, clonal selection, and a documented 2- to 8-fold elevation in SCC of the skin and lung in previously exposed populations.

More recent studies have converged on ROS-driven mutagenesis and stem-cell niche disruption as central molecular drivers of vesicant-induced carcinogenesis. The same alkylation chemistry responsible for battlefield injuries also forms the basis of therapeutic repurposing: as Szinicz highlighted, NMs served as the prototype for modern anticancer chemotherapy. Yet this scientific transformation—from chemical weapon to chemotherapeutic tool—raises continuing ethical and biosecurity concerns, including risks of secondary malignancies, especially in communities with historical exposure.

Future research priorities include the development of sensitive early-detection biomarkers for long-term cancer surveillance, targeted pharmacologic strategies such as PARP and RIPK1 inhibitors to mitigate acute tissue injury, and the design of safer alkylating chemotherapeutics. In

Table 3. Nitrogen mustards as chemotherapeutic agents

Agent	Structure	ICL Formation	FDA approval	Cancers
Mechlorethamine	HN_2 analog	2-5%/10^6 bp	1949	Lymphoma
Cyclophosphamide	Cyclic prodrug	1-3%/10^6 bp	1959	Breast, Ovarian
Melphalan	Phe analog	3-6%/10^6 bp	1964	Myeloma

ICL: Interstrand cross-links, FDA: Food and Drug Administration

particular, next-generation antibody–drug conjugates offer promise in preserving antitumor potency while minimizing off-target genomic toxicity.

Ultimately, the story of SM represents both a cautionary lesson and an example of scientific redemption. Understanding the mechanisms of a devastating vesicant has become a foundation for therapeutic advances, illustrating how the biology of a poison can guide the development of modern cancer therapies.

Footnotes

Conflict of Interest: Ahu Pakdemirli, the author of this manuscript, is the language editor of the Forbes Journal of Medicine. However, she was not involved in any stage of the review or publication decision process for this manuscript. The manuscript was independently evaluated by editors from other institutions.

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

REFERENCES

1. Fitzgerald GJ. Chemical warfare and medical response during World War I. Am J Public Health. 2008;98:611-625. Erratum in: Am J Public Health. 2008;98:1158.
2. Szinicz L. History of chemical and biological warfare agents. Toxicology. 2005;214:167-81.
3. Balali-Mood M, Hefazi M. The pharmacology, toxicology, and medical treatment of sulphur mustard poisoning. Fundam Clin Pharmacol. 2005;19:297-315.
4. Jan Y, Heck D, Laskin D, Laskin J. DNA damage signaling in the cellular responses to mustard vesicants. Toxicology letters. 2020.
5. Kehe K, Balszuweit F, Steinritz D, Thiermann H. Molecular toxicology of sulfur mustard-induced cutaneous inflammation and blistering. Toxicology. 2009;263:12-9.
6. Ghanei M, Harandi AA. Long term consequences from exposure to sulfur mustard: a review. Inhal Toxicol. 2007;19:451-6.
7. Khateri S, Ghanei M, Keshavarz S, Soroush M, Haines D. Incidence of lung, eye, and skin lesions as late complications in 34,000 Iranians with wartime exposure to mustard agent. J Occup Environ Med. 2003;45:1136-43.
8. Laskin JD, Black AT, Jan YH, et al. Oxidants and antioxidants in sulfur mustard-induced injury. Ann N Y Acad Sci. 2010;1203:92-100.
9. Sezigen S, Eyison RK, Ortatatli M, Kilic E, Kenar L. Myelosuppression and acute hematological complications of sulfur mustard exposure in victims of chemical terrorism. Toxicol Lett. 2020;318:92-8.
10. Balali-Mood M, Hefazi M. Comparison of early and late toxic effects of sulfur mustard in Iranian veterans. Basic Clin Pharmacol Toxicol. 2006;99:273-82.
11. Xiaoji Z, Xiao M, Rui X, et al. Mechanism underlying acute lung injury due to sulfur mustard exposure in rats. Toxicol Ind Health. 2016;32:1345-57.
12. Beigi Harchegani A, Khor A, Tahmasbpour E, Ghatrehsamani M, Bakhtiari Kaboutaraki H, Shahriary A. Role of oxidative stress and antioxidant therapy in acute and chronic phases of sulfur mustard injuries: a review. Cutan Ocul Toxicol. 2019;38:9-17.
13. Panahi Y, Roshandel D, Sadoughi MM, Ghanei M, Sahebkar A. Sulfur Mustard-Induced Ocular Injuries: Update on Mechanisms and Management. Curr Pharm Des. 2017;23:1589-97.
14. Munro NB, Talmage SS, Griffin GD, et al. The sources, fate, and toxicity of chemical warfare agent degradation products. Environ Health Perspect. 1999;107:933-74.
15. Noort D, Benschop HP, Black RM. Biomonitoring of exposure to chemical warfare agents: a review. Toxicol Appl Pharmacol. 2002;184:116-26.
16. Ludlum DB, Austin-Ritchie P, Hagopian M, Niu TQ, Yu D. Detection of sulfur mustard-induced DNA modifications. Chem Biol Interact. 1994;91:39-49.
17. Jost P, Svobodova H, Stetina R. Induction and repair of DNA cross-links induced by sulfur mustard in the A-549 cell line followed by a comet assay. Chem Biol Interact. 2015;237:31-37.
18. Deans AJ, West SC. DNA interstrand crosslink repair and cancer. Nat Rev Cancer. 2011;11:467-80.
19. Smith WJ, Dunn MA. Medical defense against blistering chemical warfare agents. Arch Dermatol. 1991;127:1207-13.
20. Layali I, Shahriary A, Rahmani Talatappe N, Tahmasbpour E, Rostami H, Beigi Harchegani A. Sulfur mustard triggers oxidative stress through glutathione depletion and altered expression of glutathione-related enzymes in human airways. Immunopharmacol Immunotoxicol. 2018;40:290-296.
21. Ramos E, Gil-Martín E, De Los Ríos C, et al. Melatonin as modulator for sulfur and nitrogen mustard-induced inflammation, oxidative stress and DNA damage: molecular therapeutics. Antioxidants (Basel). 2023;12:397.
22. Mol MA, van de Ruit AM, Kluivers AW. NAD+ levels and glucose uptake of cultured human epidermal cells exposed to sulfur mustard. Toxicol Appl Pharmacol. 1989;98:159-65.
23. Gould NS, White CW, Day BJ. A role for mitochondrial oxidative stress in sulfur mustard analog 2-chloroethyl ethyl sulfide-induced lung cell injury and antioxidant protection. J Pharmacol Exp Ther. 2009;328:732-9.
24. Rosenthal DS, Simbulan-Rosenthal CM, Iyer S, Spoonde A, Smith W, Ray R, Smulson ME. Sulfur mustard induces markers of terminal differentiation and apoptosis in keratinocytes via a Ca2+-calmodulin and caspase-dependent pathway. J Invest Dermatol. 1998;111:64-71.
25. Ghabili K, Agutter PS, Ghanei M, Ansarin K, Panahi Y, Shoja MM. Sulfur mustard toxicity: history, chemistry, pharmacokinetics, and pharmacodynamics. Crit Rev Toxicol. 2011;41:384-403.
26. Zhang B, Zhang Y. Necroptosis in sulfur mustard-induced toxicity. *J Appl Toxicol*. 2021;41(1):33–41.
27. Chappell G, Pogribny IP, Guyton KZ, Rusyn I. Epigenetic alterations induced by genotoxic occupational and environmental human chemical carcinogens: a systematic literature review. Mutat Res Rev Mutat Res. 2016;768:27-45.
28. Zubel T, Hochgesand S, John H, et al. A mass spectrometric platform for the quantitation of sulfur mustard-induced nucleic acid adducts as mechanistically relevant biomarkers of exposure. Arch Toxicol. 2019;93:61-79.
29. Panahi Y, Fattahi A, Zarei F. Ghasemzadeh N. Next-generation sequencing approaches for the study of genome and epigenome toxicity induced by sulfur mustard. Arch Toxicol 2018;92:3443-3457 .

30. Yu W, Tu Y, Long Z, et al. Reactive oxygen species bridge the gap between chronic inflammation and tumor development. Oxid Med Cell Longev. 2022;2022:2606928.
31. Zafarghandi MR, Soroush MR, Mahmoodi M, Naieni KH, Ardalan A, Dolatyari A, Falahati F, Mirmohammadkhani M, Mousavi B, Ghanei M. Incidence of cancer in Iranian sulfur mustard exposed veterans: a long-term follow-up cohort study. Cancer Causes Control. 2013;24:99-105.
32. Hosseini-khalili A, Haines DD, Modirian E, Soroush M, Khateri S, et al. Mustard gas exposure and carcinogenesis of lung. Mutat Res. 2009;678:1-6.
33. Owens M, Thyagarajan A, Travers JB, Sahu RP. Mechanistic insights and pharmacological approaches for nitrogen and sulfur mustards and their implications as therapeutic agents. J Appl Toxicol. 2025;45:1417-1425.
34. Gilman A. The initial clinical trial of nitrogen mustard. Am J Surg. 1963;105:574-8.
35. Chabner BA, Roberts TG Jr. Timeline: chemotherapy and the war on cancer. Nat Rev Cancer. 2005;5:65-72.
36. Pedersen-Bjergaard J, Andersen MK, Christiansen DH. Therapy-related acute myeloid leukemia and myelodysplasia after high-dose chemotherapy and autologous stem cell transplantation. Blood. 2000;95:3273-9.

DOI: 10.4274/forbes.galenos.2025.38233
Forbes J Med 2025;6(3):207-212

Original Article / Orijinal Araştırma

Metastatic Lymph Node Ratio as a Prognostic Factor in Gastric Adenocarcinoma

Mide Adenokarsinomunda Prognostik Faktör: Metastatik Lenf Oranı

Can UÇ[1], Tufan GÜMÜŞ[2], Volkan SAYUR[2], Recep TEMEL[2], Başak DOĞANAVŞARGİL[4], Taylan Özgür SEZER[2], Özgür FIRAT[3], Sinan ERSİN[3]

[1]İzmir Buca Seyfi Demirsoy Training and Research Hospital, Clinic of General Surgery, İzmir, Türkiye
[2]Ege University Hospital, Department of General Surgery, İzmir, Türkiye
[3]Acıbadem Kent Hospital, Clinic of General Surgery, İzmir, Türkiye
[4]Ege University Hospital, Department of Pathology, İzmir, Türkiye

Cite as: Uç C, Gümüş T, Sayur V, Temel R, Doğanavşargil B, Sezer TÖ, Fırat Ö, Ersin S. Metastatic lymph node ratio as a prognostic factor in gastric adenocarcinoma. Forbes J Med. 2025;6(3):207-212

ABSTRACT

Objective: Gastric cancer remains an important health problem. Although the classical staging system is the tumor, node, metastasis (TNM) system, stage migration phenomenon may occur depending on the number of lymph nodes removed. Therefore, although there are opinions that lymph node metastasis rate can be used as an alternative to TNM, a clear cut-off value has not been determined. In our study, we aimed to evaluate the survival in TNM staging according to the metastatic lymph node rate.

Methods: Patients operated for gastric adenocarcinoma between 2015 and 2022 were evaluated retrospectively. Patients were divided into quartiles according to the metastatic lymph node rate. Mean survival times were compared between the two groups according to TNM stages and metastatic lymph node rate.

Results: One hundred ninety-nine patients were included in the study. When survival was evaluated, it was observed that survival decreased with increasing disease stage and metastatic lymph node ratio. All patients with stage 1 and 2 were in the 1st quartile (25%) according to metastatic lymph node ratio. In the evaluation of stage 3 patients, it was seen that survival decreased as the metastatic lymph node ratio increased.

Conclusion: Our study revealed that metastatic lymph node ratio of 25% and above is an important parameter in predicting prognosis. Especially in stage 3 patients, metastatic lymph node ratio was shown to be a valuable parameter in predicting survival.

Keywords: Gastric adenocarcinoma, metastatic lymh node ratio, prognosis

ÖZ

Amaç: Mide kanseri önemli bir sağlık sorunu olmaya devam etmektedir. Klasik evreleme sistemi tumor, node, metastasis (TNM) sistemi olmasına rağmen, çıkarılan lenf nodu sayısına bağlı olarak evre kayması fenomeni ortaya çıkabilmektedir. Bu nedenle lenf nodu metastaz oranının TNM'ye alternatif olarak kullanılabileceği yönünde görüşler olmasına rağmen net bir kesme değeri belirlenmemiştir. Çalışmamızda metastatik lenf nodu oranına göre, evreler içerisinde sağkalımı değerlendirmeyi amaçladık.

Yöntem: 2015-2022 yılları arasında mide adenokarsinomu nedeniyle opere edilen hastalar retrospektif olarak değerlendirildi. Hastalar metastatik lenf nodu oranına göre çeyreklere ayrıldı. İki grup arasında ortalama sağ kalım süreleri TNM evrelerine ve metastatik lenf nodu oranına göre karşılaştırıldı.

Bulgular: Çalışmaya 199 hasta dahil edildi. Sağkalım değerlendirildiğinde hastalık evresi ve metastatik lenf nodu oranı arttıkça sağkalımın azaldığı görüldü. Evre 1 ve 2 olan tüm hastalar metastatik lenf nodu oranına göre 1. çeyrekte (%25) yer aldı. Evre 3 hastaların değerlendirilmesinde metastatik lenf nodu oranı arttıkça sağ kalımın azaldığı görüldü.

Received/Geliş: 21.06.2025
Accepted/Kabul: 08.08.2025
Epub: 03.09.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Can UÇ, MD,
İzmir Buca Seyfi Demirsoy Training and Research Hospital, Clinic of General Surgery, İzmir, Türkiye
✉ drcanuc@gmail.com
ORCID: 0000-0002-9641-0156

Sonuç: Çalışmamız %25 ve üzeri metastatik lenf nodu oranının prognozu tahmin etmede önemli bir parametre olduğunu ortaya koydu. Özellikle evre 3 hastalarda metastatik lenf nodu oranının sağ kalımı tahmin etmede değerli bir parametre olduğu gösterildi.

Anahtar Kelimeler: Mide kanseri, metastatik lenf nodu oranı, prognoz

INTRODUCTION

Gastric cancer continues to be a major health problem worldwide. According to GLOBOCAN data, gastric cancer is the fifth most common malignancy and the third most common cause of cancer-related death.[1] Curative surgical resection is the mainstay of gastric cancer management. Many factors affect survival after surgery, including tumor type, invasion of the gastric wall, and lymph node metastases.[2]

Evaluation of prognostic factors is critical for disease management. Classical staging is based on tumor, node, metastasis (TNM) staging. TNM staging evaluates the infiltration of the tumor in the stomach wall, the number of pathological lymph nodes, and the presence of distant metastasis. Although the 8th edition of the American Joint Committee on Cancer (AJCC) is superior to the 7th edition, it has some limitations.[3] At least 16 lymph nodes are required for effective staging. Lymph node metastasis is associated with poor prognosis.[4]

More extensive lymph node dissection in Japan compared to more limited lymph node dissection in Western countries may lead to differences in the staging of patients. In this situation, known as "stage migration phenomenon", the survival times of patients may vary depending on the stage. The stage migration phenomenon occurs in 10-25% of cases, and is one of the major limitations of TNM staging.[3-5]

To overcome this limitation of TNM staging, the use of metastatic lymph node ratio, or logarithmic metastatic lymph node ratio, has been proposed. The presence of metastatic lymph nodes is one of the most important prognostic parameters for predicting recurrence and survival. Although there are studies suggesting that metastatic lymph node ratio is superior to classical TNM staging, a common cut-off value has not been determined.[6-9] There is still a need for studies showing the prognostic value of metastatic lymph node ratio in the literature. The aim of our study was to evaluate the survival times within stages according to the metastatic lymph node ratio.

METHODS

Patient Selection Criteria

Between 2015 and 2022, patients operated with a prediagnosis of gastric cancer were retrospectively evaluated. Patients with non-adenocarcinoma pathology, neoadjuvant systemic therapy, perioperative peritoneal or liver metastasis, preoperative distant metastasis, and fewer than 16 lymph nodes in the pathology examination were excluded. Patients over 18 years old with pathological results of gastric adenocarcinoma and 16 or more lymph nodes were included in the study. Preoperative distant metastasis assessment was performed by whole-body computed tomography or fluorodeoxyglucose positron emission tomography scan.

Metastatic Lymph Node Ratio

Metastatic lymph node ratio was defined as the ratio of metastatic lymph nodes to the total number of resected lymph nodes.

Metastatic lymph node ratio: (number of metastatic lymph nodes/total number of resected lymph nodes).

According to the metastatic lymph node ratio, patients were divided into four groups: 0-24% (first quartile), 25-49% (second quartile), 50-74% (third quartile), and 75-100% (fourth quartile).

Tumor, Node, Metastasis Staging

TNM staging of the patients was performed according to the AJCC 8th edition. Patients were grouped as stage 1, stage 2, and stage 3.

Statistical Analysis

Retrospectively obtained data were analyzed using the Statistical Package for Social Sciences (SPSS, Version 26.0. Armonk, NY: IBM Corp.). Demographic data were analyzed by descriptive statistical methods. Patients were grouped according to their stage and metastatic lymph node ratio. Survival of both groups was compared. Survival analysis was evaluated using the Kaplan-Meier test. Cox regression analysis was used to evaluate the factors affecting survival. Survival analysis was conducted using stage 3 patients who were defined as locally advanced according to the metastatic lymph node ratio using the same statistical methods.

Ethical Statement and Informed Consent

This study was approved by the Ege University Medical Research Ethics Committee (decision number: 23-10T/33, date: 05.10.2023) and written informed consent was obtained from all participants.

RESULTS

A total of 199 patients [77 females (38.7%) and 122 males (61.3%)] were included in the study. The mean age of the patients was 65.81 (±10.72) years. One hundred thirty-four patients (67.3%) underwent total gastrectomy, 34 (17.1%) distal gastrectomy, and 31 (15.6%) proximal gastrectomy. The mean number of total resected lymph nodes was 27.04 (±10.57), and the mean number of metastatic lymph nodes was 7.72 (±9.47). Pathologic features are shown in Table 1.

During the follow-up period, 19 patients (9.5%) developed local recurrence, while 51 patients (25.6%) developed distant organ metastasis. One hundred patients (50.3%) died during follow-up. The mean disease-free survival was 31.93±24.19 months, while overall survival was 55.29±24.30 months. The evaluation of the patients according to the (AJCC) 8th edition and metastatic lymph node ratio is shown in Tables 2 and 3.

When evaluated, it was observed that the survival time shortened with the advancement of the disease stage and the metastatic lymph node rate. According to AJCC, all patients in stages 1 and 2, were in the first quartile for metastatic lymph node ratio. When stage 3 patients were evaluated individually, it was seen that survival decreased as the metastatic lymph node ratio increased (p: 0.001). (p: 0.001). Survival times according to stages and metastatic lymph node ratios are shown in Tables 4 and 5, with stage 3 patients' survival times according to metastatic lymph node ratio in Table 6. When the factors affecting survival were evaluated, it was observed that distant organ metastasis was associated with survival (p<0.001).

Table 1. Pathological features

Invasion	Number of patients (n)	%
Lymphovascular invasion	108	54.3
Perineural invasion	125	62.8
Venous invasion	58	29.1
Subtype		
Poorly cohesive	47	23.6
Tubuler	49	24.6
Mixed	38	19.1
Poorly differentiated	53	26.6
Mucinous	7	3.5
Signet cell	5	2.5

Table 2. Staging according to AJCC 8th edition

AJCC staging	Number of patients	%
1A	20	9.9
1B	18	9.1
2A	22	11.1
2B	22	11.1
3A	27	13.5
3B	27	13.5
3C	63	31.8

AJCC: American Joint Committee on Cancer

DISCUSSION

Lymph node dissection in gastric cancer surgery reduces the risk of local recurrence and provides more accurate staging. Determining the patient's stage is an important step in individualized treatment planning. The metastatic lymph node ratio can be used to predict survival and recurrence, and to determine the treatment protocol.[7]

There are studies advocating the superiority of metastatic lymph node ratio over the classical TNM classification. Hou et al.[6] showed that the metastatic lymph node ratio can be used as a prognostic marker even if fewer than 15 lymph nodes are removed. Deng et al.[10] also compared the metastatic lymph node ratio with TNM staging and showed that the metastatic lymph node ratio was an effective factor in survival and significantly associated with recurrence. Similarly, in a study conducted in Italy involving 463 patients, metastatic lymph node ratio was shown to be

Table 3. Quartiles according to metastatic lymph node ratio

Metastatic lymph node ratio quarters	Number of patients	%
1st quartile (0-24%)	110	55.3
2nd quartile (25-49%)	41	20.6
3rd quartile (50-74%)	29	14.6
4th quartile (75-100%)	19	9.5

Table 4. Mean survival time according to AJCC staging (p<0.001)

Staging	Mean survival (month)	Std. Error	95% CI	
			Min.	Max.
1A	84.3	3.60	77.23	91.36
1B	74.34	7.20	60.22	88.45
2A	70.64	7.24	56.45	84.83
2B	58.06	8.38	41.63	74.48
3A	58.59	6.86	45.13	72.04
3B	39.50	7.56	24.68	54.33
3C	33.14	3.79	25.71	40.58
Mean	55.29	3.79	49.46	61.13

AJCC: American Joint Committee on Cancer, CI: Confidence interval, Min.: Minimum, Max.: Maximum

a better prognostic marker than standard staging.[11] Ergenç et al.[9] also found that metastatic lymph node ratio was more sensitive in predicting survival, especially in stage 3 disease. Our study supports these studies. Metastatic lymph node ratio was found to be an effective parameter in predicting survival. When stage 3 patients were evaluated by stage, it was observed that survival decreased as the metastatic lymph node ratio increased. Metastatic lymph node ratio can be used as a valuable parameter, especially in the subgroup analysis of stage 3 patients. Gulmez et al.[12] obtained similar results to our study; they showed that metastatic lymph node ratio was superior to TNM staging for overall survival in stage 3 disease, and discussed that it should be included in the treatment algorithm.

It may not always be possible to remove at least 16 lymph nodes required for TNM staging. In this case, the proportion of metastatic lymph nodes can be used as a predictor.[6,8] In the study of the Spanish group, a metastatic lymph node was found to be a prognostic factor, and compared to other studies, a metastatic lymph node rate of more than 25% was shown to be an important parameter, especially in cases where less than 16 lymph nodes were removed.[8]

The metastatic lymph node rate is not accurate in all cases. For example, if two lymph nodes are metastatic in patients with four lymph nodes dissected, the patients may fall into a lymph node metastasis rate group of 50%, which may cause patients to receive more treatment than necessary. In patients with inadequate lymph node dissection, the metastatic lymph node rate can also yield inaccurate results. Although the Italian study group proposed a logarithmic lymph node ratio to solve this problem, the same limitation applies to this method.[3] In contrast to these studies, we believe that at least 16 lymph nodes should be examined in the evaluation of metastatic lymph node ratio, as in TNM staging, and for this reason, patients with less than 16 lymph nodes were excluded in our study. In patients with less than 16 lymph nodes, the minimum number of lymph nodes should be clarified in future studies so that the metastatic lymph node ratio can be accepted as a prognostic factor. Extensive lymph node dissection during surgery is necessary for accurate staging and to reduce local recurrence.[7]

In patients who have undergone adequate surgical resection, the metastatic lymph node ratio will be useful in the choice of treatment; however, there is no consensus in the literature regarding its threshold values. Various studies have used different thresholds. In our study, the threshold value that increased mortality was 25%. Marchet et al.[13] used N0 (0%), N1 (1%-9%), N2 (10%-25%), N3 (>25%). Persiani et al.[14] and Zhao et al.[15], N0 (0%), N1 (1%-15%), N2 (16%-

Table 5. Mean survival time according to metastatic lymph node ratio (p<0.001)

Metastatic lymph node ratio	Mean survival (month)	Std. Error	95% CI	
			Min.	Max.
1st quartile (0-24%)	72.26	3.95	64.50	80.01
2nd quartile (25-49%)	44.82	5.62	33.79	55.85
3rd quartile (50-74%)	33.74	5.70	22.56	44.93
4th quartile (75-100%)	22.66	4.78	13.28	32.03
Mean	55.29	2.97	49.46	61.13

CI: Confidence interval, Min.: Minimum, Max.: Maximum

Table 6. Stage 3 patients' mean survival time according to metastatic lymph node ratio

Stage		Mean survival (month)	Std. Error	95% CI		p value
				Min.	Max.	
3	1st quartile (0-24%)	57.96	7.54	43.19	72.75	0.001
	2nd quartile (25-49%)	44.82	5.63	33.79	55.86	
	3rd quartile (50-74%)	33.75	5.71	22.56	44.93	
	4th quartile (75-100%)	22.66	4.78	13.28	32.03	
Total		41.52	3.33	34.98	48.06	

CI: Confidence interval, Min.: Minimum, Max.: Maximum

40%), N3 (>40%). Zhang et al.[16] used N0 (0%), N1 (1%-25%), N2 (26%-50%), N3 (>50%), similar to our study. Pedrazzani et al.[17] used 25% percentiles similarly. In all of the studies, the prognosis worsens as the metastatic lymph node ratio increases. Indeed, no clear threshold value has been found in the literature, and further studies on this subject are still needed. In our study, a metastatic lymph node rate of 25% was associated with an aggressive prognosis. We advocate that 25% should be accepted as the threshold value when evaluating metastatic lymph node rate.

Study Limitations

Adjuvant treatment protocols were not evaluated in our study. This is one of the limitations of the study, however, since we evaluated survival in the same patient group, adjuvant treatments of the patients do not affect the results of the study. Hwang et al.[18] argued that the metastatic lymph node ratio can be used to determine the chemotherapy protocol. In their study, they suggested that the prognosis of patients with a metastatic lymph node ratio of 10% or more was worse, especially in the stage 3 patient group, and that platinum-based chemotherapy protocols should be given to those patients who also had lymphovascular invasion. In stage 3 disease, the metastatic lymph node ratio may guide the selection of the chemotherapy protocol by evaluating it together with other risk factors; however, further studies on this subject are needed.

In our study, patients with neoadjuvant systemic treatment were also excluded. Indeed, metastatic lymph node ratio continues to be an important prognostic parameter after neoadjuvant systemic therapy. In the study by Jiang et al.[19], the 3-year survival of patients with a metastatic lymph node ratio of 30% or less after neoadjuvant chemotherapy was 81.9%, while the 3-year survival of patients with a metastatic lymph node ratio of more than 30% decreased to 18.5%. The metastatic lymph node ratio was also found to be effective in predicting disease-free survival.

The retrospective design is one of the main limitations; however, this limitation is offset by the fact that it was conducted on a large group of patients. An important factor that increases the reliability of our study is that the data were obtained from hospital records in a complete and objective manner.

CONCLUSION

In patients with adequate lymph node dissection, a metastatic lymph node rate above 25% is an important prognostic factor. When additional evaluation is performed in stage 3 patients, especially, themetastatic lymph node ratio can be used to determine the treatment protocol.

Ethics

Ethics Committee Approval: This study was approved by the Ege University Medical Research Ethics Committee (decision number: 23-10T/33, date: 05.10.2023).

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

Footnotes

Authorship Contributions

Surgical and Medical Practices: B.D., T.Ö.S., Ö.F., S.E., Concept: C.U., T.G., B.D., S.E., Design: C.U., T.G., S.E., Data Collection or Processing: C.U., R.T., Analysis or Interpretation: C.U., V.S., R.T., Literature Search: C.U., V.S., Writing: C.U., R.T., S.E.

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

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

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71:209-49.
2. Hu HM, Tsai HJ, Ku HY, et al. Survival outcomes of management in metastatic gastric adenocarcinoma patients. Sci Rep. 2021;11:23142.
3. Agnes A, Biondi A, Cananzi FM, et al. Ratio-based staging systems are better than the 7th and 8th editions of the TNM in stratifying the prognosis of gastric cancer patients: a multicenter retrospective study. J Surg Oncol. 2019;119:948-57.
4. Son T, Sun J, Choi S, et al. Multi-institutional validation of the 8th AJCC TNM staging system for gastric cancer: analysis of survival data from high-volume Eastern centers and the SEER database. J Surg Oncol. 2019;120:676-84.
5. Degiuli M, De Manzoni G, Di Leo A, et al. Gastric cancer: current status of lymph node dissection. World J Gastroenterol. 2016;22:2875-93.
6. Hou Y, Wang X, Chen J. Prognostic significance of metastatic lymph node ratio: the lymph node ratio could be a prognostic indicator for patients with gastric cancer. World J Surg Oncol. 2018;16:198.
7. Anonymous. Japanese gastric cancer treatment guidelines 2018 (5th edition). Gastric Cancer 2021;24:1-21.
8. Díaz Del Arco C, Estrada Muñoz L, Sánchez Pernaute A, et al. Towards standardization of lymph-node ratio classifications: validation and comparison of different lymph node ratio classifications for predicting prognosis of patients with resected gastric cancer. Ann Diagn Pathol. 2021;52:151738.
9. Ergenç M, Uprak TK, Akın Mİ, et al. Prognostic significance of metastatic lymph node ratio in gastric cancer: a Western-center analysis. BMC Surg. 2023;23:220.
10. Deng JY, Liang H. Clinical significance of lymph node metastasis in gastric cancer. World J Gastroenterol. 2014;20:3967-75.
11. Giuffrida M, Viani L, Iapichino GG, et al. The metastatic lymph node ratio is a better prognostic factor than the number of

metastatic limph node after curative resection for gastric cancer. Acta Biomed. 2021;92:e2021284.

12. Gulmez S, Senger AS, Uzun O, et al. Prognostic significance of the metastatic lymph node ratio compared to the TNM classification in stage III gastric cancer. Niger J Clin Pract. 2021;24:1602-08.
13. Marchet A, Mocellin S, Ambrosi A, et al. Italian Research Group for Gastric Cancer Study (GIRCG). The prognostic value of N-ratio in patients with gastric cancer: validation in a large, multicenter series. Eur J Surg Oncol. 2008;34:159-65.
14. Persiani R, Rausei S, Antonacci V, et al. Metastatic lymph node ratio: a new staging system for gastric cancer. World J Surg 2009;33(10):2106-11
15. Zhao LY, Li CC, Jia LY, et al. Superiority of lymph node ratio-based staging system for prognostic prediction in 2575 patients with gastric cancer: validation analysis in a large single center. Oncotarget. 2016;7:51069-81.
16. Zhang M, Wang J, Shi W, et al. Prognostic significance of metastatic lymph nodes ratio in patients with gastric adenocarcinoma after curative gastrectomy. Chin Med J (Engl). 2014;127:1874-8.
17. Pedrazzani C, Sivins A, Ancans G, et al. Ratio between metastatic and examined lymph nodes (N ratio) may have low clinical utility in gastric cancer patients treated by limited lymphadenectomy: results from a single-center experience of 526 patients. World J Surg. 2010;34:85-91.
18. Hwang JE, Kim H, Shim HJ, et al. Lymph-node ratio is an important clinical determinant for selecting the appropriate adjuvant chemotherapy regimen for curative D2-resected gastric cancer. J Cancer Res Clin Oncol. 2019;145:2157-66.
19. Jiang Q, Zeng X, Zhang C, et al. Lymph node ratio is a prospective prognostic indicator for locally advanced gastric cancer patients after neoadjuvant chemotherapy. World J Surg Oncol. 2022;20:261.

DOI: 10.4274/forbes.galenos.2025.29981
Forbes J Med 2025;6(3):213-220

Orijinal Araştırma / Original Article

Toplum Kökenli Pnömonili Çocuklarda Yoğun Bakım Gereksinimini Öngörmede Hemogram Türevli Enflamatuvar İndekslerin Tanısal Değeri

Diagnostic Value of Hemogram-Derived Inflammatory Indices in Predicting Intensive Care Requirement in Children with Community-Acquired Pneumonia

Leyla ŞERO, Muhammet ASENA

Sağlık Bilimleri Üniversitesi, Gazi Yaşargil Eğitim ve Araştırma Hastanesi, Pediatri Kliniği, Diyarbakır, Türkiye

Atıf: Şero L, Asene M. Diagnostic value of hemogram-derived inflammatory indices in predicting intensive care requirement in children with community-acquired pneumonia. Forbes J Med. 2025;6(3):213-220

ÖZ

Amaç: Bu çalışma, pediatrik toplum kökenli pnömonide (TKP) nötrofil lenfosit oranı (NLR), platelet lenfosit oranı (PLR) ve sistemik immün enflamasyon indeksinin (SII) yoğun bakım gereksinimini öngörmedeki tanısal değerini araştırmayı amaçlamaktadır.

Yöntem: Bu retrospektif çalışmada, pediatrik TKP tanısı almış hastaların demografik verileri, klinik bulguları ve laboratuvar sonuçları değerlendirildi. TKP'li hastalar; pediatri servisinde (Grup 1) ve pediatrik yoğun bakım ünitesinde (Grup 2) takip ve tedavi edilenler olmak üzere iki gruba ayrıldı. Hemogram parametrelerinden hesaplanan NLR, PLR, SII, yatıştaki kan gazı (pH, serum laktat düzeyi), C-reaktif protein ve prokalsitonin (PCT) analiz edildi. Bulgular, hastalığın şiddetini tahmin etmedeki öngördürücülüğü açısından değerlendirildi.

Bulgular: Toplam 116 hasta çalışmaya dahil edildi. Bunların 56'sı (%48,3) Grup 1'de, 60'ı (%51,7) Grup 2'de yer aldı. Grup 2'deki hastalarda NLR (p=0,039), PLR (p=0,039) ve SII (p=0,026) değerleri anlamlı düzeyde yüksek bulundu. Alıcı işletim karakteristiği analizinde NLR için eğri altında kalan alan (AUC): 0,718; SII için AUC: 0,705 ve PLR için AUC: 0,684 olarak hesaplandı. NLR, kan pH'ı, serum laktat ve PCT'nin yoğun bakım gereksinimini öngörmede bağımsız belirteçler olarak saptandı. Hipoksi, komorbid hastalıklar, anormal vital bulgular ve yüksek enflamasyon indeksleri, yoğun bakım yatışını öngörmede diğer anlamlı risk faktörleri olarak gözlendi.

Sonuç: NLR, PLR ve SII, çocukluk çağı TKP'sinde hastalık şiddetini değerlendirmede ve yoğun bakım ihtiyacını öngörmede düşük maliyetli, hızlı karar vermede yardımcı araç olarak kullanılabilir ve klinik karar sürecini destekleyebilir.

Anahtar Kelimeler: Toplum kökenli pnömoni, çocuk, nötrofil lenfosit oranı (NLR), platelet lenfosit oranı (PLR), sistemik enflamatuvar indeks (SII)

ABSTRACT

Objective: This study aimed to investigate the diagnostic value of the neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and systemic immune-inflammation index (SII) in predicting intensive care unit (ICU) requirements in pediatric community-acquired pneumonia (CAP).

Methods: In this retrospective study, demographic data, clinical findings, and laboratory results of pediatric patients diagnosed with CAP were evaluated. Patients were divided into two groups: those treated in the pediatric ward (Group 1) and those managed in the pediatric ICU (PICU) (Group 2). NLR,

Received/Geliş: 06.07.2025
Accepted/Kabul: 30.09.2025
Epub: 23.10.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Dr. Leyla ŞERO,
Sağlık Bilimleri Üniversitesi, Gazi Yaşargil Eğitim ve Araştırma Hastanesi, Pediatri Kliniği, Diyarbakır, Türkiye
ozlemleylasero@gmail.com
ORCID: 0000-0001-5224-8583

PLR, and SII calculated from hemogram parameters, as well as arterial blood gas values (pH, serum lactate), C-reactive protein, and procalcitonin (PCT), were analyzed. Their predictive roles in determining disease severity were assessed.

Results: A total of 116 patients were included, of whom 56 (48.3%) were in Group 1 and 60 (51.7%) in Group 2. NLR (p=0.039), PLR (p=0.039), and SII (p=0.026) levels were significantly higher in Group 2. Receiver operating characteristic analysis revealed area under the curve values of 0.718 for NLR, 0.705 for SII, and 0.684 for PLR. NLR, blood pH, serum lactate, and PCT were identified as independent predictors of ICU admission. Hypoxia, comorbid conditions, abnormal vital signs, and elevated inflammatory indices were additional significant risk factors associated with PICU admission.

Conclusion: NLR, PLR, and SII may serve as inexpensive and readily available markers to assess disease severity and predict intensive care needs in pediatric CAP. These indices can support rapid clinical decision-making and facilitate risk stratification in practice.

Keywords: Community-acquired pneumonia, pediatric, neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), systemic immune-inflammatory index (SII)

GİRİŞ

Toplum kökenli pnömoni (TKP), enfeksiyon etkeninin hastane dışı koşullarda edinildiği ve akciğer parankimini tutan akut bir enfeksiyon tablosudur. Çoğu çocuk tamamen iyileşse de bazı olgularda akciğerde lokal veya sistemik komplikasyonlar gelişebilir.[1] Özellikle yaygın pulmoner lezyonların varlığı, hipoksemi, sistemik komplikasyonlar ve ağır klinik bulgularla seyreden "ciddi pnömoni", karmaşık bir enfeksiyon tablosu olarak kabul edilir.[2]

Ciddi pnömoni, çocukluk çağında önemli bir mortalite nedenidir.[3] Çocukluk döneminde yaş grubuna bağlı olarak ciddi pnömoni gelişmesi açısından farklı duyarlılıklar söz konusudur. Bağışıklık sistemi henüz gelişmekte olan bebekler ve küçük çocuklar, hava yollarının daha dar olması nedeniyle bu duruma daha yatkındır.[3,4] Klinik kötüleşme, beslenme güçlüğü veya dehidratasyon, bilinç değişikliği, belirgin takipne, santral siyanoz, solunum sıkıntısı, birden fazla akciğer lobunun veya tek bir akciğerin üçte ikisinin tutulumuyla birlikte görülen plevral efüzyon, oksijen satürasyonunun %92'nin altına düşmesi ya da herhangi bir ekstrapulmoner komplikasyon ciddi hastalık belirtileridir.[5,6]

TKP erken tanısı ve özellikle yoğun bakım ünitesi (YBÜ) ihtiyacının öngörülebilmesi, mortaliteyi azaltmak ve sekelsiz iyileşmeyi desteklemek açısından kritik öneme sahiptir. Tanı genellikle hastanın öyküsü, fizik muayene bulguları ve göğüs radyografisi ile konur. Etkenin belirlenmesinde ise mikrobiyolojik, serolojik ve moleküler yöntemler kullanılmaktadır. Akut faz reaktanları olarak değerlendirilen beyaz kan hücresi (WBC), mutlak nötrofil sayısı, C-reaktif protein (CRP) ve prokalsitonin (PCT), enflamasyon ve enfeksiyonun şiddetini değerlendirmek için sıklıkla kullanılan parametrelerdir. Ancak bu belirteçler, her zaman tanı koymada yeterli veya zamanında bilgi sağlamayabilir.[7,8] Ayrıca, bu parametrelerin pnömoniye özgül olmaması, daha spesifik ve duyarlı biyobelirteçlerin araştırılmasını gerekli kılmıştır.[7]

Özellikle sistemik immün enflamasyon indeksi (SII), nötrofil, lenfosit ve trombosit sayılarını entegre ederek sistemik enflamasyon düzeyini ve immün yanıt dengesini yansıtan yeni nesil bir parametredir.

Güncel çalışmalar, rutin tam kan sayımı parametrelerinden elde edilen enflamatuvar oranların; özellikle nötrofil-lenfosit oranı (NLR), trombosit-lenfosit oranı (PLR) ve SII gibi göstergelerin; kanser, koroner arter hastalığı, sepsis ve 2019 koronavirüs hastalığı (COVID-19) enfeksiyonu gibi çeşitli sistemik enflamatuvar hastalıkların tanı ve prognoz süreçlerinde anlamlı katkı sağladığını göstermektedir.[9-12] Bu indeksler arasında özellikle NLR, PLR ve SII sistemik enflamasyon düzeyini ve immün yanıtın dengesini yansıtan, ekonomik, hızlı ve hesaplaması kolay parametreler olarak öne çıkmaktadır.

Bu çalışmanın amacı, pnömoni tanısı almış çocuk hastalarda hemogram parametrelerinden hesaplanan enflamatuvar indekslerin yoğun bakım ihtiyacını öngörmedeki tanısal değerini belirlemektir.

YÖNTEM

Bu retrospektif çalışmaya, 1 Ocak 2024-1 Ocak 2025 tarihleri arasında üçüncü düzey bir pediatrik YBÜ (PYBÜ) ile pediatri servislerine TKP tanısıyla kabul edilen, 3 ay ile 18 yaş arasındaki çocuk hastalar dahil edildi. Veriler, hastane bilgi yönetim sisteminden geriye dönük olarak elde edildi. Bronkopulmoner displazi, serebral palsi, nöromüsküler hastalıklar, kistik fibrozis gibi kronik akciğer hastalıkları bulunan, majör konjenital anomalisi olan veya eksik veriye sahip hastalar çalışma dışı bırakıldı. Bu hastalar, altta yatan kronik akciğer hastalıkları ve sistemik etkiler nedeniyle yüksek düzeyde bazal enflamatuvar aktivite sergileyebileceğinden, sistemik enflamatuvar indekslerin yoğun bakım yatış endikasyonunu öngörmedeki özgüllüğünü ve güvenilirliğini anlamlı biçimde düşürebileceğinden çalışma dışı bırakıldı.

Pnömoni tanısı; ateş, solunum sistemi ile ilişkili semptomlar ve fizik muayene bulguları ile birlikte ve/veya akciğer grafisinde parankim tutulumunun gösterilmesiyle konuldu. Hastalar iki gruba ayrıldı: pediatri servisinde takip edilenler (Grup 1) ve PYBÜ'de izlenenler (Grup 2). Kaydedilen klinik ve demografik veriler arasında başvuru özellikleri, tekrarlayan

yatış öyküsü, doğumla ilgili bilgiler (düşük doğum ağırlığı, prematürite, doğum şekli), aşılanma ve beslenme durumu, hipotansiyon, siyanoz, solunum güçlüğü şiddeti, akciğer grafisi bulguları ve hastanede yatış süresi yer aldı.

Laboratuvar analizleri aşağıdaki parametreleri içermekteydi:

• Kan gazı: pH, laktat (mmol/L)

• Tam kan sayımı: WBC ($10^3/mm^3$), nötrofil ($10^3/mm^3$), lenfosit ($10^3/mm^3$), trombosit ($10^3/mm^3$)

• Biyokimya: CRP (mg/dL), PCT (ng/mL)

Her hasta için aşağıdaki sistemik enflamasyon indeksleri hesaplandı:[9]

• NLR: Nötrofil/Lenfosit

• PLR: Trombosit/Lenfosit

• SII: (Trombosit × Nötrofil)/Lenfosit)

Elde edilen bulgular, gruplar arasında demografik veriler, klinik özellikler ve enflamatuvar indeksler açısından karşılaştırıldı.

Bu çalışma için, Sağlık Bilimleri Üniversitesi, Gazi Yaşargil Eğitim ve Araştırma Hastanesi, Klinik Araştırmalar Etik Kurulu'ndan onay alınmıştır (karar numarası: 369, tarih: 28.02.2025). Araştırma, Helsinki Deklarasyonu'na uygun olarak gerçekleştirilmiş ve tüm katılımcılardan bilgilendirilmiş onam alınmıştır.

İstatistiksel Analiz

Tüm istatistiksel analizler IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp., Armonk, NY, ABD) yazılımı kullanılarak gerçekleştirildi. Sürekli değişkenler, dağılım özelliklerine göre ortalama ± standart sapma veya medyan (çeyreklik aralık) olarak sunuldu. Değişkenlerin normal dağılıma uyumu Kolmogorov-Smirnov testi ile değerlendirildi. Kategorik değişkenler frekans (n) ve yüzde (%) ile ifade edildi. Kategorik değişkenler, uygun olduğunda ki-kare testi veya Fisher kesin testi ile karşılaştırıldı. Normal dağılım gösteren sürekli değişkenler için Student's t testi, normal dağılmayan sürekli değişkenler için Mann-Whitney U testi kullanıldı. Lojistik regresyon analizinde olasılık oranı (OR) ve %95 güven aralıkları (GA) hesaplandı.

Hematolojik parametreler ve enflamatuvar indeksler için tanısal doğruluğu belirlemek amacıyla alıcı işletim karakteristiği (ROC) eğrisi analizleri yapıldı ve eğri altında kalan alan (AUC) değerleri hesaplandı. AUC değerleri 0,5-0,6 arasında ise zayıf, 0,6-0,7 arasında ise orta, 0,7-0,8 ve üzeri ise güçlü ayrım gücüne sahip olarak değerlendirildi. En iyi kesim noktası, youden indeksi kullanılarak belirlendi.

Tüm analizlerde iki yönlü $p<0,05$ değeri istatistiksel olarak anlamlı kabul edildi.

BULGULAR

Çalışmaya TKP tanısı konulan 116 pediatrik hasta dahil edildi. Hastaların 56'sı (%48,3) serviste, 60'ı (%51,7) ise PYBÜ'de takip edilmiştir. Grup 1'de hastaların yaşları medyan 6 ay (1-86 ay), Grup 2'de ise 8 ay (1-82 ay) olup; her iki grupta da yaşlar benzerdi (p=0,530). Hastaların demografik ve klinik özellikleri Tablo 1'de özetlenmiştir.

Çalışmaya dahil edilen hastaların 94'ü (%81,0) bir yaş altı, 57'si (%49,1) ise altı ay altı bebeklerden oluşmaktaydı. Gruplar arasında cinsiyet dağılımı, prematürite öyküsü, doğum şekli, emzirme durumu ve aşılanma oranları açısından istatistiksel fark yoktu (p>0,050).

Grup 2'de, klinik bulgulardan siyanoz (%30; Grup 1'de %9; p=0,040) ve hipotansiyon (%21,7; Grup 1'de %9,0; p=0,049) daha sık görülmüştür. Ayrıca, Grup 2'de hastanede kalış süresi medyan 6,5 gün (2-54 gün) iken, Grup 1'de bu süre 5 gün (2-8 gün) olarak saptanmıştır (p<0,010) (Tablo 2).

YBÜ yatırılan hastaların altısında (%10,0) mortalite gelişmiş, diğer altısı (%10,0) ise evde solunum desteği gereksinimi devam ederek sekel ile taburcu edilmiştir. Mortalite gelişen hastaların ikisi ve evde cihaz desteği ile taburcu edilen hastaların biri altı aydan küçük bebeklerdi. Grup 2'de, başvuru sırasında 65 hasta (%56,0), solunum sıkıntısı nedeniyle yüksek akımlı nazal kanül ile izlenmiştir.

Tablo 1. Grupların demografik ve klinik özelliklerinin karşılaştırılması

	Grup 1 (n=56)	Grup 2 (n=60)	p
Yaş (ay)*	6 (1-86)	8 (1-82)	0,530
Cinsiyet, erkek, n (%)	29 (51,7)	33 (55,0)	0,850
Düşük doğum ağırlığı, n (%)	2 (3,6)	8 (13,3)	0,056
Prematürite, n (%)	18 (32,0)	16 (26,7)	0,560
Doğum şekli, C/S, n (%)	27 (48,2)	26 (65,0)	0,709
Anne sütü alma oranı, n (%)	51 (91,0)	54 (91,0)	1,000
Aşı öyküsü, tam, n (%)	51 (91,0)	53 (90,0)	0,760
Hastanede yatış öyküsü, n (%)	12 (21,4)	15 (25,0)	0,820

*Medyan (minimum-maksimum), C/S: Sezaryen doğum

Laboratuvar bulguları incelendiğinde, Grup 2'de arteriyel pH düzeyleri anlamlı olarak daha düşüktü (7,24±0,12'ye karşı 7,38±0,05; p<0,010). Laktat düzeyi anlamlı şekilde daha yüksekti (3,4±1,6 mmol/L'ye karşı 1,9±0,7 mmol/L; p<0,010). PCT düzeyleri de Grup 2'de belirgin olarak yüksekti (9,8±3,6 ng/mL'ye karşı 0,7±0,15 ng/mL; p=0,019). Ancak CRP düzeyleri arasında anlamlı fark bulunmamıştır (p=0,330). Her iki grupta da CRP düzeyleri yüksek saptanmış olmakla birlikte, gruplar arası fark istatistiksel olarak anlamlı değildi (46±7,0'e karşı 33±4,4 mg/dL; p=0,330) (Tablo 3).

Enflamatuvar indeks analizine göre, NLR (p=0,039), PLR (p=0,039) ve SII (p=0,026) değerleri yoğun bakım grubunda anlamlı olarak daha yüksek bulunmuştur (Tablo 4). ROC analizi sonuçları sırasıyla; NLR için AUC=0,718 (%95 GA: 0,324-0,532), SII için AUC=0,705 (%95 GA: 0,351-0,562), PLR için AUC=0,684 (%95 GA: 0,317-0,527) olarak saptanmıştır. Kesme değerleri; NLR için 2,7 (duyarlılık %74, özgüllük %74), SII için 1.006 (duyarlılık %76, özgüllük %74) ve PLR için 123 (duyarlılık %75, özgüllük %60) olarak belirlenmiştir (Tablo 5, Şekil 1).

PYBÜ kabulünün bağımsız öngörücülerini belirlemek amacıyla çok değişkenli lojistik regresyon analizi yapılmıştır. Dahil edilen değişkenler arasında hipotansiyon, siyanoz, arteriyel pH, laktat, PCT, NLR, PLR ve SII yer almıştır. Analiz sonucunda, arteriyel pH (OR: 5,3; %95 GA: 2,2-12,9; p<0,001), laktat (OR: 3,7; %95 GA: 1,692-4,212; p=0,001), PCT (OR: 4,1; %95 GA: 0,916-1,668; p=0,002) ve NLR (OR: 2,5; %95 GA: 1,017-1,461; p=0,031) bağımsız risk faktörleri olarak belirlenmiştir. Buna karşın, hipotansiyon, siyanoz, PLR ve SII açısından istatistiksel olarak anlamlı fark saptanmamıştır (Tablo 4, Şekil 1).

Etiyolojik değerlendirmede; 45 hastada (%38,8) *Haemophilus influenzae*, 27 hastada (%23,2) *Streptococcus pneumoniae*, 11 hastada (%9,5) *Respiratuvar Sinsityal Virüs* (RSV), 8 hastada (%6,9) *Rinovirüs*, 7 hastada (%6,0) *İnfluenza A*, 4 hastada (%3,4) *Adenovirüs*, 3 hastada (%2,6) *Bocavirüs* ve 1 hastada (%0,8) *İnfluenza B* saptanmıştır. Ayrıca, 26 hastada (%22,4) hem viral hem de bakteriyel etkenlerin birlikte bulunduğu miks enfeksiyonlar gözlenmiştir.

Sekel ile seyreden hastalarda etken olarak *Haemophilus influenzae*, *Adenovirüs* ve *İnfluenza A* tespit edilmiştir. Eksitus olan hastalarda ise etkenler *Haemophilus influenzae* ve *Streptococcus pneumoniae* olarak belirlenmiştir.

Tablo 2. Grupların klinik özelliklerinin karşılaştırılması

	Grup 1 (n=56)	Grup 2 (n=60)	p
Hipotansiyon, n (%)	5 (9,0)	13 (21,7)	0,049
Siyanoz, n (%)	5 (9,0)	18 (30,0)	0,040
Atelektazi, n (%)	0	1 (1,7)	0,740
Plevral efüzyon, n (%)	0	2 (3,3)	0,260
Hastanede yatış süresi, gün*	5 (2-8)	6,5 (2-54)	<0,010

*Medyan (minimum-maksimum)

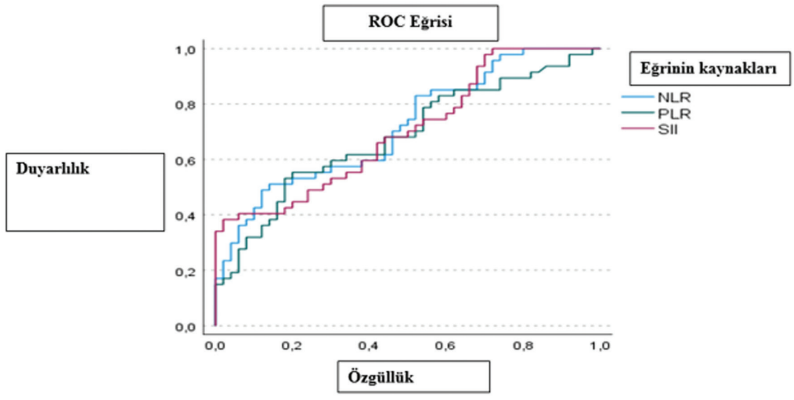
Tablo 3. Grupların laboratuvar özelliklerinin karşılaştırılması

	Grup 1 (n=56)	Grup 2 (n=60)	p
WBC ($10^3/mm^3$)*	13,6±7,4	13,7±6,6	0,960
Nötrofil sayısı ($10^3/mm^3$)*	7,2±4,2	7,6±3,9	0,640
Lenfosit sayısı ($10^3/mm^3$)*	5,3±3,5	4,9±3,0	0,490
Trombosit sayısı ($10^3/mm^3$)*	419±145	426±192	0,830
pH*	7,38±0,05	7,24±0,12	<0,010
Laktat (mmol/L)*	1,9±0,70	3,4±1,60	<0,010
CRP (mg/dL)*	33±4,40	46±7,00	0,330
Prokalsitonin (ng/mL)*	0,7±0,15	9,8±3,60	0,019

*Ortalama ± standart sapma.
WBC: Beyaz kan hücresi, CRP: C-reaktif protein

Tablo 4. Grupların enflamasyon belirteçleri açısından karşılaştırılması			
	Grup 1 (n=56)	**Grup 2 (n=60)**	**p**
NLR*	1,8±0,25	2,8±0,40	0,039
PLR*	99±18	138±25	0,039
SII*	684,6±81,5	1091,6±157,5	0,026
*Ortalama ± standart sapma. NLR: Nötrofil/lenfosit oranı, PLR: Trombosit/lenfosit oranı, SII: Sistemik enflamasyon indeksi [(Trombosit × Nötrofil)/Lenfosit]			

Tablo 5. Grupların ROC analizi kesme değerleri, sensitivite ve spesifiteleri				
Enflamatuvar belirteçler	**AUC**	**Kesme değeri**	**Sensitivite (%)**	**Spesifisite (%)**
SII	0,705	1006,5	76	74
NLR	0,718	2,7	74	74
PLR	0,683	123	75	60
NLR: Nötrofil/lenfosit oranı, PLR: Trombosit/lenfosit oranı, SII: Sistemik enflamasyon indeksi [(Trombosit × Nötrofil)/Lenfosit]				



Şekil 1. SII, NLR ve PLR indekslerinin ROC analizi

NLR: Nötrofil/lenfosit oranı, PLR: Trombosit/lenfosit oranı, SII: Sistemik enflamasyon indeksi [(Trombosit × Nötrofil) / Lenfosit], ROC: Alıcı işletim karakteristiği

TARTIŞMA

Bu çalışmada, TKP tanısıyla hastaneye yatırılan çocukların klinik ve laboratuvar özellikleri incelenmiş ve PYBÜ kabulün öngördürücüleri olarak enflamatuvar belirteçlerin rolüne odaklanılmıştır. Bulgularımız, NLR, PLR ve SII gibi enflamatuvar indekslerin risk sınıflandırmasında kullanılabilirliğini vurgulayan güncel literatür ile uyumludur.

TKP, beş yaş altı çocuklarda halen önemli bir morbidite ve mortalite nedenidir. Hastalık, genellikle hafif-orta şiddette başlayıp, yoğun bakımda takip ve tedavi gerektiren ciddi bir klinik tabloya dönüşebilmektedir.[13,14] Bu nedenle, erken risk sınıflandırması ve yüksek riskli olguların zamanında tanımlanması büyük önem taşımaktadır.

Hematolojik parametreler, pediatrik pnömonide en sık başvurulan enflamasyon belirteçleri arasındadır. Literatürde bu parametrelerin hastalık şiddeti ve yoğun bakım gereksinimi ile ilişkili olduğu çok sayıda çalışma ile gösterilmiştir.[15-21] İlginç bir şekilde, CRP gibi geleneksel akut faz reaktanlarının artmasına rağmen, CRP düzeylerinin yoğun bakım ve servis grupları arasında anlamlı bir fark göstermediği saptanmıştır. Bu durum, CRP'nin tek başına hastalık şiddetini belirlemede yeterince spesifik olmadığını ortaya koymaktadır. Son dönem çalışmalarda CRP'nin diğer enflamatuvar belirteçlerle birlikte değerlendirilmesinin klinik öngörü gücünü artırdığı ve biyobelirteç panellerinin önemini vurguladığı belirtilmiştir.[15-21]

Yakın tarihli çalışmalarda; WBC, nötrofil, lenfosit sayıları ile birlikte NLR, PLR ve SII değerlerinin yoğun bakım ihtiyacı olan hastalarda anlamlı düzeyde yüksek olduğu, bu belirteçlerin PCT ve CRP ile pozitif; kritik hasta skorları ile negatif korelasyon gösterdiği bildirilmiştir.[15,19,22]

Wang ve ark.nın[23] yaptığı çalışmada, özellikle NLR'nin pnömoni hastalarında hastalık şiddeti, yoğun bakım yatışı, intrakraniyal komplikasyonlar ve kötü prognoz ile

ilişkili olduğu ortaya konmuştur. Elmeazawy ve ark.'nın[17] prospektif çalışmasında ise nekrotizan pnömoni geçiren çocuklarda NLR, PLR, SII ve sistemik enflamatuvar yanıt indeksi (SIRI) anlamlı şekilde yüksek bulunmuş ve SII + SIRI + D-dimer kombinasyonunun ROC analizinde yüksek ayırt edicilik gösterdiği belirtilmiştir.

Bizim çalışmamızda da PYBÜ'de takip edilen çocuklarda NLR, PLR ve SII değerlerinin anlamlı olarak yüksek olduğu ve bu bulguların literatürle tutarlı olduğu görülmüştür. Benzer şekilde, monosit-lenfosit oranı ve SIRI'nin çocuklarda *Mycoplasma pneumoniae* enfeksiyonu için risk belirteci olduğu, SII değeri ≥699,00 olan çocukların şiddetli Mikoplazma pnömonisi geliştirme olasılığının daha yüksek olduğu ve SII'nin hastalık şiddetini ayırt etmede faydalı bir belirteç olabileceği bildirilmiştir.[24,25]

COVID-19 hastalarında yapılan çalışmalarda da hastalık şiddeti arttıkça NLR, PLR ve SII seviyelerinde belirgin yükselme olduğu, yoğun bakım gereksinimi olan olgularda bu parametrelerin anlamlı düzeyde yüksek olduğu bildirilmiştir.[26-29] Parainfluenza ilişkili viral pnömonilerde ise NLR, PLR ve türetilmiş NLR [d-NLR; nötrofil/(WBC-nötrofil)] oranlarının yoğun bakım ihtiyacı ile ilişkili olarak yüksek bulunduğu bildirilmiştir.[30]

Bununla birlikte, PCT ve CRP gibi biyobelirteçler de yoğun bakım ihtiyacını öngörmede yüksek prognostik değere sahiptir. Literatürde, PCT'nin özellikle bakteriyel enfeksiyonlarda CRP'ye kıyasla daha yüksek doğruluk sağladığı ve Mikoplazma pnömonilerinde her iki belirtecin de hastalık şiddetiyle korele olduğu gösterilmiştir.[17,24,25] RSV pnömonili çocuklarda yapılan çalışmalarda ise PCT düzeyinin bakteriyel koenfeksiyon ve hastalık şiddeti ile anlamlı korelasyon gösterdiği belirtilmiştir.[31] Çalışmamızda da PCT düzeylerinin yoğun bakım grubunda anlamlı şekilde yüksek olduğu, CRP düzeylerinin ise gruplar arasında fark göstermediği tespit edilmiştir. Bu bulgu, CRP'nin hastalık şiddetini öngörmede tek başına sınırlı değere sahip olduğunu göstermekte ve mevcut literatür ile uyumlu bulunmuştur.[15,19,32-34]

Verilerimiz, genel servislerde takip edilenlere kıyasla PYBÜ'de izlenen çocuklarda anlamlı şekilde daha düşük kan pH'ı ve daha yüksek serum laktat düzeyleri olduğunu göstermektedir. Bu durum, kritik hastalarda daha belirgin metabolik bozukluk ve doku hipoksisini yansıtmaktadır.[26-28] Bulgularımız, artmış serum laktat seviyesinin pnömonide kötü prognozun güvenilir bir göstergesi olduğunu belirten literatür ile örtüşmektedir.[35-37] Aynı şekilde, PCT düzeylerindeki artış da ciddi sistemik enfeksiyon ve sepsisin varlığını desteklemektedir.

Çalışmamızda PYBÜ grubunda siyanoz ve hipotansiyon daha sık görülmüş olsa da bu bulgular çok değişkenli analizde bağımsız öngörücü olarak anlamlı bulunmamıştır. Bu durum, klinik belirtilerin yoğun bakım ihtiyacını öngörmede tek başına yeterli olmadığını göstermektedir.

Bu çalışmanın bazı sınırlılıkları mevcuttur. Öncelikle çalışmanın retrospektif tasarımlı ve tek merkezli olması, verilerin genellenebilirliğini sınırlamaktadır. ROC analizi ile NLR, PLR ve SII için cut-off değerleri belirlenmiş olmakla birlikte, bu eşik değerler farklı merkezlerde ve popülasyonlarda değişkenlik gösterebilir. Bu nedenle, klinik uygulamada kesin referans aralıkları henüz netleşmemiştir. Pnömoni etkenleri arasında bakteriyel ve viral patojenlerin yanı sıra miks enfeksiyonların da görülmesi, enflamatuvar indekslerin yorumunu güçleştirmiş ve bu belirteçlerin özgüllüğünü azaltmış olabilir. Ayrıca, çalışmamızda yalnızca hemogram türevi belirteçler (NLR, PLR, SII) değerlendirilmiş; ferritin, interlökin-6, D-dimer gibi diğer enflamatuvar biyobelirteçler veya klinik skor sistemleri (örneğin; PRISM, PIM) ile karşılaştırma yapılmamıştır. Bu da elde edilen parametrelerin prognostik değerinin daha kapsamlı değerlendirilmesini sınırlandırmıştır.

SONUÇ

Sonuç olarak, bu çalışma TKP'li çocuklarda yoğun bakım ihtiyacının erken öngörülmesinde, pratik ve kolay erişilebilir parametreler olan NLR, PLR ve SII gibi enflamatuvar indekslerin; kan pH'ı, serum laktat ve PCT düzeyleri ile birlikte değerlendirilmesinin faydalı olabileceğini ortaya koymuştur. Bu belirteçlerin klinik protokollere entegre edilmesi, ciddi olguların hızlı tanımlanmasını kolaylaştırabilir ve hasta sonuçlarını iyileştirebilir.

Ancak bu parametrelerin yaygın ve güvenilir klinik kullanımı için, prospektif, çok merkezli ve yapay zeka destekli modeller ile klinik risk skor sistemlerine entegre edilmiş ileri düzey çalışmalara ihtiyaç vardır.

Etik

Etik Kurul Onayı: Bu çalışma için, Sağlık Bilimleri Üniversitesi, Gazi Yaşargil Eğitim ve Araştırma Hastanesi, Klinik Araştırmalar Etik Kurulu'ndan onay alınmıştır (karar numarası: 369, tarih: 28.02.2025).

Hasta Onayı: Retrospektif çalışma.

Dipnotlar

Yazarlık Katkıları

Cerrahi ve Medikal Uygulama: L.Ş., M.A., Konsept: L.Ş., Dizayn: L.Ş., M.A., Veri Toplama veya İşleme: L.Ş., M.A., Analiz veya Yorumlama: L.Ş., M.A., Literatür Arama: L.Ş., M.A., Yazan: L.Ş., M.A.

Çıkar Çatışması: Yazarlar bu makale ile ilgili olarak herhangi bir çıkar çatışması bildirmemiştir.

Finansal Destek: Çalışmamız için hiçbir kurum ya da kişiden finansal destek alınmamıştır.

KAYNAKLAR

1. GBD 2016 Lower Respiratory Infections Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis. 2018;18:1191-210.
2. de Benedictis FM, Kerem E, Chang AB, Colin AA, Zar HJ, Bush A. Complicated pneumonia in children. Lancet. 2020;396:786-98.
3. Shi T, Chen C, Huang L, et al. Risk factors for mortality from severe community acquired pneumonia in hospitalized children transferred to the pediatric intensive care unit. Pediatr Neonatol. 2020;61:577-83.
4. Shah SS, Bradley JS. Pediatric community-acquired pneumonia. In: Cherry JD, Harrison G, Kaplan SL, et al. editors. Feigin and Cherry's Textbook of Pediatric Infectious Diseases, 8th ed. Philadelphia: Elsevier; 2018. p. 208-19.
5. Subspecialty Group of Respiratory Diseases, The Society of Pediatrics, Chinese Medical Association; Editorial Board, Chinese Journal of Pediatrics. Guidelines for management of community acquired pneumonia in children (the revised edition of 2013) (I) [in Chinese]. Zhonghua Er Ke Za Zhi. 2013;51:745–52.
6. British Thoracic Society of Standards of Care Committee. BTS guidelines for the community acquired pneumonia in childhood. Thorax. 2002;57(Suppl 1):1-24.
7. Madhi SA, Kohler M, Kuwanda L, Cutland C, Klugman KP. Usefulness of C-reactive protein to define pneumococcal conjugate vaccine efficacy in the prevention of pneumonia. Pediatr Infect Dis J. 2006;25:30-6.
8. Perren A, Cerutti B, Lepori M, et al. Influence of steroids on procalcitonin and C-reactive protein in patients with COPD and community-acquired pneumonia. Infection. 2008;36:163-6.
9. Zhu M, Chen L, Kong X, et al. The systemic inflammation response index as an independent predictor of survival in breast cancer patients: a retrospective study. Front Mol Biosci. 2022;9:856064.
10. Urbanowicz T, Michalak M, Olasińska-Wiśniewska A, et al. Neutrophil counts, neutrophil-to-lymphocyte ratio, and systemic inflammatory response index (SIRI) predict mortality after off-pump coronary artery bypass surgery. Cells. 2022;11:1124.
11. Yang J, Wang H, Hua Q, Wu J, Wang Y. Diagnostic value of systemic inflammatory response index for catheter-related bloodstream infection in patients undergoing haemodialysis. J Immunol Res. 2022;2022:7453354.
12. Nooh HA, Abdellateif MS, Refaat L, et al. The role of inflammatory indices in the outcome of COVID-19 cancer patients. Med Oncol. 2021;39:6.
13. Rudan I, Boschi-Pinto C, Biloglav Z, Mulholland K, Campbell H. Epidemiology and etiology of childhood pneumonia. Bull World Health Organ. 2008;86:408-16.
14. Wardlaw T, Salama P, Johansson EW, Mason E. Pneumonia: the leading killer of children. Lancet. 2006;368:1048-50.
15. Florin TA, Ambroggio L, Brokamp C, et al. Biomarkers and disease severity in children with community-acquired pneumonia. Pediatrics. 2020;145:e20193728. Erratum in: Pediatrics. 2020;146:e2020011452.
16. Lee H, Kim I, Kang BH, Um SJ. Prognostic value of serial neutrophil-to-lymphocyte ratio measurements in hospitalized community-acquired pneumonia. PLoS One. 2021;16:e0250067.
17. Elmeazawy R, Ayoub D, Morad LM, El-Moazen AMF. Role of systemic immune-inflammatory index and systemic inflammatory response index in predicting the diagnosis of necrotizing pneumonia in children. BMC Pediatr. 2024;24:496.
18. Wu J, Yan L, Chai K. Systemic immune-inflammation index is associated with disease activity in patients with ankylosing spondylitis. J Clin Lab Anal. 2021;35:e23964.
19. Li L, Miao H, Chen X, Yang S, Yan X. Research on the correlation of peripheral blood inflammatory markers with PCT, CRP, and PCIS in infants with community-acquired pneumonia. Evid Based Complement Alternat Med. 2022;2022:9024969.
20. Acar E, Gokcen H, Demir A, Yildirim B. Comparison of inflammation markers with prediction scores in patients with community acquired pneumonia. Bratisl Lek Listy. 2021;122:418-23.
21. Yadav RK, Kumar D, Gupta A, Sharma P. C-reactive protein and procalcitonin: as predictor biomarkers of severity and outcome in children with community-acquired pneumonia. Trop Doct. 2024;54:262-67.
22. Esposito S, Tagliabue C, Picciolli I, et al. Procalcitonin measurements for guiding antibiotic treatment in pediatric pneumonia. Respir Med. 2011;105:1939-45.
23. Wang RH, Wen WX, Jiang ZP, et al. The clinical value of neutrophil-to-lymphocyte ratio (NLR), systemic immune-inflammation index (SII), platelet-to-lymphocyte ratio (PLR) and systemic inflammation response index (SIRI) for predicting the occurrence and severity of pneumonia in patients with intracerebral hemorrhage. Front Immunol. 2023;14:1115031.
24. Shao L, Yu B, Lyu Y, Fan S, Gu C, Wang H. The clinical value of novel inflammatory biomarkers for predicting mycoplasma pneumoniae ınfection in children. J Clin Lab Anal. 2025;39:e25150.
25. Wang S, Wan Y, Zhang W. The clinical value of systemic immune inflammation ındex (SII) in predicting the severity of hospitalized children with mycoplasma pneumoniae pneumonia: a retrospective study. Int J Gen Med. 2024;17:935-942.
26. Ozdemir A, Kocak SY, Karabela SN, Yılmaz M. Can systemic immune inflammation index at admission predict in-hospital mortality in chronic kidney disease patients with SARS-CoV-2 infection? Nefrologia (Engl Ed). 2022;42:549-58.
27. Fois AG, Paliogiannis P, Scano V, et al. The systemic inflammation index on admission predicts in-hospital mortality in COVID-19 patients. Molecules. 2020;25:5725.
28. Sayed AA. Assessing the diagnostic values of the neutrophil-to-lymphocyte ratio (NLR) and systematic immunoinflammatory index (SII) as biomarkers in predicting COVID-19 severity: a multicentre comparative study. Medicina (Kaunas). 2024;60:602.
29. Yuan H, Tian J, Wen L. Meta-analysis of the systemic immune-inflammatory index and in-hospital mortality of COVID-19 patients. Heliyon. 2023;10:e23441.
30. Chu FL, Li C, Liu Y, Dong B, Qiu Y, Fan G. Peripheral blood parameters for predicting PICU admission and mechanical ventilation in pediatric inpatients with human parainfluenza virus-induced pneumonia. J Med Virol. 2023;95:e28752.
31. Do Q, Dao TM, Nguyen TNT, Tran QA, Nguyen HT, Ngo TT. Procalcitonin identifies bacterial coinfections in Vietnamese

children with severe respiratory syncytial virus pneumonia. Biomed Res Int. 2020;2020:7915158.

32. Yiğit E, Demir Yiğit Y. Diagnostic importance of serum C-reactive protein and procalcitonin in sepsis after burn. Int J Burns Trauma. 2021;11:391-6.
33. Ratageri VH, Panigatti P, Mukherjee A, et al. Role of procalcitonin in diagnosis of community acquired pneumonia in Children. BMC Pediatr. 2022;22:217.
34. McLaughlin JM, Khan F, Schmitt HJ, et al. Respiratory syncytial virus-associated hospitalization rates among us infants: a systematic review and meta-analysis. J Infect Dis. 2022;225:1100-11.
35. Sauer CM, Gómez J, Botella MR, et al. Understanding critically ill sepsis patients with normal serum lactate levels: results from U.S. and European ICU cohorts. Sci Rep. 2021;11:20076.
36. Song Y, Wang N, Xie X, Tian Y, Wang Y. Relationship between lactate levels and 28-day mortality in pediatric sepsis: results from the pediatric intensive care database. BMC Pediatr. 2024;24:712.
37. Xie M, Min Z, Jiang W, He Z, Xia X. Prognostic value of multivariate logistic regression analysis and amyloid a lactate monitoring in patients with severe pneumonia-associated sepsis. BMC Pulm Med. 2025;25:191.

DOI: 10.4274/forbes.galenos.2025.35492
Forbes J Med 2025;6(3):221-226

Original Article / Orijinal Araştırma

Medial Patellofemoral Ligament Reconstruction with and without Tibial Tubercle Osteotomy in Patellofemoral Instability Surgery: Clinical Outcomes

Patellofemoral İnstabilite Cerrahisinde Tibial Tüberkül Osteotomisi ile ve Osteotomisiz Medial Patellofemoral Ligament Rekonstrüksiyonu: Klinik Sonuçlar

Yılmaz ÖNDER, Mahmut TUNÇEZ, Tuğrul BULUT

İzmir Katip Çelebi University, Atatürk Training and Research Hospital, Department of Orthopaedics and Traumatology, İzmir, Türkiye

Cite as: Önder Y, Tunçez M, Bulut T. Medial patellofemoral ligament reconstruction with and without tibial tubercle osteotomy in patellofemoral instability surgery: clinical outcomes. Forbes J Med. 2025;6(3):221-226

ABSTRACT

Objective: The aim of this study was to evaluate the clinical outcomes of medial patellofemoral ligament (MPFL) reconstruction with and without tibial tubercle osteotomy (TTO) in patellofemoral instability surgery.

Methods: Between January 2020 and January 2025, 23 patients who underwent surgery for recurrent patellar dislocation due to patellofemoral instability, were retrospectively evaluated. Seventeen patients over 18 years of age, without any additional injuries, who had not undergone any surgery other than MPFL reconstruction ± TTO, with at least 12 months of follow-up, were included in the study. Twelve patients underwent MPFL reconstruction (Group 1), while five patients underwent MPFL reconstruction combined with TTO (Group 2). Demographic data of the patients were recorded. Clinical outcomes were assessed using Kujala, Lysholm, and IKDC 2000 scores, as well as visual analog scale (VAS).

Results: Of the patients included in the study, 11 were female (65%) and 6 were male (35%), with a median age of 22 years (18-39 years), body mass index of 25.3 (18.2-32.5), and follow-up duration of 32 months (15-150 months). Comparison of preoperative and postoperative clinical scores demonstrated significant improvement in Kujala, Lysholm, and IKDC 2000 scores in both groups ($p<0.001$), along with a significant decrease in VAS scores ($p<0.001$).

Conclusion: MPFL reconstruction with or without TTO provides satisfactory outcomes in patellofemoral instability surgery. This reconstructive procedure may be considered an effective treatment option for recurrent patellar dislocation and can be combined with TTO in appropriate cases.

Keywords: Medial patellofemoral ligament, reconstruction, tibial tubercle osteotomy, patellar instability

Received/Geliş: 22.09.2025
Accepted/Kabul: 14.10.2025
Epub: 23.10.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Tuğrul BULUT, Prof., MD
İzmir Katip Çelebi University, Atatürk Training and Research Hospital, Department of Orthopaedics and Traumatology, İzmir, Türkiye
drtugrulbulut@yahoo.com
ORCID: 0000-0002-7075-0873

ÖZ

Amaç: Bu çalışmanın amacı; patellafemoral instabilite nedeniyle tibial tüberkül osteotomisi (TTO) ile birlikte veya osteotomisiz medial patellofemoral ligament (MPFL) rekonstrüksiyonu yapılan hastaların klinik sonuçlarını değerlendirmektir.

Yöntem: Ocak 2020 ile Ocak 2025 yılları arasında patellafemoral instabiliteye bağlı tekrarlayan patella çıkığı nedeniyle ameliyat edilen 23 hasta retrospektif olarak incelendi. En az 12 aylık takibi olan, 18 yaş üstü, ek yaralanması olmayan ve MPFL rekonstrüksiyonu ± TTO harici bir cerrahi uygulanmamış olan 17 hasta çalışmaya dahil edildi. On iki hastaya MPFL rekonstrüksiyonu (Grup 1), 5 hastaya MPFL rekonstrüksiyonu ile birlikte TTO (Grup 2) uygulanmıştı. Hastaların demografik verileri kayıt altına alındı. Klinik sonuçlar Kujala, Lysholm ve IKDC 2000 skorları ile görsel analog skala (VAS) kullanılarak değerlendirildi.

Bulgular: Çalışmaya dahil edilen hastaların; 11'i kadın (%65), 6'sı erkek (%35), median yaş 22 yıl (aralık;18-39 yıl), vücut kitle indeksi 25,3 (aralık;18,2-32,5), ve takip süresi 32 aydı (aralık;15-150 ay). Hastaların ameliyat öncesi ve sonrası klinik skorları karşılaştırıldığında her iki grubun da Kujala, Lysholm ve IKDC 2000 skorlarında anlamlı yükselme/iyileşme (p<0,001), VAS skorlarında ise anlamlı düşme/düzelme izlendi (p<0,001).

Sonuç: Patellafemoral instabilite cerrahisinde MPFL rekonstrüksiyonu ± TTO tatminkâr sonuçlara sahiptir. Bu rekonstrüksiyon cerrahisi tekrarlayan patellar çıkık olgularında etkin bir tedavi seçeneği olarak değerlendirilebilir ve uygun olgularda TTO ile kombine edilebilir.

Anahtar Kelimeler: Medial patellofemoral ligament, rekonstrüksiyon, tibial tüberkül osteotomisi, patellar instabilite

INTRODUCTION

Recurrent patellar dislocation secondary to patellofemoral instability is common and is associated with medial patellofemoral ligament (MPFL) injury in approximately 90% of cases.[1] The MPFL is the primary restraint preventing lateral displacement of the patella between 0° and 30° of knee flexion; it is compromised in nearly 90% of dislocations.[2] Patients with recurrent patellar dislocation frequently present with underlying pathoanatomical abnormalities that predispose them to patellofemoral instability.[3,4] These abnormalities include lower extremity malalignment syndromes such as trochlear dysplasia and tibial external rotation, as well as osseous anomalies such as patella alta.[5-7] Most patients harbor more than one of these pathoanatomical risk factors.[6-8] The recurrence rate following a primary dislocation ranges from 17% to 69%.[1]

Given its multifactorial etiology, the management of recurrent patellar dislocation due to patellofemoral instability can be challenging. In recent years, there has been a growing interest in MPFL reconstruction for the surgical treatment of recurrent patellar dislocation.[2] Indeed, MPFL reconstruction has increasingly become the surgical technique of first choice for many surgeons.[9] However, in certain cases, concomitant osseous malalignment may render isolated MPFL reconstruction insufficient. In such patients, additional corrective procedures may be required.[10,11] Among these, tibial tubercle osteotomy (TTO) is the most commonly performed.[10]

The aim of this study is to evaluate the clinical outcomes of patients undergoing MPFL reconstruction with and without concomitant TTO for recurrent patellar dislocation.

METHODS

This retrospective study was approved by the İzmir Katip Çelebi University Health Research Institutional Review Board (decision no: 0037, date: 16.01.2025). Due to its retrospective design, informed consent from patients and healthcare providers was not required.

Patient Inclusion

Between January 2020 and January 2025, 23 patients who underwent surgery for recurrent patellar dislocation were retrospectively analyzed. Exclusion criteria were: having undergone other surgical procedures such as isolated medial plication, being under 18 years of age, having a follow-up period of less than 12 months, having concomitant injuries in the ipsilateral lower extremity, or undergoing additional surgical interventions including trochleoplasty, anterior cruciate ligament reconstruction, or meniscus repair. Seventeen patients who did not meet these criteria were included in the study. Twelve patients underwent isolated MPFL reconstruction, while five patients underwent MPFL reconstruction combined with TTO.

Evaluation of Patients

Clinical outcomes were assessed using the Kujala score, Lysholm score, visual analog scale (VAS), and the International Knee Documentation Committee Subjective Knee Form (IKDC 2000).

The IKDC 2000 score is a patient-reported outcome measure that evaluates knee-related symptoms, daily activities, and sports function. As it reflects patients' own experiences, it is considered a reliable and valid tool for both clinical studies and follow-up. Scores range from 0 to 100, with higher scores indicating better knee function.[12]

The Kujala score is a patient-based questionnaire developed to assess patellofemoral disorders and anterior knee pain. It includes parameters related to pain, function, walking, stair climbing, and kneeling. The total score ranges from 0 to 100, with higher scores reflecting better knee function.[13]

The Lysholm score is a questionnaire used to evaluate clinical outcomes after knee surgery. It comprises eight parameters, including pain, instability, locking, swelling, stair climbing, and daily activities. The total score ranges from 0 to 100, with higher scores indicating better knee function and patient satisfaction.[14]

The VAS score is used to assess patients' pain levels. For this purpose, a 10-cm VAS ruler was employed, with the left end marked as "no pain" (0 cm) and the right end marked as "very severe pain" (10 cm). Patients were asked to indicate the point on the ruler that best represented their pain intensity.

In patients who underwent TTO, bone union was evaluated using plain radiographs.

Surgical Technique and Follow-up Protocol

All procedures were initiated with routine arthroscopic examination. In patients with arthroscopically confirmed lateral patellar tightness (9/17), lateral release was performed using radiofrequency ablation. MPFL reconstruction was performed with a free autologous gracilis tendon graft. After graft harvesting, a 2-cm medial mini-incision was made over the patella, and two sockets with a depth of 1 cm were created using a 3.2-mm drill bit. The position of the sockets was verified under fluoroscopy (Figure 1a). The graft was pulled into the sockets and fixed with two 3.5-mm non-absorbable suture anchors. Subsequently, a 2-cm medial incision was made over the medial epicondyle, and a guide wire was advanced under fluoroscopic guidance from the isometric point near the medial femoral epicondyle, as described in previous studies (Figure 1b).[15] If the tibial tubercle-trochlear groove (TT-TG) distance exceeded 20 mm, tibial tubercle anteromedialization was performed as described by Fulkerso [16] (Figure 2a). For TTO, a paramedial skin incision was made to expose the patellar tendon and tibial tubercle. Vertical and longitudinal step-cut osteotomies were performed, blending into the anterior cortex of the tibia. The bone block was then corrected depending on the required distal transfer and medialized according to the TT-TG distance. Finally, the tubercle was fixed with two or three 4.0-mm cancellous screws (Figure 2b).

In patients who underwent isolated MPFL reconstruction, weight-bearing as tolerated was allowed with the knee immobilized at 60° of flexion for the first 3 weeks. After 3 weeks, flexion restriction was removed and strengthening exercises were initiated. In patients who underwent TTO, no weight-bearing was permitted and flexion was limited to 60° for the first 4 weeks. After 4 weeks, partial weight-bearing and mobilization were initiated, and strengthening exercises were started with unrestricted flexion.

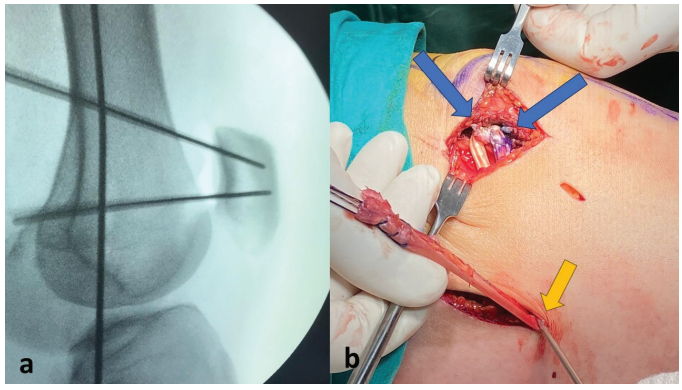


Figure 1. a) Fluoroscopic control of the patellar suture anchor placement using a K-wire. b) Gracilis tendon fixation with two suture anchors (blue arrows) and a guide wire (yellow arrow) for isometric femoral epicondyle fixation of the tendon

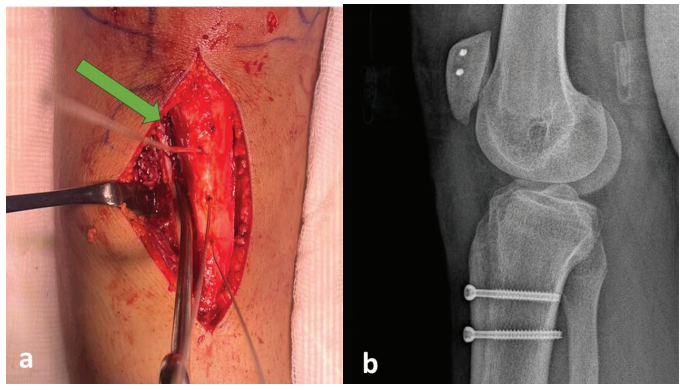


Figure 2. a) Medial and anterior transfer of the tibial tubercle after osteotomy (green arrow) and fixation with guide wires. b) Radiographic image of the tibial tubercle osteotomy line in the second postoperative year

Statistical Analysis

Data analysis was performed using IBM SPSS version 26. The suitability of continuous variables to a normal distribution was tested with the Shapiro-Wilk test. As the data were non-normally distributed, non-parametric tests were used. The Mann-Whitney U test was employed for comparisons between independent groups, while the Wilcoxon signed-rank test was applied for repeated measurements. Distribution of categorical variables was analyzed using Pearson's chi-squared test or Fisher's exact test. A type I error margin of α=0.05 was set, and all tests were performed two-tailed.

RESULTS

The median age of the 17 patients included in the study was 22 years (range: 18-39 years), with a median follow-up period of 32 months (range: 15-150 months). The median body mass index was 25.3 (range: 18.2-32.5). Eleven patients (65%) were female, and six (35%) were male. Nine patients (53%) underwent arthroscopic lateral release (Table 1).

When preoperative and postoperative clinical scores of all patients were evaluated, a significant improvement was observed (Table 2). The median Kujala score improved from 56 (range: 48-64) preoperatively to 84 (range: 78-90) at the final follow-up ($p<0.001$). Similarly, the median IKDC 2000 score increased from 66 (range: 50-78) to 86 (range: 82-92), showing a significant difference ($p<0.001$). The median Lysholm score improved from 50 (range: 48-60) to 90 (range: 80-96), which was statistically significant ($p<0.001$). The median VAS score decreased significantly from 6 (range, 4-8) preoperatively to 0 (range, 0-4) postoperatively ($p<0.001$) (Table 2).

When comparing patients (n=12, Group 1) who underwent isolated MPFL reconstruction with those (n=5, Group 2) who

underwent MPFL reconstruction combined with TTO, both groups demonstrated improvement in knee scores, but no statistically significant difference was detected between the groups (p>0.05). Similarly, VAS scores improved in both groups, with no significant intergroup difference (p>0.05) (Table 2).

All patients who underwent TTO achieved bone union, with a median bone union time of 12 weeks (range: 8-16 weeks). One patient developed screw head irritation, which required hardware removal at 30 weeks, after confirmed bone union. Another patient experienced recurrent dislocation and underwent revision MPFL reconstruction with a palmaris longus tendon autograft at 12 months postoperatively. No cases of infection, wound complications, deep vein thrombosis, or osteotomy union problems were observed.

Table 1. Demographic and clinical data of patients

	Total (n=17)	Group 1 (n=12)	Group 2 (n=5)	p
Age (year), Med. (min; max)	22 (16;39)	21.5 (16;39)	23 (22;24)	0.489[U]
BMI Med. (min;max)	25.3 (18.2;32.5)	23.65 (18.2;32.5)	27.3 (23.7;31.7)	0.065[U]
Follow-up (mounth), Med. (min; max)	32 (15;150)	43 (15;150)	24 (15;32)	**0.020[U]**
Sex, n (%)				
Female	11 (65%)	8 (67%)	3 (60%)	1.000[¥]
Male	6 (35%)	4 (33%)	2 (40%)	
Side, n (%)				
Right	8 (47%)	4 (33%)	4 (80%)	0.131[¥]
Left	9 (53%)	8 (67%)	1 (20%)	
Lateral release, n (%)				
Yes	9 (53%)	4 (33%)	5 (100%)	**0.029[¥]**
No	8 (47%)	8 (67%)	0 (0%)	

[¥]: Fisher's exact test, [U]: Mann-Whitney U test, Med: Median, Min: Minumum, Max: Maximum, BMI: Body mass index

Table 2. Comparison of clinical scores of patients

		Total (n=17)	Group 1 (n=12)	Group 2 (n=5)	p[U]
		Med. (min;max)	Med. (min;max)	Med. (min;max)	
Kujala	Preop	56 (48;64)	56 (48;64)	52 (50;62)	0.133
	Postop	84 (78;90)	82 (78;90)	86 (78;88)	0.789
	[Δ]Med. (min;max)	26 (16;40)	26 (16;40)	32 (26;36)	0.118
	p[w]	**<0.001**	**0.002**	**0.041**	
Lysholm	Preop	50 (48;60)	54 (48;60)	50 (50;52)	0.162
	Postop	90 (80;96)	90 (80;96)	90 (86;96)	0.872
	[Δ] Med. (min;max)	38 (20;46)	38 (20;46)	40 (36;46)	0.425
	p[w]	**<0.001**	**0.002**	**0.041**	
IKDC 2000	Preop	66 (50;78)	64 (50;72)	66 (56;78)	0.672
	Postop	86 (82;92)	86 (82;92)	86 (84;90)	0.788
	[Δ]Med. (min;max)	22 (12;36)	23 (12;36)	22 (12;30)	0.491
	p[w]	**<0.001**	**0.002**	**0.043**	
VAS	Preop	6 (4;8)	6 (4;8)	4 (4;8)	0.360
	Postop	0 (0;4)	1 (0;4)	0 (0;2)	0.634
	[Δ]Med. (min;max)	-4 (-6;-2)	-5 (-6;-2)	-4 (-6;-4)	0.562
	p[w]	**<0.001**	**0.002**	**0.034**	

[U]: Mann-Whitney U test, [w]: Wilcoxon signed ranks test, [Δ]: The amount of postoperative change in score value, Med: Median, Min: Minumum, Max: Maximum, Preop: Preoperative, Postop: Postoperative, VAS: Visual analog scale, IKDC: International knee documentation committee subjective knee form

DISCUSSION

In this study, satisfactory outcomes were obtained in patients who underwent MPFL reconstruction due to recurrent patellar dislocation associated with patellofemoral instability. Clinical scores significantly improved during the postoperative period, and no statistically significant differences were observed between patients who underwent isolated MPFL reconstruction and those who underwent MPFL reconstruction combined with TTO. Significant improvements were observed in all patients' postoperative Kujala, IKDC 2000, and Lysholm scores, while VAS pain scores showed a notable decrease. These results suggest that this technique is effective in enhancing patellofemoral stability and improving patients' functional capacity.

Systematic reviews and meta-analyses specifically addressing this topic indicate that whether MPFL reconstruction is performed alone or in combination with TTO, functional outcomes are generally favorable, and recurrent instability/dislocation rates are low in both techniques.[17,18] Su et al.[18] reported that although recurrent instability rates were similar in patients undergoing MPFL reconstruction with, and without TTO, the mean Kujala score was slightly higher in the isolated MPFL group. In our study, clinical scores were similar in both groups; however, one patient in the MPFL reconstruction with TTO group required revision surgery due to recurrent dislocation. We attribute this to the higher baseline instability in patients undergoing TTO and the increased complexity of the surgical procedure, which may pose challenges during postoperative rehabilitation. Patients undergoing TTO have an increased TT-TG distance, indicating advanced and more complex instability. Moreover, the addition of TTO renders the surgical procedure more complex, and considerations such as waiting for bone union after osteotomy further complicate postoperative follow-up and rehabilitation. Consequently, these patients may be at a higher risk for recurrent dislocation during the postoperative period. Therefore, both the surgeon and the patient must exercise considerable patience and maintain strict control throughout the follow-up and rehabilitation process to achieve a successful and satisfactory outcome.

When reviewing the relevant literature, some studies have reported that patients undergoing MPFL reconstruction with TTO may experience slightly higher rates of surgical complications, particularly related to TTO, such as joint stiffness, implant irritation, and the need for reoperation.[19] In our study, no patient experienced infection, wound dehiscence, deep vein thrombosis, or osteotomy union problems. However, one patient in the TTO group required implant removal due to screw head irritation. We attribute this to the thin skin and subcutaneous fat over the anterior tibial surface. This weakness is particularly pronounced in lean male patients, making the screw heads easily palpable postoperatively. Therefore, patients of this type should be explicitly warned that they may experience discomfort from implant irritation during activities such as kneeling, which could necessitate a secondary procedure (implant removal). Such complications can lead to patient dissatisfaction. In these cases, the surgeon may also consider the use of headless screws as an alternative. To achieve a successful outcome, the surgeon must select an appropriate surgical technique and manage the postoperative process in a controlled manner, maintaining open communication with the patient and proactively addressing potential risks such as screw head irritation.

Study Limitations

This study has some limitations. The most significant of these are its retrospective design and the relatively small sample size. Another limitation is that the study was not conducted on a homogeneous cohort, and there was no control group. However, despite these limitations, the follow-up period ranging from 15 to 150 months allows for preliminary insights into the mid- to long-term outcomes and this could be considered a major strength of the study. In the future, prospective randomized studies on larger, homogeneous patient groups will provide more concrete data.

CONCLUSION

Both isolated MPFL reconstruction and MPFL reconstruction combined with TTO are effective techniques in the surgical management of recurrent patellofemoral dislocation and instability. Satisfactory clinical and functional outcomes can be achieved with either approach. However, it should be noted that isolated MPFL reconstruction may not always be sufficient without patient-specific patoanatomical evaluation. Therefore, preoperative assessment of bony morphological abnormalities and, when appropriate, combining soft tissue procedures with osteotomy is a critical component of a comprehensive surgical strategy.

Ethics

Ethics Committee Approval: This study was approved by the İzmir Katip Çelebi University Health Research Institutional Review Board (decision number: 0037, date: 16.01.2025).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: M.T., Concept: Y.Ö., M.T., Design: Y.Ö., M.T., T.B., Data Collection or Processing: Y.Ö.,

M.T., Analysis or Interpretation: Y.Ö., M.T., Literature Search Y.Ö., M.T., T.B., Writing: Y.Ö., M.T., T.B.

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

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

REFERENCES

1. Gravesen KS, Kallemose T, Blønd L, Troelsen A, Barfod KW. High incidence of acute and recurrent patellar dislocations: a retrospective nationwide epidemiological study involving 24.154 primary dislocations. Knee Surg Sports Traumatol Arthrosc. 2018;26:1204-09.
2. Conlan T, Garth WP Jr, Lemons JE. Evaluation of the medial soft-tissue restraints of the extensor mechanism of the knee. J Bone Joint Surg Am. 1993;75:682-93.
3. Migliorini F, Trivellas A, Colarossi G, Eschweiler J, Tingart M, Rath B. Single- versus double-bundle patellar graft insertion for isolated MPFL reconstruction in patients with patellofemoral instability: a systematic review of the literature. Arch Orthop Trauma Surg. 2020;140:769-76.
4. Migliorini F, Rath B, Tingart M, Meisen N, Eschweiler J. Surgical management for recurrent patellar dislocations in skeletally immature patients. Eur J Orthop Surg Traumatol. 2019;29:1815-22.
5. Dejour H, Walch G, Neyret P, Adeleine P. La dysplasie de la trochlée fémorale [Dysplasia of the femoral trochlea]. Rev Chir Orthop Reparatrice Appar Mot. 1990;76:45-54.
6. Dejour H, Walch G, Nove-Josserand L, Guier C. Factors of patellar instability: an anatomic radiographic study. Knee Surg Sports Traumatol Arthrosc. 1994;2:19-26.
7. Insall J, Goldberg V, Salvati E. Recurrent dislocation and the high-riding patella. Clin Orthop Relat Res. 1972;88:67-9.
8. Smith TO, Walker J, Russell N. Outcomes of medial patellofemoral ligament reconstruction for patellar instability: a systematic review. Knee Surg Sports Traumatol Arthrosc. 2007;15:1301-14.
9. Nomura E, Inoue M, Kobayashi S. Long-term follow-up and knee osteoarthritis change after medial patellofemoral ligament reconstruction for recurrent patellar dislocation. Am J Sports Med. 2007;35:1851-8.
10. Rosso F, Rossi R, Cottino U, Bonasia DE. Tibial tubercle osteotomy for patellofemoral malalignment and chondral disease provided good outcomes: a systematic review. J ISAKOS. 2022;7:78-86.
11. Syed AN, Orellana KJ, Kell D, Dejneka A, Alayleh A, Patel NM, et al. Complication rates following tibial tubercle osteotomy with and without distalization in young patients with patellar ınstability: a multicenter study. J Pediatr Orthop. 2025;45:e506-12.
12. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, et al. Development and validation of the international knee documentation committee subjective knee form. Am J Sports Med. 2001;29:600-13.
13. Kujala UM, Jaakkola LH, Koskinen SK, Taimela S, Hurme M, Nelimarkka O. Scoring of patellofemoral disorders. Arthroscopy. 1993;9:159-63.
14. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am J Sports Med. 1982;10:150-4.
15. Schöttle PB, Schmeling A, Rosenstiel N, Weiler A. Radiographic landmarks for femoral tunnel placement in medial patellofemoral ligament reconstruction. Am J Sports Med. 2007;35:801-4.
16. Fulkerson JP. Anteromedialization of the tibial tuberosity for patellofemoral malalignment. Clin Orthop Relat Res. 1983;176-81.
17. Meng X, Ji Z, Wu P, Fang H, Zhao P, Ding Y, et al. Combining tibial tubercle osteotomy with medial patellofemoral ligament reconstruction often yields better outcomes in treating patellofemoral instability: a systematic review and meta-analysis of case-control studies. J Orthop Surg Res. 2024;19:695.
18. Su P, Yao D, Zhang L, Li G. Results of medial patellofemoral ligament reconstruction with and without tibial tubercle osteotomy in patellar instability: a systematic review and single-arm meta-analysis. BMC Musculoskelet Disord. 2024;25:642.
19. Markes AR, Ghanta RB, Zhang AL, Ma CB, Feeley BT, Lansdown DA. Combined medial patellofemoral ligament reconstruction and tibial tubercle osteotomy has a lower risk of recurrent ınstability requiring revision stabilization at 2 years than either procedure alone. Arthrosc Sports Med Rehabil. 2024;6:100994.

DOI: 10.4274/forbes.galenos.2025.58751
Forbes J Med 2025;6(3):227-236

Original Article / Orijinal Araştırma

Anti-Müllerian Hormone and Insulin-Like Peptide 3 Levels in Adolescent Polycystic Ovary Syndrome: Correlation with Androgen Levels and Ultrasonographic Features

Polikistik Over Sendromlu Adolesanlarda Anti-Müllerian Hormon ve İnsülin Like Peptit 3 Düzeyleri: Androjen Düzeyleri ve Ultrasonografik Bulgularla İlişkisi

Ayça KÖMÜRLÜOĞLU TAN[1,2], E. Nazlı GÖNÇ[3], Z. Alev ÖZÖN[3], Nurgün KANDEMİR[3], H. Nursun ÖZCAN[4], Ayfer ALİKAŞİFOĞLU[3]

[1]Hacettepe University Faculty of Medicine, Department of Pediatrics, Ankara, Türkiye
[2]Sivas Cumhuriyet University Faculty of Medicine, Department of Pediatrics, Sivas, Türkiye
[3]Hacettepe University Faculty of Medicine, Department of Pediatrics, Division of Pediatric Endocrinology, Ankara, Türkiye
[4]Hacettepe University Faculty of Medicine, Department of Radiology, Division of Pediatric Radiology, Ankara, Türkiye

Cite as: Kömürlüoğlu Tan A, Gönç EN, Özön ZA, Kandemir N, Özcan HN, Alikaşifoğlu A. Anti-Müllerian hormone and insulin-like peptide 3 levels in adolescent polycystic ovary syndrome: correlation with androgen levels and ultrasonographic features. Forbes J Med. 2025;6(3):227-236

ABSTRACT

Objective: Polycystic ovary syndrome (PCOS) is a common endocrine disorder characterized by chronic anovulation and hyperandrogenism. In adolescents, the diagnosis is challenging because of overlap with physiological pubertal changes. Anti-Müllerian hormone (AMH) and insulin-like peptide 3 (INSL3) are proposed biomarkers in adults, but their utility in adolescence is unclear. This study investigated serum AMH and INSL3 levels in adolescents with PCOS and healthy controls and assessed associations with ultrasonographic and biochemical features.

Methods: This cross-sectional study included 50 adolescents with PCOS (diagnosed according to the Rotterdam criteria) and 25 healthy controls. Anthropometric measurements, hormonal/biochemical assays, and transabdominal ultrasonography were performed. Correlations between biomarkers, biochemical hyperandrogenism, and ovarian morphology were evaluated. Receiver operating characteristic (ROC) analysis was performed to evaluate the diagnostic performance of AMH.

Results: Adolescents with PCOS had significantly higher serum AMH levels than controls (11.1±5.4 vs. 3.8±1.8 ng/mL; p<0.001), whereas INSL3 levels were similar (p=0.806). AMH correlated positively with total testosterone, androstenedione, free androgen index, Ferriman-Gallwey score, luteinizing hormone/follicle-stimulating hormone ratio, ovarian volume, and antral follicle count, and negatively with sex hormone-binding globulin. In the PCOS group, AMH also correlated with INSL3 (r=0.35, p=0.012), particularly in the overweight/obese subgroups. ROC analysis identified an AMH cut-off value of 5.05 ng/mL, with 94% sensitivity and 80% specificity (area under the curve: 0.938, 95% confidence interval: 0.88-0.99).

Conclusion: Serum AMH is significantly elevated in adolescents with PCOS and strongly correlates with clinical and biochemical hyperandrogenism and with ovarian morphology, suggesting its role as a supportive but not a standalone diagnostic biomarker. Unlike in adult PCOS, INSL3 showed no diagnostic value during adolescence. Incorporating AMH into clinical assessment may help identify adolescents at risk for persistent PCOS features and future metabolic complications.

Keywords: Adolescent, PCOS, AMH, INSL3, polycyctic morphology

Received/Geliş: 30.09.2025
Accepted/Kabul: 17.11.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Ayça KÖMÜRLÜOĞLU TAN, MD, PhD,
Sivas Cumhuriyet University Faculty of Medicine, Department of Pediatrics, Sivas, Türkiye
akomurluoglu@hotmail.com
ORCID: 0000-0002-9034-4848

This study was previously presented as an oral presentation at the 59th Turkish National Pediatrics Congress, Antalya, on November 4-8, 2015. This article is also derived from the pediatric residency thesis of Dr. Ayça Kömürlüoğlu, completed at the Hacettepe University, Department of Pediatrics, in 2015.

ÖZ

Amaç: Polikistik over sendromu (PKOS), kronik anovülasyon ve hiperandrojenizmle karakterize yaygın bir endokrin bozukluktur. Adolesanlarda, fizyolojik pubertal değişikliklerle örtüşmesi nedeniyle tanısı güçtür. Anti-Müllerian hormon (AMH) ve insülin like peptit-3 (INSL3), erişkin PKOS'ta potansiyel biyobelirteçler olarak önerilmiştir, ancak adolesan dönemdeki tanısal değerleri belirsizdir. Bu çalışmada, PKOS tanılı adolesanlarda ve sağlıklı kontrol grubunda serum AMH ve INSL3 düzeyleri araştırılmış, ultrasonografik ve biyokimyasal bulgularla ilişkileri değerlendirilmiştir.

Yöntem: Bu kesitsel çalışmaya, Rotterdam kriterlerine göre PKOS tanısı almış 50 adolesan ve 25 sağlıklı kontrol dahil edildi. Tüm katılımcılarda antropometrik ölçümler, hormonal/biyokimyasal analizler ve transabdominal ultrasonografi gerçekleştirildi. Biyobelirteçler, biyokimyasal hiperandrojenizm ve over morfolojisi arasındaki korelasyonlar değerlendirildi. AMH'nin tanısal performansı için alıcı işletim karakteristiği eğrisi (ROC) analizi yapıldı.

Bulgular: PKOS'lu adolesanlarda serum AMH düzeyleri kontrol grubuna göre anlamlı derecede yüksekti (11,1±5,4 vs. 3,8±1,8 ng/mL, p<0,001), INSL3 seviyeleri ise benzerdi (p=0,806). AMH, total testosteron, androstenedion, serbest androjen indeksi, Ferriman-Gallwey skoru, lüteinizan hormon/folikül uyarıcı hormon oranı, over hacmi ve antral folikül sayısı ile pozitif, seks hormonu bağlayıcı globulin ile negatif korelasyon gösterdi. PKOS grubunda AMH, INSL3 ile de anlamlı bir korelasyon gösterdi (r=0,35, p=0,012); özellikle fazla kilolu/obez alt gruplarında bu ilişki daha belirgindi. ROC analizi, %94 duyarlılık ve %80 özgüllük (eğri altındaki alan: 0,938, %95 güven aralığı: 0,88-0,99) ile 5,05 ng/mL'lik bir AMH kesme değeri belirledi.

Sonuç: Serum AMH, PKOS grubundaki adolesanlarda anlamlı düzeyde yüksektir ve hem klinik/biyokimyasal hiperandrojenizm hem de over morfolojisi ile güçlü bir korelasyon göstermektedir. Bu bulgu, AMH'nın tek başına tanısal bir biyobelirteç olmadığını, ancak PKOS tanısında destekleyici bir rol oynayabileceğini düşündürmektedir. Erişkin PKOS çalışmalarının aksine, adolesan dönemde INSL3 tanısal bir değer göstermemiştir. AMH'nın klinik değerlendirmeye katılması, farklı PKOS fenotiplerinin ve gelecekteki metabolik komplikasyonlar açısından risk altındaki adolesanların belirlenmesine katkı sağlayabilir.

Anahtar Kelimeler: Adolesan, PKOS, AMH, INSL3, polikistik morfoloji

INTRODUCTION

Polycystic ovary syndrome (PCOS) is a widely recognized endocrine disorder resulting from genetic and environmental factors and characterized by diagnostic features, including chronic anovulation and hyperandrogenism confirmed clinically or biochemically. The estimated prevalence among adolescent girls worldwide is approximately 8-10%.[1,2] PCOS is increasingly recognized as a condition associated with long-term metabolic complications, including insulin resistance, type 2 diabetes mellitus, abnormal lipid profiles, and an increased risk of cardiovascular disease. Diagnosing PCOS during adolescence can be difficult because many symptoms and signs of the syndrome may resemble normal pubertal findings.[3]

There are various diagnostic guidelines and criteria. According to the Rotterdam criteria, established in 2003, PCOS is diagnosed if two or more of the following are present: ovulatory dysfunction, clinical and/or biochemical hyperandrogenism, and polycystic ovarian morphology (PCOM).[4] Diagnosis in adolescents is more challenging than in adults because androgens have different reference ranges and there are no universal thresholds; therefore, it should be evaluated according to age and pubertal stage. Physiological multicystic ovarian appearance in adolescence can be confused with PCOM; excessive reliance on pelvic ultrasound (USG) can lead to overdiagnosis. Furthermore, anovulatory cycles in the first years after menarche are physiological, and most resolve with pubertal maturation.[3,5] In 2015, the Pediatric Endocrine Society convened to establish clear criteria for diagnosing adolescent PCOS and highlighted the drawbacks of directly applying adult criteria to adolescents.[6] Diagnostic criteria may lead to unnecessary overdiagnosis, and conversely, ignoring diagnostic features may lead to delayed or missed diagnoses with unfavorable long-term consequences. The 2018 and 2023 international evidence-based PCOS guidelines emphasize the necessity of using a combination of persistent menstrual irregularity and clinical or biochemical hyperandrogenism, while discouraging the use of USG morphology as a diagnostic criterion within eight years of menarche.[7,8] Hyperandrogenism is a required criterion for the diagnosis of PCOS in adolescents. According to current guidelines, biochemical hyperandrogenism should be assessed by measuring total and free testosterone, with the free androgen index (FAI) employed to estimate free testosterone levels.[8]

Anti-Müllerian hormone (AMH) is a glycoprotein of the transforming growth factor-β family, secreted predominantly by granulosa cells in preantral and small antral follicles of the ovaries in women. Because it is stable across the menstrual cycle and directly indicates ovarian reserve, AMH is commonly used in reproductive medicine to evaluate ovarian reserve, predict menopause, guide infertility treatment, and support the diagnosis of PCOS.[9-11] In individuals with PCOS, AMH concentrations are generally elevated, reflecting the increased caunt of preantral and small antral follicles. However, in adolescents, the abundance of physiological follicles and pubertal variability limit its diagnostic utility.

Insulin-like peptide 3 (INSL3), a new member of the insulin/relaxin family, is secreted by theca interna cells within antral follicles and by the corpus luteum and ovarian stromal tissue. It is thought to be associated with follicle number and development.[12] Several studies have demonstrated that INSL3 levels are higher in adults with PCOS and that

they are associated with antral follicle number.[13] However, studies of adolescents are limited, and available evidence suggest that INSL3 does not provide significant diagnostic value for PCOS in this age group.[14]

While current guidelines support the use of PCOM and AMH for diagnosing adult PCOS, using PCOM and AMH is not recommended in the diagnosis of adolescent PCOS.[5,8] However, it is important to identify adolescents who do not fully meet the diagnostic criteria but are predisposed to PCOS and its chronic complications. Our study aimed to evaluate AMH and INSL3 concentrations in adolescents with PCOS and healthy controls, and to compare these concentrations with ultrasonographic and biochemical findings. It was hypothesized that serum AMH and INSL3 would be elevated in adolescents diagnosed with PCOS and would show positive correlations with ultrasonographic findings and androgen levels.

METHODS

Study Population

Fifty adolescent girls aged 12-18 years who were admitted to the Hacettepe University, Division of Pediatric Endocrinology, between July 2014 and March 2015, participated in the study. PCOS was diagnosed in accordance with the Rotterdam consensus, which requires at least two of the three diagnostic criteria: oligo/anovulation, clinical or biochemical hyperandrogenism, and PCOM.[4] Participants with comorbidities, such as chronic systemic disease, thyroid abnormalities, hyperprolactinemia, congenital adrenal hyperplasia, androgen-secreting tumors, Cushing's syndrome, or other endocrine disorders, or with prior treatment for PCOS, were excluded.

Twenty-five healthy adolescent girls with regular menstrual cycles for at least two years were selected as controls. Menstrual cycles were considered eumenorrheic when the interval ranged between 21 and 45 days, oligomenorrheaic when the interval exceeded 45 days or there were fewer than eight menstrual episodes within a year, and amenorrheaic when the absence of menses persisted more than three months.[5] Only participants who were at least two years post-menarche were enrolled in the study. Age at menarche and family history of PCOS were recorded. Written informed consent was obtained from both the adolescents and their parents.

Clinical Evaluation

Anthropometric Evaluation

Weight and height assessments were carried out by a single physician (A.K.T) using a calibrated wall-mounted Harpenden stadiometer, and an electronic scale (sensitivity 0.1 kg) after 12 hours of fasting, barefoot, and wearing normal clothing. The body mass index (BMI) was calculated using the formula: weight (kg)/height^2 (m^2). Participants were grouped by BMI percentile: 5th-85th percentile as normal weight, 85th-95th percentile as overweight, and ≥95th percentile as obese. Standard deviation scores (SDS) for anthropometric parameters were computed.[15] Patients were only eligible for enrollment if at least two years had elapsed since menarche. A detailed medical history and systemic physical examination were obtained from all participants by the same physician (A.K.T).

Assessment of Clinical Hyperandrogenism

Clinical hyperandrogenism was assessed according to the Ferriman-Gallwey scoring system. A score of ≥8 was accepted as clinical hyperandrogenism.[16]

Laboratory Evaluation

Blood Sampling

Blood samples were collected from all participants on days 2-3 of the menstrual cycle, corresponding to the early follicular phase. In patients with amenorrhea, samples were collected after a 12-hour overnight fast on the day of the clinical evaluation at 08:00 AM. Serum was promptly separated, frozen, and preserved at -80 °C until analysis.

Biochemical Analyses

Serum concentrations of follicle-stimulating hormone (FSH), luteinizing hormone (LH), estradiol (E2), total testosterone, sex hormone-binding globulin (SHBG), dehydroepiandrosterone sulfate (DHEAS), androstenedione, 17-hydroxyprogesterone (17OHP), AMH, and INSL3 were measured in all patients with PCOS and in controls. To exclude differential diagnoses, basal 17OHP levels were assessed for adrenal enzymatic abnormalities, free thyroxine (T4) and thyroid-stimulating hormone (TSH) for thyroid dysfunction, and prolactin for hyperprolactinemia. Serum FSH, LH, E2, prolactin, TSH, and free T4 were measured using a two-step chemiluminescence microparticle immunoassay; SHBG and 17OHP by immunoradiometric assay; DHEAS and total testosterone by solid-phase chemiluminescence immunoassay; and AMH and INSL3 by enzyme-linked immunosorbent assay (Beckman Coulter®, Webster USA; Elabscience®, USA)

FAI was calculated using the formula: [total testosterone (nmol/L) ÷ SHBG (nmol/L)] × 100. Serum testosterone levels measured in ng/dL were converted to nmol/L using the conversion factor 1 ng/dL = 0.0347 nmol/L.

Pelvic Ultrasound

On the day of hormonal and biochemical assessments, pelvic USG examinations were also conducted.

Prospective transabdominal USG scans were performed in both the PCOS and control groups by a pediatric radiologist (H.N.Ö), who was blinded to the clinical and laboratory information of the participants. Imaging was conducted using a Sonoline Elegra US scanner (Siemens, Erlangen, Germany) with a 2-5 MHz convex-array broadband transducer in the department of radiology.

For each subject, uterine dimensions, endometrial thickness, and ovarian dimensions [longitudinal, transverse, and anteroposterior (AP) diameters] were measured in the sagittal and coronal planes. Uterine and ovarian volumes were computed according to the ellipsoid formula: length × AP diameter × transverse diameter × π/6. The follicles were grouped by size into three categories: <5 mm, 6-10 mm, and >10 mm. Follicle counts were documented for each group. Antral follicles (<10 mm) were counted, and cysts (>10 mm) were noted. PCOM was described by the presence of ≥12 follicles measured <10 mm, and/or ovarian volume >10 mL.[4,17]

Statistical Analysis

Statistical analyses were performed using SPSS version 21 (Chicago, IL, USA). Data are given as mean ± standard deviation (SD) or as median (minimum-maximum). Normality was examined with the Kolmogorov-Smirnov test. Categorical data are expressed as percentages (%). Group comparisons were made using Student's t-test or Mann-Whitney U test. Categorical data were analyzed with the chi-square test. Correlations were assessed using Pearson's correlation for normally distributed variables and Spearman's rank correlation when at least one variable was not normally distributed. A p-value <0.05 was accepted as statistically significant.

Biomarkers demonstrating significant differences between groups were further evaluated using receiver operating characteristic (ROC) curve analysis to determine their diagnostic values (sensitivity and specificity). An area under the curve (AUC) was considered statistically significant at a type 1 error level of <5%.

Ethical Approval

The study was approved by the Ethics Committee of the Faculty of Medicine at Hacettepe University (approval no: GO 14/256-12, date: 04.06.2014) Written informed consent was obtained from all participants and their parents.

RESULTS

50 adolescents with PCOS and 25 healthy controls were included in the analysis. The mean age and age at menarche were similar among the groups (p>0.05). Compared to controls, adolescents with PCOS had significantly higher weight SDS (p=0.003) and BMI SDS (p<0.001), whereas height SDS were similar between groups (Table 1). All participants in the control group had normal BMI values, whereas approximately half of the adolescents in PCOS group were overweight or obese.

Serum LH concentrations and the LH/FSH ratio were significantly elevated in adolescents with PCOS (LH: 10.47±6.29 vs. 6.53±4.98 mIU/mL, p=0.008; LH/FSH ratio: 1.83±1.07 vs. 1.22±1.11, p=0.025), whereas FSH levels were similar (p=0.454). The PCOS group had significantly elevated levels of serum total testosterone (64.04±21.43 vs. 27.91±9.64 ng/mL; p<0.001), DHEAS (275.66±114.35 vs. 170.09±89.98 µg/dL; p<0.001), and androstenedione (4.65±2.36 vs. 1.93±0.79 ng/mL; p<0.001) compared with controls. Additionally, the FAI was significantly elevated in PCOS cases (p<0.001). Serum SHBG and E2 concentrations were similar between groups. 17OHP levels were also significantly elevated in adolescents with PCOS (p<0.001) (Table 1).

Serum AMH levels were significantly higher in the PCOS group (11.1±5.42 ng/mL) than in the controls (3.76±1.75 ng/mL) (p<0.001). INSL3 levels were similar across groups (p=0.806). Anthropometric, clinical, and laboratory data of the PCOS and control groups are presented in Table 1. The mean ovarian volume was significantly greater in the PCOS group compared with the control group (13.99±5.22 vs. 9.70±3.64 mL, p<0.001). Both the right and left ovarian volumes were significantly higher in adolescents with PCOS (right: 14.76±6.69 vs. 10.60±4.78 mL, p=0.003; left: 13.22±5.29 vs. 8.75±3.80 mL, p<0.001). The number of follicles <5 mm and the overall antral follicle number were significantly higher in PCOS cases than in controls (both p<0.001) (Table 2). Adolescents with amenorrhea in the PCOS group had significantly higher antral follicle counts (AFCs) than those without amenorrhea (p<0.001).

Normal ovarian morphology was observed in 5 (10%) of the PCOS group, whereas PCOM was observed in 3 (12%) of the control group. Ultrasonographic features of patients with PCOS and controls are summarized in Table 2.

Correlation Analysis

Correlation analysis showed significant positive correlations between serum AMH concentrations and LH (r=0.462, p<0.001), LH/FSH ratio (r=0.448, p<0.001), total testosterone (r=0.590, p<0.001), androstenedione (r=0.524, p<0.001), FAI (r=0.315, p=0.006), Ferriman-Gallwey score (r=0.490, p<0.001), total ovarian volume (r=0.496, p<0.001), and AFC (r=0.620, p<0.01). In contrast, AMH was negatively correlated with SHBG (r=-0.241, p=0.037).

No significant correlations were detected between INSL3 and other hormonal and ultrasonographic parameters. However, in the PCOS group, AMH showed a moderate

positive correlation with INSL3 (r=0.35, p=0.012). Subgroup analysis according to BMI demonstrated that this correlation was stronger in overweight patients (r=0.532, p=0.050) and obese patients (r=0.595, p=0.032).

Diagnostic Value of Serum Anti-Müllerian Hormone Level

ROC curve analysis was used to assess the diagnostic ability of AMH in adolescents with PCOS (Figure 1). At a cut-off level of 5.05 ng/mL, AMH demonstrated a sensitivity

Table 1. Anthropometric, clinical, and laboratory characteristics of adolescents with PCOS and control group

Parameter	PCOS (n=50)	Controls (n=25)	p
Age (years)	15.9±1.13	15.7±1.21	0.362
Age at menarche (years)	12.2±0.11	12.4±1.25	0.546
Weight (kg)	66.76±17.02	54.44±7.6	<0.001
Weight SDS	0.77±1.18	-0.02±0.79	0.003
Height (cm)	160.1±6.3	161.7±7.16	0.342
Height SDS	-0.32±0.93	-0.74±1.07	0.251
BMI (kg/m^2)	25.9±6.08	20.7±2.23	<0.001
BMI SDS	1.16±1.74	-0.31±1.12	<0.001
Ferriman-Gallwey score	16.14±6.7	4.92±2.06	<0.001
FSH (mIU/mL)	6.11±2.39	5.72±1.99	0.454
LH (mIU/mL)	10.47±6.29	6.53±4.98	0.008
LH/FSH ratio	1.83±1.07	1.22±1.11	0.025
Total testosterone (ng/dL)	64.04±21.43	27.91±9.64	<0.001
DHEAS (µg/dL)	275.66±114.35	170.09±89.98	<0.001
Androstenedione (ng/mL)	4.65±2.36	1.93±0.79	<0.001
SHBG (nmol/L)	37.98±36.41	45.30±18.57	0.253
Estradiol (pg/mL)	50±30	48±30	0.798
17OHP (ng/mL)	1.47±0.73	0.89±0.44	<0.001
FAI	11.20±10.51	2.98±2.96	<0.001
AMH (ng/mL)	11.1±5.42	3.76±1.75	<0.001
INSL3 (ng/mL)	0.23±0.18	0.22±0.15	0.806

Values are presented as mean ± standard deviation. DHEAS: Dehydroepiandrosterone sulfate, SHBG: Sex hormone-binding globulin, 17OHP: 17-hydroxyprogesterone, AMH: Anti-Müllerian hormone, INSL3: Insulin-like peptide-3, PCOS: Polycystic ovary syndrome, BMI: Body mass index, SDS: Standard deviation scores, FSH: Follicle-stimulating hormone, LH: Luteinizing hormone. FAI: Free androgen index, calculated as [total testosterone (nmol/L) ÷ SHBG (nmol/L)] × 100

Table 2. USG features of PCOS and control group

Parameter	PCOS (n=50)	Controls (n=25)	p
Mean ovarian volume (mL)	13.99±5.22	9.70±3.64	<0.001
Left ovarian volume (mL)	13.22±5.29	8.75±3.80	<0.001
Right ovarian volume (mL)	14.76±6.69	10.60±4.78	0.003
Endometrial thickness (mm)	6.63±2.39	5.78±3.04	0.226
Number of follicles <5 mm (count, per ovary)	10.00±4.42	3.14±2.76	<0.001
Number of follicles 6-10 mm (count, per ovary)	3.42±4.66	4.16±2.22	0.450
Antral follicles <10 mm (count, per ovary)	13.42±4.6	7.30±2.28	<0.001

Data are presented as mean ± standard deviation. "Per-ovary" mean = (right + left)/2
PCOS: Polycystic ovary syndrome, USG: Ultrasound

of 94% and a specificity of 80%. The AUC was 0.938 [95% confidence interval (CI): 0.88-0.99, p<0.001].

DISCUSSION

PCOS is one of the most frequently diagnosed endocrine disorders in adolescent females and continues to pose a diagnostic challenge owing to the overlap between physiological pubertal changes and the early manifestations of the syndrome.[1] In our adolescent cohort diagnosed with PCOS according to the Rotterdam criteria, serum AMH levels were significantly elevated compared with those of their healthy peers. This finding, although not included in current diagnostic criteria, highlights the potential role of AMH as a supportive marker of PCOS in adolescents.

INSL3 and AMH are produced by granulosa and theca cells in mammalian gonads.[10,12] In relation with the elevated preantral and antral follicle caunts in PCOS, studies showed that AMH is widely reported in individuals with PCOS, especially in adults.[14,18,19] In line with our results, previous studies have consistently reported elevated AMH concentrations in adolescents with PCOS. In a recent meta-analysis including 15 studies, Tsukui et al.[20] showed that serum AMH concentrations were significantly higher in adolescents with PCOS than in controls, with an average

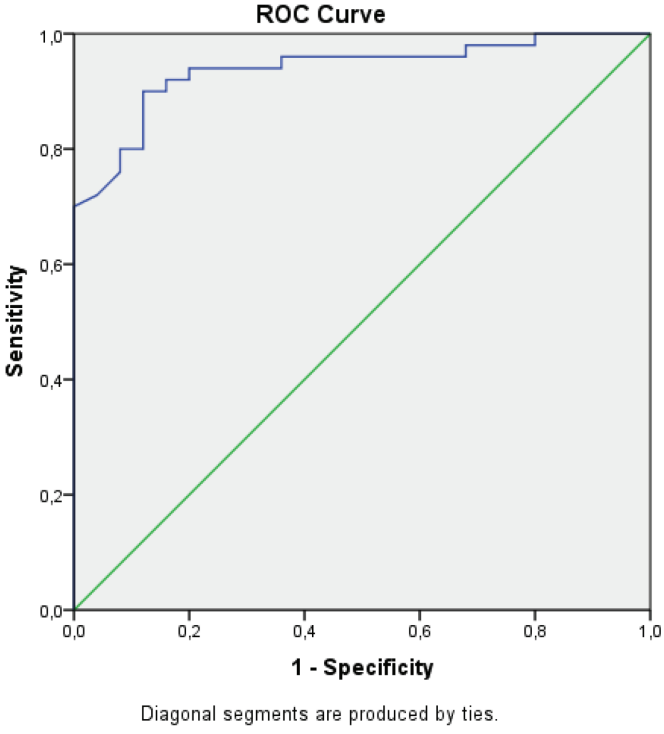


Figure 1. ROC curve of serum anti-Müllerian hormone levels in adolescents with polycystic ovary syndrome. The curve plots sensitivity (y-axis) against 1-specificity (x-axis)

ROC: Receiver operating characteristic

weighted difference of 2.91 ng/mL (95% CI: 0.74-5.09). In the same study, AFC was also significantly higher in the PCOS group, with a mean difference of 7.14 follicles (95% CI: 2.70-11.59). In our study, a cut-off level of 5.05 ng/mL was determined for serum AMH in adolescents with PCOS, demonstrating 94% sensitivity and 80% specificity. An examination of cut-off values reported in adolescent PCOS studies in the literature reveals that they are generally within the range of 5.8-7.25 ng/mL, although some studies have reported higher (≥8-10 ng/mL) or lower values (Table 3).[14,21-29] Although the cut-off value obtained in this study is at the lower end of this range, it is particularly notable for its high sensitivity. Differences between studies are thought to result from measurement kits, variations in diagnostic criteria, sample characteristics (obesity, time since menarche, ethnic distribution), and study designs. In contrast to our findings, Kocaay et al.[30] found no difference in AMH levels among adolescent PCOS cases and the control group. Despite these differences, the findings suggest that AMH is a valuable complementary marker for diagnosing PCOS in adolescents and may increase its diagnostic power, especially when evaluated together with androgen levels and ultrasonographic findings.

Data in the literature on the relationship between serum AMH and androgen levels are contradictory. Although the majority of studies exhibited a positive correlation among AMH and androgen concentrations, some other studies had shown no correlation.[31-34] In this study, AMH was significantly correlated with androgens (including total testosterone and androstenedione), as well as with FAI and the Ferriman-Gallwey score, suggesting that AMH may also reflect the degree of clinical and biochemical hyperandrogenism in adolescents with PCOS.

The FAI, defined as total testosterone divided by SHBG, provides an estimate of free testosterone. It is recommended by current guidelines as a diagnostic marker for biochemical hyperandrogenism.[7] In their study on adolescents, Sağsak et al.[35] showed that individuals with an FAI above 6.15 warrant evaluation for PCOS. Villarroel et al.[36] study on adolescent PCOS also found that an FAI ≥6.1% is valuable for diagnosing PCOS. On the other hand Özer et al.[37] found that FAI measurements were found similar among adolescents with PCOS and hyperinsulinemia/obesity, while SHBG levels were lower in hyperinsulinemia/obesity adolescents with oligomenorrhea. These findings suggest that neither FAI nor SHBG alone is a reliable diagnostic marker for PCOS in the presence of metabolic disorders. Several studies have also demonstrated the diagnostic utility of FAI, reporting higher levels in adolescents with PCOS compared to healthy controls.[14,28] In our study, no difference in SHBG levels was identified among the PCOS group and controls,

Table 3. AMH cut-off values for adolescent PCOS determined by ROC analysis in the literature

Author-year (ref)	AMH cut-off (ng/mL)	AUC	95% CI	Sensitivity	Specificity	p value
Li et al.[21]	8 ng/mL	0.664	0.551-0.778	70%	61.7%	
Deveer et al.[22]	6.66 ng/mL	0.820		62%	76%	
Tokmak et al.[23]	7.11 ng/mL	0.763	0.607-0.920	84%	66.7%	0.004
Yetim et al.[14]	6.1 ng/mL	0.880	0.800-0.960	81.1%	92.3%	<0.001
Savas-Erdeve et al.[24]	7.25 ng/mL	0.700	0.591-0.808	58%	72.5%	0.001
Kim et al.[25]	6.26 ng/mL	0.788	0.687-0.868	67%	81%	<0.001
Merino et al.[26]	7.03 ng/mL	0.758		58.8%	82.1%	0.0001
Yetim Şahin et al.[27]	10 ng/mL	0.326	0.171-0.482	50%	69%	0.047
Khashchenko et al.[28]	7.20 ng/mL			76%	89%	
Tunç and Özkan[29]	5.8 ng/mL			86.9%	70%	
Kömürlüoğlu Tan (present study)	5.05 ng/mL	0.938	0.880-0.990	94%	80%	<0.001

AMH: Anti-Müllerian hormone, AUC: Area under the curve, CI: Confidence interval, ROC: Receiver operating characteristic, PCOS: Polycystic ovary syndrome

whereas FAI was significantly elevated in adolescents in the PCOS group compared with controls. These results are consistent with the literature and suggest that FAI can be used as a supportive marker in adolescent PCOS. However, due to differences in pubertal stage, obesity, and measurement methods, FAI alone is not diagnostic and should be interpreted in conjunction with other clinical and biochemical findings.

Özer et al.[37] found the LH/FSH ratio was also determined as a predictive marker for PCOS. Previous studies have consistently reported elevated LH/FSH ratios in adolescents diagnosed PCOS, supporting our findings of higher ratios in this population.[14,28] In our cohort, the LH/FSH ratio was significantly elevated in PCOS cases than in controls, and AMH was correlated positively with the LH/FSH ratio ($r=0.448$; $p<0.01$). Since the LH/FSH ratio is variable during adolescence, it does not provide diagnosis alone; it is best to interpret this ratio together with high FAI and ovarian USG findings or AMH.

Previous studies have reported a potential role for INSL3 in the pathogenesis of PCOS. Data on INSL3 levels in adolescent PCOS are limited compared with those from studies in adults. In adult women with PCOS, elevated INSL3 concentrations have consistently been reported, reflecting increased theca cell activity.[12] In this study, Gambineri et al.[12] also reported significantly higher INSL3 levels in adult women with PCOS, with positive correlations to androgen concentrations. Similarly, Pelusi et al.[38] found that INSL3 was associated with androgen levels in adolescents with anovulatory cycles. These findings indicate that INSL3 may reflect ovarian steroidogenic activity and hyperandrogenism at an early stage. However, in our cohort, INSL3 levels did not differ among adolescents with PCOS and controls, and were not associated with ultrasonographic features or androgen levels, suggesting that its diagnostic utility in adolescence remains limited. Similar to our results Tunç and Özkan[29] and Yetim et al.[14] found INSL3 levels to be similar to the control group in their study on adolescent PCOS patients, and did not detect any correlation with any laboratory or clinical parameters. These discrepancies between adolescent and adult populations may be due to theca cell activity and INSL3 secretion not yet being fully established during early adolescence, resulting in a weaker association with androgenic and morphological traits. Furthermore, differences in analysis methodologies and sample sizes across studies may contribute to the inconsistent findings. When considered together, these findings indicate that INSL3 may be a biomarker of hyperandrogenism in adults, but its role in adolescents appears less reliable and requires further investigation in larger longitudinal studies.

Pelusi et al.[13] demonstrated a significant association between AMH and INSL3 levels in their study of adult women with PCOS, particularly among those with amenorrhea. Similarly, we found a significant correlation between AMH and INSL3 in adolescents with PCOS; this association was particularly evident in overweight/obese subgroups. This may reflect the combined effect of metabolic and gonadal factors on ovarian function and suggest that body composition may modulate the relationship between AMH and INSL3. These results support the hypothesis that INSL3 and AMH are co-regulated and suggest that they provide a complementary perspective on the pathophysiology of PCOS.

Many studies of adult women have shown a positive association among serum AMH levels and follicle counts in PCOS patients, reflecting increased AMH production

resulting from the elevated preantral and AFCs in polycystic ovaries.[39,40] Findings in adolescents have been less consistent. Pawelczak et al.[34] demonstrated a positive association among AMH, ovarian volume and peripheral follicle distribution. Similary, Villarroel et al.[36] noted a positive link between among AMH and the caunt of 2-5 mm follicles. In contrast, Yetim et al.[14] showed no correlation among AMH and ultrasonographic features, and Savas-Erdeve et al.[24] also documented no significant association among AMH, ovarian volume, and number of follicles. In our study, serum AMH was positively correlated with both ovarian volume and antral follicle number in adolescent girls with PCOS, supporting its role as a marker reflecting ovarian morphology in this population.

Follicle counts are physiologically high and variable in adolescents, so they are unreliable in diagnosing PCOS. Current guidelines do not recommend pelvic USG for diagnostic purposes before 8 years postmenarche.[7] If USG is used during this period, ovarian volume should be the preferred morphological criterion rather than the follicle count. Previous studies have reported that mean ovarian volumes in adolescent patients with PCOS were higher than those in healthy individuals.[24,28] In our cohort, mean ovarian volumes were significantly greater in adolescents with PCOS than in controls (right ovary: 14.8 vs. 10.6 cm^3; left ovary: 13.2 vs. 8.8 cm^3), supporting ovarian volume as a more informative morphological indicator in this age group.

The relationship between obesity and serum AMH levels is controversial. Obesity may affect serum AMH concentrations through several metabolic and inflammatory mechanisms. Studies in adult women with PCOS have reported lower AMH levels due to leptin elevation, insulin resistance, and impaired follicular development secondary to chronic inflammation.[41,42] In an adolescent study conducted in our country, normal-weight adolescents with PCOS exhibited higher AMH levels than their overweight or obese peers.[43] Conversely, other studies have reported elevated AMH levels in obese adolescents with PCOS.[25,44] In our study, no significant differences in AMH levels were found among normal-weight, overweight, and obese participants. These conflicting findings suggest that the relationship among obesity and AMH is complex and can be affected by age-related, environmental, hormonal, and metabolic factors.

Although AMH is not currently recommended as a diagnostic criterion in adolescents due to test variability and the physiological changes of puberty, our findings highlight its potential clinical utility and suggest that AMH can be considered a supportive criterion for PCOS. In our cohort, AMH showed strong correlations not only with AFC and ovarian volume, but also with biochemical and clinical hyperandrogenism (testosterone, androstenedione levels, FAI, and Ferriman-Gallwey score) and with the LH/FSH ratio. These findings suggest that AMH reflects both the morphological and endocrine features of PCOS. The high sensitivity of the 5.05 ng/mL cut-off value suggests that AMH may serve as a supportive biomarker for identifying adolescents at risk of persistent PCOS features. It has also been shown in the literature that high AMH in adolescence may be associated with future development of PCOS.[31] However, due to its limited specificity and lack of test standardization, AMH should not be used alone but in conjunction with established clinical and biochemical criteria.

Study Limitations

This study is valuable because it enables objective comparison of findings between adolescents with PCOS and healthy controls using standardized pelvic USG performed by a blinded pediatric radiologist. Simultaneously measured AMH, INSL3, gonadotropins, and androgen levels were correlated with detailed ultrasonographic measurements. We demonstrated robust diagnostic performance of AMH in adolescents (cut-off point of 5.05 ng/mL; AUC of 0.938; sensitivity of 94%; specificity of 80%) and consistent positive correlations between AMH and biochemical and clinical hyperandrogenism (total testosterone, androstenedione, FAI, Ferriman-Gallwey score), the LH/FSH ratio, and mean ovarian volumes. In contrast to data from adults, we found that INSL3 did not distinguish adolescents with PCOS from controls. These results indicate that biomarker levels and their interpretation, in addition to hyperandrogenism and oligo/amenorrhea criteria, can contribute significantly to the clinical decision-making process for diagnosing adolescent PCOS.

This study has some limitations. First, applying the Rotterdam criteria derived from adults to adolescents carries the risk of misclassification because anovulatory cycles are common in the early postmenarche period, and physiologically high or variable follicle counts may exaggerate polycystic morphology, potentially leading to overdiagnosis. Transabdominal USG, used because transvaginal USG, is unsuitable for virgin adolescents, may underestimate follicle count; to mitigate this limitation, we prioritized ovarian volume as a morphological descriptor. The limited sample size and the cross-sectional, single-center design restrict causal inference and the generalizability of the findings. Finally, we were unable to perform longitudinal follow-up to determine the transition of adolescent phenotypes into adulthood and did not consider all metabolic confounders (e.g., insulin resistance) in subgroup analyses.

CONCLUSION

In this adolescent cohort, serum AMH levels were significantly elevated in the PCOS group compared with controls, and correlated with ovarian morphology (ovarian volume, AFC), biochemical/clinical hyperandrogenism (total testosterone, androstenedione, FAI, Ferriman-Gallwey score), and the LH/FSH ratio. FAI was significantly elevated in PCOS, whereas INSL3 did not distinguish adolescents with PCOS from controls. These findings support the use of AMH as an adjunct test in the diagnosis of PCOS in adolescence, rather than as a standalone parameter. AMH may be particularly helpful for identifying at-risk adolescents prior to a definitive PCOS diagnosis and for the early stratification of future metabolic risk. Given test variability and limited specificity, AMH should be considered together with clinical data and other biochemical assessments. Larger, prospective multicenter studies are required to biomarkers in adolescent PCOS cases and to evaluate assess the long-term metabolic consequences of these biomarkers.

Ethics

Ethics Committee Approval: The study was approved by the Ethics Committee of the Faculty of Medicine at Hacettepe University (approval no: GO 14/256-12, date: 04.06.2014).

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

Footnotes

Authorship Contributions

Concept: A.K.T., E.N.G., A.A., Design: A.K.T., Z.A.Ö., N.K., A.A., Data Collection or Processing: A.K.T., H.N.Ö., Analysis or Interpretation: A.K.T., E.N.G., Z.A.Ö., N.K., A.A., Literature Search: A.K.T., H.N.Ö., A.A., Writing: A.K.T.

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

Financial Disclosure: The study was supported by the Hacettepe University Scientific Research Projects Coordination Unit (Project ID: 668, Project Code: 014 D07 101 004-668).

REFERENCES

1. Ibáñez L, de Zegher F. Polycystic ovary syndrome in adolescent girls. Pediatr Obes. 2020;15:e12586.
2. Teede HJ, Misso ML, Costello MF, et al; International PCOS Network. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Hum Reprod. 2018;33:1602-18.
3. Ibáñez L, Oberfield SE, Witchel S, et al. An international consortium update: pathophysiology, diagnosis, and treatment of polycystic ovarian syndrome in adolescence. Horm Res Paediatr. 2017;88:371-95.
4. Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). Hum Reprod. 2004;19:41-7.
5. Peña AS, Witchel SF, Hoeger KM, et al. Adolescent polycystic ovary syndrome according to the international evidence-based guideline. BMC Med. 2020;18:72.
6. Witchel SF, Oberfield S, Rosenfield RL, et al. The Diagnosis of Polycystic Ovary Syndrome during Adolescence. Horm Res Paediatr. 2015.
7. Peña AS, Witchel SF, Boivin J, et al. International evidence-based recommendations for polycystic ovary syndrome in adolescents. BMC Med. 2025;23:151.
8. Teede HJ, Tay CT, Laven JJE, et al. Recommendations from the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome. J Clin Endocrinol Metab. 2023;108:2447-69.
9. Iwase A, Hasegawa Y, Tsukui Y, et al. Anti-Müllerian hormone beyond an ovarian reserve marker: the relationship with the physiology and pathology in the life-long follicle development. Front Endocrinol (Lausanne). 2023;14:1273966.
10. Weenen C, Laven JS, Von Bergh AR, et al. Anti-Müllerian hormone expression pattern in the human ovary: potential implications for initial and cyclic follicle recruitment. Mol Hum Reprod. 2004;10:77-83.
11. Leader B, Baker VL. Maximizing the clinical utility of antiMüllerian hormone testing in women's health. Curr Opin Obstet Gynecol. 2014;26:226-36.
12. Gambineri A, Patton L, De lasio R, Palladoro F, Pagotto U, Pasquali R. Insulin-like factor 3: a new circulating hormone related to luteinizing hormone-dependent ovarian hyperandrogenism in the polycystic ovary syndrome. J Clin Endocrinol Metab. 2007;92:2066-73.
13. Pelusi C, Fanelli F, Pariali M, Zanotti L, Gambineri A, Pasquali R. Parallel variations of insulin-like peptide 3 (INSL3) and antimüllerian hormone (AMH) in women with the polycystic ovary syndrome according to menstrual cycle pattern. J Clin Endocrinol Metab. 2013;98:E1575-82.
14. Yetim A, Yetim Ç, Baş F, et al. Anti-Müllerian hormone and inhibin-a, but not inhibin-b or insulin-like peptide-3, may be used as surrogates in the diagnosis of polycystic ovary syndrome in adolescents: preliminary results. J Clin Res Pediatr Endocrinol. 2016;8:288-97.
15. Neyzi O, Bundak R, Gökçay G, et al. Reference values for weight, height, head circumference, and body mass index in Turkish children. J Clin Res Pediatr Endocrinol. 2015;7:280-93.
16. Ferriman D, Gallwey JD. Clinical assessment of body hair growth in women. J Clin Endocrinol Metab. 1961;21:1440-7.
17. Fulghesu AM, Canu E, Casula L, Melis F, Gambineri A. Polycystic ovarian morphology in normocyclic non-hyperandrogenic adolescents. J Pediatr Adolesc Gynecol. 2021;34:610-6.
18. Dewailly D, Gronier H, Poncelet E, et al. Diagnosis of polycystic ovary syndrome (PCOS): revisiting the threshold values of follicle count on ultrasound and of the serum AMH level for the definition of polycystic ovaries. Hum Reprod. 2011;26:3123-9.
19. Woo HY, Kim KH, Rhee EJ, Park H, Lee MK. Differences of the association of anti-Müllerian hormone with clinical or biochemical characteristics between women with and without polycystic ovary syndrome. Endocr J. 2012;59:781-90.
20. Tsukui Y, Kitahara Y, Hasegawa Y, Kobayashi M, Osuka S, Iwase A. Anti-Müllerian hormone levels in the diagnosis of adolescent

polycystic ovarian syndrome: a systematic review and meta-analysis. Endocr J. 2022;69:897-906.

21. Li L, Chen X, Mo Y, Chen Y, Wenig M, Yang D. Elevated serum antiMullerian hormone in adolescent and young adult Chinese patients with polycystic ovary syndrome. Wien Klin Wochenschr. 2010;122:519-24.
22. Deveer M, Deveer R, Basaran O, et al. Serum copeptin, pentraxin 3, antiMüllerian hormone levels with echocardiography and carotid artery intima-media thickness in adolescents with polycystic ovary syndrome. J Clin Med Res. 2015;7:989-94.
23. Tokmak A, Kokanali D, Timur H, Kuntay Kokanali M, Yilmaz N. Association between antiMüllerian hormone and insulin resistance in non-obese adolescent females with polycystic ovary syndrome. Gynecol Endocrinol. 2016;32:926-30.
24. Savas-Erdeve S, Keskin M, Sagsak E, Cenesiz F, Cetinkaya S, Aycan Z. Do the Anti-Müllerian hormone levels of adolescents with polycystic ovary syndrome, those who are at risk for developing polycystic ovary syndrome, and those who exhibit isolated oligomenorrhea differ from those of adolescents with normal menstrual cycles? Horm Res Paediatr. 2016;85:406-11.
25. Kim JY, Tfayli H, Michaliszyn SF, Lee S, Nasr A, Arslanian S. Anti-Müllerian hormone in obese adolescent girls with polycystic ovary syndrome. J Adolesc Health. 2017;60:333-9.
26. Merino PM, Villarroel C, Jesam C, López P, Codner E. New diagnostic criteria of polycystic ovarian morphology for adolescents: impact on prevalence and hormonal profile. Horm Res Paediatr. 2017;88:401-7.
27. Yetim Şahin A, Baş F, Yetim Ç, et al. Determination of insulin resistance and its relationship with hyperandrogenemia,anti-Müllerian hormone, inhibin A, inhibin B, and insulin-like peptide-3 levels in adolescent girls with polycystic ovary syndrome. Turk J Med Sci. 2019;49:1117-25.
28. Khashchenko E, Uvarova E, Vysokikh M, et al. The relevant hormonal levels and diagnostic features of polycystic ovary syndrome in adolescents. J Clin Med. 2020;9:1831.
29. Tunç S, Özkan B. Analysis of new biomarkers for the diagnosis of polycystic ovary syndrome in adolescents. J Curr Pediatr. 2021;19:311-8.
30. Kocaay P, Siklar Z, Buyukfirat S, Berberoglu M. The diagnostic value of anti-Müllerian hormone in early post menarche adolescent girls with polycystic ovarian syndrome. J Pediatr Adolesc Gynecol. 2018;31:362-6.
31. Pinola P, Morin-Papunen LC, Bloigu A, et al. Anti-Müllerian hormone: correlation with testosterone and oligo- or amenorrhoea in female adolescence in a population-based cohort study. Hum Reprod. 2014;29:2317-25.
32. Piouka A, Farmakiotis D, Katsikis I, Macut D, Gerou S, Panidis D. AntiMüllerian hormone levels reflect severity of PCOS but are negatively influenced by obesity: relationship with increased luteinizing hormone levels. Am J Physiol Endocrinol Metab. 2009;296:E238-43.
33. Siow Y, Kives S, Hertweck P, Perlman S, Fallat ME. Serum Müllerian-inhibiting substance levels in adolescent girls with normal menstrual cycles or with polycystic ovary syndrome. Fertil Steril. 2005;84:938-44.
34. Pawelczak M, Kenigsberg L, Milla S, Liu YH, Shah B. Elevated serum anti-Müllerian hormone in adolescents with polycystic ovary syndrome: relationship to ultrasound features. J Pediatr Endocrinol Metab. 2012;25:983-9.
35. Sağsak E, Keskin M, Çetinkaya S, Erdeve ŞS, Aycan Z. The diagnostic value of free androgen index in obese adolescent females with idiopathic hirsutism and polycystic ovary syndrome. J Acad Res Med. 2021;11:81-5.
36. Villarroel C, Merino PM, López P, et al. Polycystic ovarian morphology in adolescents with regular menstrual cycles is associated with elevated antiMüllerian hormone. Hum Reprod. 2011;26:2861-8.
37. Özer E, Taş D, Çakır Gündoğan S, Boyraz M, Gürbüz F. Clinical utility of FAI and SHBG in differentiating PCOS from anovulatory cycles in adolescent girls. Front Pediatr. 2025;13:1581060.
38. Pelusi C, Stancampiano M, Fanelli F, Pariali M, Gambineri A, Pasquali R. Anti-müllerian hormone and insulin-like 3 levels in healthy normal-weight ovulatory and anovulatory eumenorrheic late adolescent females: potential early biomarkers of ovarian dysfunction? Eur J Obstet Gynecol Reprod Biol. 2015;195:188-92.
39. Pigny P, Merlen E, Robert Y, et al. Elevated serum level of antiMüllerian hormone in patients with polycystic ovary syndrome: relationship to the ovarian follicle excess and to the follicular arrest. J Clin Endocrinol Metab. 2003;88:5957-62.
40. Fanchin R, Schonäuer LM, Righini C, Guibourdenche J, Frydman R, Taieb J. Serum anti-Müllerian hormone is more strongly related to ovarian follicular status than serum inhibin B, estradiol, FSH and LH on day 3. Hum Reprod. 2003;18:323-7.
41. Zhao H, Zhou D, Liu C, Zhang L. The relationship between insulin resistance and obesity and serum antiMüllerian hormone level in Chinese women with polycystic ovary syndrome: a retrospective, single-center cohort study. Int J Womens Health. 2023;15:151-66.
42. Bernardi LA, Carnethon MR, de Chavez PJ, et al. Relationship between obesity and anti-Müllerian hormone in reproductive-aged African American women. Obesity (Silver Spring). 2017;25:229-35.
43. Büyükyılmaz G, Koca SB, Toksoy Adıgüzel K, Boyraz M, Gurbuz F. The role of the AMH, SHBG, free androgen index and LH/FSH ratio in the diagnosis of polycystic ovary syndrome in adolescent. Turkish J Pediatr Dis. 2024;18:34-40.
44. Shrikhande BA, Shrikhand LA. Association between weight loss in obese Indian adolescent girls with polycystic ovarian syndrome and decline in anti-Müllerian hormone concentrations. J South Asian Feder Obst Gynaecol. 2019;11:110-2.

DOI: 10.4274/forbes.galenos.2025.01033
Forbes J Med 2025;6(3):237-241

Original Article / Orijinal Araştırma

Artificial Intelligence Generated Questions in Medical Education: How Prompt Design in Different Chatbots Shapes Assessment in Obstetrics and Gynecology?

Tıp Eğitiminde Yapay Zeka Tarafından Oluşturulan Sorular: Farklı Chatbotlarda Prompt Tasarımı Kadın Hastalıkları ve Doğum Alanında Değerlendirmeyi Nasıl Şekillendiriyor?

Zuhal YAPICI COŞKUN[1], Yavuz Selim KIYAK[2], Özlem COŞKUN[2], Mehmet KOCA[3], İrem BUDAKOĞLU[2], Özhan ÖZDEMİR[4]

[1]Ankara Bilkent City Hospital, Clinic of Obstetrics and Gynecology, Ankara, Türkiye
[2]Gazi University Faculty of Medicine, Department of Medical Education and Informatics, Ankara, Türkiye
[3]Republic of Türkiye, Ministry of Health, Ankara, Türkiye
[4]University of Health Sciences Türkiye, Gülhane Faculty of Medicine, Department of Obstetrics and Gynecology, Ankara, Türkiye

Cite as: Yapıcı Coşkun Z, Kıyak YS, Coşkun Ö, Koca M, Budakoğlu İ, Özdemir Ö. Artificial intelligence generated questions in medical education: how prompt design in different chatbots shapes assessment in obstetrics and gynecology? Forbes J Med. 2025;6(3):237-241

ABSTRACT

Objective: The aim of this study is to assess the difficulty level of artificial intelligence (AI)-generated multiple-choice questions (MCQs) created by large language models (LLMs) using different prompts across various chatbots, compared to human-written questions.

Methods: We generated case-based MCQs on obstetrics and gynecology using two distinct prompts across four LLM-based chatbots. Expert-reviewed MCQs were administered to 97 medical students who were undergoing clerkship training in obstetrics and gynecology. Subsequently, item difficulty indices were calculated for each MCQ.

Results: The mean difficulty index of the AI-generated questions was 0.30. One prompt produced questions with a difficulty index of 0.34 (classified as difficult), while the other produced a lower difficulty index of 0.25 (classified as more difficult). In contrast, the mean difficulty index of the human-written questions was 0.63, indicating a moderate level of difficulty.

Conclusion: Our study highlights the challenges of using AI-generated MCQs in medical education. Although AI offers promising benefits for question generation, the questions produced were generally too difficult for undergraduate medical students. This underscores the need for more detailed and contextually informed prompt designs to better align AI outputs with assessment requirements. Although BDM-based chatbots enhance efficiency in question generation, expert review remains essential to ensure the appropriateness and quality of the items.

Keywords: Artificial intelligence, large language models, medical education, obstetrics and gynecology

ÖZ

Amaç: Bu çalışmada çeşitli chatbotlarda (yapay zeka robotu) farklı promptlar (istemler) kullanılarak büyük dil modelleri (BDM) ile yapay zeka tarafından üretilen çoktan seçmeli soruların (ÇSS) zorluk seviyesinin, insan tarafından yazılmış sorularla karşılaştırmalı bir şekilde değerlendirilmesi amaçlanmıştır.

Yöntem: Dört BDM tabanlı chatbotta (yapay zeka robotu) iki farklı istem kullanarak obstetrik ve jinekoloji üzerine vaka tabanlı ÇSS'lar oluşturulmuştur. Uzman grubu tarafından incelendikten sonra, ÇSS'lar kadın hastalıkları ve doğum anabilim dalında staj yapan 97 tıp öğrencisine uygulanmıştır. Daha sonra her bir ÇSS için madde (soru) güçlük indeksleri hesaplanmıştır.

Received/Geliş: 15.10.2025
Accepted/Kabul: 18.11.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Zuhal YAPICI COŞKUN, MD,
Ankara Bilkent City Hospital, Clinic of Obstetrics and Gynecology, Ankara, Türkiye
✉ zuhalyapici@yahoo.com
ORCID: 0000-0002-9338-4669

Bulgular: Yapay zeka tarafından üretilen soruların ortalama zorluk endeksi 0,30'dur. İstemlerden biri 0,34 zorluk indeksine sahip (zor olarak sınıflandırılan) sorular üretirken, diğeri 0,25'lik daha düşük bir zorluk indeksi (çok zor olarak kabul edilen) ile sonuçlanmıştır. Buna karşılık, insan tarafından yazılan soruların ortalama zorluk endeksi 0,63'tür ve bu da orta düzeyde bir zorluğa işaret etmektedir.

Sonuç: Çalışmamız, tıp eğitiminde yapay zeka tarafından üretilen ÇSS'lar ile insan üretimi olan ÇSS'lar yerine kullanımında karşılaşılabilecek zorlukları vurgulamaktadır. Yapay zeka, soru üretimi açısından umut verici görünmekle birlikte, üretilen soruların genellikle tıp öğrencileri için yüksek zorlukta olduğu gözlemlenmiştir. Bu sonuçlar ölçme değerlendirme gereksinimlerini karşılayabilecek yapay zeka çıktılarına ulaşabilmek için daha detaylandırılmış ve bağlamla uyumlu istemlerin yapılması gereksinimini vurgulamaktadır. Ayrıca, BDM tabanlı chatbotlar verimlilik açısından destek sağlarken, soruların uygunluğunu ve kalitesini sağlamak için uzman incelemesi önemini korumaktadır.

Anahtar Kelimeler: Yapay zeka, büyük dil modelleri, tıp eğitimi, kadın hastalıkları ve doğum

INTRODUCTION

Case-based multiple-choice questions (MCQs) are widely used in medical education,[1] particularly in fields such as obstetrics and gynecology,[2,3] because of their efficiency in assessing knowledge and cognitive skills. However, creating high-quality MCQs is a labor-intensive and time-consuming process that demands significant expertise. To address this challenge, leveraging artificial intelligence (AI) for the development of MCQs has been proposed as a promising solution, given its potential to play a pivotal role in medical education.[4-6]

Large language models (LLMs) have demonstrated significant potential to automate the generation of MCQs for medical education.[7-12] More specifically, recent literature provides evidence supporting the validity of ChatGPT-generated MCQs.[13] However, the findings suggest that employing simple prompts, such as "write four MCQs about this topic," often leads to suboptimal outputs.[14,15] Therefore, well designed[11] and detailed[16] prompts should be used to generate higher-quality MCQs. However, it remains unclear whether these commands can generate MCQs at targeted difficulty levels—defined as the proportion of test-takers who answer an item correctly—across different LLM-based chatbots. This study aims to fill this gap by examining the difficulty levels of AI-generated MCQs in obstetrics and gynecology using various prompts across four chatbots.

This study examines how prompt design and model selection influence the difficulty of AI-generated MCQs in obstetrics and gynecology and addresses how these compare with human-written questions. Our goal was to enhance understanding of how AI can be effectively utilized for case-based MCQ generation and to provide evidence supporting the development of better MCQs that are appropriately challenging for medical students.

METHODS

Study Setting, Design and Participants

This study was conducted in the Department of Obstetrics and Gynecology at University of Health Sciences Türkiye, Gülhane Faculty of Medicine, during the 2023–2024 academic year. Undergraduate medical education at the school spans six years.

The first three years focus on basic medical sciences, offering limited clinical exposure. In the fourth year, the gynecology and obstetrics clerkship is conducted in six rotations with 35–50 students per group, and the passing grade is calculated based on an assessment consisting of a multiple-choice examination. To avoid bias, only fourth-year students who had completed the obstetrics and gynecology clerkship were included. Using convenience sampling, we recruited 97 (88.1%) volunteer participants for the study. The participants completed a test comprising 18 obstetrics and gynecology MCQs, including 14 AI-generated questions and four human-written ones. Since the participants were native Turkish speakers, the AI-generated questions in English were translated into Turkish. Following completion of the internship, the test was conducted in a supervised classroom setting after informed consent had been obtained from all participants.

Multiple-Choice Question Generation

In March 2024, we aimed to generate four MCQs for each of the four LLM-based chatbots. The chatbots used were ChatGPT-3.5, ChatGPT-4, Gemini 1.0 Pro, and Mixtral-8x7B-Instruct-v0.1. We used two prompts for each chatbot: Kıyak's[16] promptand Zuckerman et al.'s[11] prompt. Previous studies have demonstrated their effectiveness.[9,11] These prompts were also used to develop a custom GPT to generate case-based MCQs for medical education because of their proven effectiveness.[17]

To generate the questions, we selected two topics from the learning objectives of the obstetrics and gynecology clerkship: the differential diagnosis of vaginal bleeding and the management of postpartum hemorrhage. These topics were provided to the four chatbots using two distinct prompts. Ultimately, we generated 14 questions—seven each on vaginal bleeding and postpartum hemorrhage—because Gemini 1.0 Pro declined to create a question using Zuckerman et al.'s[11] prompt, likely due to its requirement for an NBME-style question. We used ChatGPT's webpage for ChatGPT-3.5, Case-based MCQ Generator[18] for ChatGPT-4, Gemini's webpage for Gemini 1.0 Pro, and Hugging Face's webpage[19] for Mixtral. We added "differential diagnosis of vaginal bleeding" and "management of postpartum hemorrhage" to the prompts and generated each question

on a separate conversation page. Additionally, since Kıyak's[16] prompt requires selecting a difficulty level, we specifically requested the generation of difficult questions.

In addition to the 14 AI-generated questions, four expert-written MCQs—two on vaginal bleeding and two on postpartum hemorrhage—were provided to the participants.

The test was administered independently of the obligatory exams in the medical school and did not affect the students' grades. Informed consent was obtained from each participant. The study was approved by the University of Health Sciences Türkiye, Gülhane Scientific Research Ethics Committee (approval no: 2024/04, date: 24.04.2014).

Expert Panel

The AI-generated questions were reviewed and revised by two obstetrician–gynecologists with expertise with LLMs to ensure clinical accuracy, consistency, and suitability for the intended student level. Revisions were primarily limited to minor refinements in wording, clinical accuracy, and internal consistency, leaving the fundamental structure of the questions unchanged. None of the items necessitated major revisions.

Statistical Analysis

In line with our research question, we calculated the item (question) difficulty index for each question as the number of correct answers divided by the total number of test-takers. The difficulty index ranges from zero to one, with one representing the easiest level and zero representing the most difficult level. We classified values <0.30 as "too difficult", 0.30–0.40 as "difficult", 0.40–0.80 as "moderate", and >0.80 as "easy".[20,21]

RESULTS

Of the fourth-year medical students who completed their obstetrics and gynecology internship, 88.1% participated in an 18-item multiple-choice test. The mean difficulty index of the 14 AI-generated items was 0.30, placing the items at the threshold of the "difficult" category. In contrast, the four expert-authored items demonstrated a substantially higher mean difficulty index of 0.63, aligning with the "moderate" difficulty range (Table 1).

Notable differences were observed in the more detailed classification based on prompt type. Items generated using Kıyak's[16] prompt (n=8) were classified as "difficult", with an average difficulty index of 0.34. However, items derived from Zuckerman et al.'s[11] prompt (n=5) showed a lower average difficulty index of 0.25 and were classified as "very difficult."

Table 1. The difficulty levels of AI-generated and human-written multiple-choice questions

Focus	LLM type or human	Prompt	Difficulty index	Difficulty level
Vaginal bleeding	ChatGPT-3.5	Kıyak[16]	0.15	Too difficult
		Zuckerman et al.[11]	0.24	Too difficult
	ChatGPT-4	Kıyak[16]	0.68	Moderate
		Zuckerman et al.[11]	0.33	Difficult
	Gemini 1.0 Pro	Kıyak[16]	0.27	Too difficult
		Zuckerman et al.[11]	-	-
	Mixtral-8x7B	Kıyak[16]	0.40	Moderate
		Zuckerman et al.[11]	0.21	Too difficult
	Human-written	-	0.73	Moderate
	Human-written	-	0.60	Moderate
Postpartum hemorrhage	ChatGPT-3.5	Kıyak[16]	0.07	Too difficult
		Zuckerman et al.[11]	0.50	Moderate
	ChatGPT-4	Kıyak[16]	0.34	Difficult
		Zuckerman et al.[11]	0.09	Too difficult
	Mixtral-8x7B	Kıyak[16]	0.36	Difficult
		Zuckerman et al.[11]	0.13	Too difficult
	Gemini 1.0 Pro	Kıyak[16]	0.48	Moderate
		Zuckerman et al.[11]	-	
	Human-written	-	0.62	Moderate
	Human-written	-	0.57	Moderate

AI: Artifical intelligence, LLM: Large language model

In light of these findings, prompt design plays a critical role in shaping the cognitive accessibility of AI-generated questions.

In the two thematic areas evaluated—differential diagnosis of vaginal bleeding (n=7) and management of postpartum hemorrhage (n=7)—items generated by AI consistently exhibited lower difficulty indices compared to those written by experts. AI-generated items have a uniform level of difficulty, items written by experts provide a more balanced spectrum of difficulty. Specifically, none of the AI-generated items reached the "easy" classification (>0.80), while two of the four items written by experts fell within the "moderate-easy" range.

DISCUSSION

This study aimed to evaluate the difficulty levels of human-generated and AI-generated MCQs in obstetrics and gynecology. The findings highlighted the combined influence of prompt architecture and human judgment on the psychometric properties of MCQs in medical education. To the best of our knowledge, this is the first study to investigate the effect of different prompts and chatbots on the difficulty of MCQs.[13] Our key findings indicate that the mean difficulty index of the 14 AI-generated questions falls into the "too difficult" category when administered to a group of undergraduate medical students. Questions generated using Kıyak's[16] prompt were classified as difficult, whereas those generated using Zuckerman et al.'s[11] prompt were classified as too difficult. In comparison, the human-written questions reflected a moderate level of difficulty.

While our study found AI-generated questions difficult, previous researches reported moderate mean difficulty levels indices such as of 0.71,11 0.69,10 and 0.689 in ChatGPT-generated case-based questions. This discrepancy may be attributed to differences among the participant groups; variations in their background knowledge, experience, and familiarity with the subject matter could have influenced their performance. Our results emphasize that prompt design is not just a technical input for AI tools, but an educational intervention that can shape learning outcomes by modulating assessment difficulty. This positions prompt engineering as an emerging field of educational design.

The findings of this study highlight a critical issue in AI-based educational content creation: mismatched difficulty levels. Assessments that fail to accurately reflect the expected competency of learners may lead to suboptimal or even negative learning outcomes. When AI-generated items consistently fall outside the optimal difficulty range—particularly in formative assessments—this can compromise both learning efficiency and student motivation. Consequently, aligning the prompt structure with the curriculum level and learner profiles is not merely desirable, but a pedagogical necessity.

Our findings indicate that AI-generated questions, particularly those created using Zuckerman et al.'s[11] prompt, tend to be too difficult for undergraduates, likely because the prompt did not specify the intended difficulty. In contrast, Kıyak's[16] prompt produced questions with varying difficulty, ranging from too difficult to moderate, despite having explicitly requested difficult questions. This variability may arise from two factors: (1) limited detail provided in the prompts and (2) inherent limitations of LLMs.[22-24] The prompt templates can be improved by updating them to allow inclusion of additional context about local needs in the curriculum and target population, through prompt-engineering tactics.[25] However, the inherent limitations of LLMs will still require experts to review and revise the outputs to ensure that the MCQs are appropriate for the target population's needs. Better prompts can reduce the effort required after generation.

Study Limitations

Several limitations should be considered when interpreting the results of this study. First, the study was conducted with a specific group of fourth-year medical students at a single institution, which may limit the generalizability of the findings. Second, the study evaluated only a small number of questions generated by four AI models. Additionally, the study focused on two specific topics in obstetrics and gynecology, and the results might differ for other medical topics. Future research should expand the study to include multiple institutions and a more diverse group of participants to enhance generalizability, investigate the performance of various AI models and versions to identify those most effective in generating optimally difficult questions, and assess AI-generated question difficulty across a wider range of medical topics. Furthermore, the reviewers' expertise with LLMs could have affected their assessment of question clarity and perceived difficulty.

CONCLUSION

In conclusion, our study highlights both the challenges and potential of utilizing AI-generated MCQs in medical education, particularly in obstetrics and gynecology. Despite the promising advancements in AI, the questions generated were generally too difficult for undergraduate medical students. This underscores the necessity for more detailed and contextually informed prompt designs to better align AI outputs with assessment requirements. Additionally, while LLM-based chatbots provide valuable

support in efficient question generation, expert review remains crucial to ensure compliance with ethical guidelines.

Ethics

Ethics Committee Approval: The study was approved by the University of Health Sciences Türkiye, Gülhane Scientific Research Ethics Committee (approval no: 2024/04, date: 24.04.2014).

Informed Consent: Informed consent was obtained from each participant.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Pugh D, De Champlain A, Touchie C. Plus ça change, plus c'est pareil: making a continued case for the use of MCQs in medical education. Med Teach. 2019;41:569-77.
2. Balaha MH, El-Ibiary MT, El-Dorf AA, El-Shewaikh SL, Balaha HM. Construction and writing flaws of the multiple-choice questions in the published test banks of obstetrics and gynecology: adoption, caution, or mitigation? Avicenna J Med. 2022;12:138-47.
3. Jud SM, Cupisti S, Frobenius W, et al. Introducing multiple-choice questions to promote learning for medical students: effect on exam performance in obstetrics and gynecology. Arch Gynecol Obstet. 2020;302:1401-6.
4. Çalışkan SA, Demir K, Karaca O. Artificial intelligence in medical education curriculum: An e-Delphi study for competencies. PLoS One. 2022;17:e0271872.
5. Gordon M, Daniel M, Ajiboye A, et al. A scoping review of artificial intelligence in medical education: BEME Guide No. 84. Med Teach. 2024;46:446-70.
6. Stadler M, Horrer A, Fischer MR. Crafting medical MCQs with generative AI: a how-to guide on leveraging ChatGPT. GMS J Med Educ. 2024;41:Doc20.
7. Cheung BHH, Lau GKK, Wong GTC, et al. ChatGPT versus human in generating medical graduate exam multiple choice questions-A multinational prospective study (Hong Kong S.A.R., Singapore, Ireland, and the United Kingdom). PLoS One. 2023;18:e0290691.
8. Coşkun Ö, Kıyak YS, Budakoğlu Iİ. ChatGPT to generate clinical vignettes for teaching and multiple-choice questions for assessment: a randomized controlled experiment. Med Teach. 2025;47:268-74.
9. Kıyak YS, Coşkun Ö, Budakoğlu Iİ, Uluoğlu C. ChatGPT for generating multiple-choice questions: Evidence on the use of artificial intelligence in automatic item generation for a rational pharmacotherapy exam. Eur J Clin Pharmacol. 2024;80:729-35.
10. Laupichler MC, Rother JF, Grunwald Kadow IC, Ahmadi S, Raupach T. Large language models in medical education: comparing ChatGPT- to human-generated exam questions. Acad Med. 2024;99:508-12.
11. Zuckerman M, Flood R, Tan RJB et al. ChatGPT for assessment writing. Med Teach. 2023;45:1224-7.
12. Sathe TS, Roshal J, Naaseh A, L'Huillier JC, Navarro SM, Silvestri C. How I GPT it: development of custom artificial intelligence (AI) chatbots for surgical education. J Surg Educ. 2024;81:772-5.
13. Kıyak YS, Emekli E. ChatGPT prompts for generating multiple-choice questions in medical education and evidence on their validity: a literature review. Postgrad Med J. 2024;100:858-65.
14. Kıyak YS. ChatGPT's ability or prompt quality: what determines the success of generating multiple-choice questions. Acad Pathol. 2024;11:100119.
15. Ngo A, Gupta S, Perrine O, Reddy R, Ershadi S, Remick D. ChatGPT 3.5 fails to write appropriate multiple choice practice exam questions. Acad Pathol. 2024;11:100099.
16. Kıyak YS. A ChatGPT prompt for writing case-based multiple-choice questions. Rev Esp Educ Med. 2023;4:98-103.
17. Kıyak YS, Kononowicz AA. Case-based MCQ generator: a custom ChatGPT based on published prompts in the literature for automatic item generation. Med Teach. 2024;46:1018-20.
18. OpenAI. Case-based MCQ generator [Internet]. United States: ChatGPT; 2024 [cited 2024 Apr 30]. Available from: https://chatgpt.com/g/g-vuyyH0jUp-case-based-mcq-generator
19. Hugging Face. Hugging Face — Chat [Internet]. United States: Hugging Face; 2024 [cited 2024 Apr 30]. Available from: https://huggingface.co/chat/
20. Franzen D, Cuddy MM, Ilgen JS. Trusting your test results: building and revising multiple-choice examinations. J Grad Med Educ. 2018;10:337-8.
21. Tavakol M, O'Brien DG, Sharpe CC, Stewart C. Twelve tips to aid interpretation of post-assessment psychometric reports. Med Teach. 2024;46:188-95.
22. Masters K. Medical Teacher's first ChatGPT's referencing hallucinations: lessons for editors, reviewers, and teachers. Med Teach. 2023;45:673-5.
23. Kirshteyn G, Golan R, Chaet M. Performance of ChatGPT vs. HuggingChat on OB-GYN topics. Cureus. 2024;16:e56187.
24. Ozgor BY, Simavi MA. Accuracy and reproducibility of ChatGPT's free version answers about endometriosis. Int J Gynaecol Obstet. 2024;165:691-5.
25. Indran IR, Paranthaman P, Gupta N, Mustafa N. Twelve tips to leverage AI for efficient and effective medical question generation: a guide for educators using Chat GPT. Med Teach. 2024;46:1021-6.

DOI: 10.4274/forbes.galenos.2025.13540
Forbes J Med 2025;6(3):242-248

Orijinal Araştırma / Original Article

Türkiye'de Sağlık Çalışanlarında Beyin Göçü

Brain Drain Among Health Workers in Türkiye

Orhan ÇAKIR[1], Metehan AKKOÇ[2], Çağrı Işıl ÇAKIR[3], Ayşe Şimal GENÇER[4]

[1]İzmir Tınaztepe Üniversitesi, Sağlık Hizmetleri Meslek Yüksekokulu, İzmir, Türkiye
[2]Hacettepe Üniversitesi Eğitim Fakültesi, Ankara, Türkiye
[3]Ege Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, İzmir Türkiye
[4]İzmir Tınaztepe Üniversitesi Sağlık Bilimleri Fakültesi, İzmir, Türkiye

Atıf: Çakır O, Akkoç M, Çakır ÇI, Gençer AŞ. Brain drain among health workers in Türkiye. Forbes J Med. 2025;6(3):242-248

ÖZ

Amaç: Beyin göçü, bir ülkede yetişmiş ve eğitimli bireylerin daha iyi iş veya eğitim olanakları amacıyla yurt dışına göç etmesi sonucu ortaya çıkan nitelikli insan gücü kaybı olarak tanımlanmaktadır. Bu süreçten en çok etkilenen meslek gruplarının başında sağlık profesyonelleri gelmektedir. Bu araştırmanın amacı, Türkiye'de sağlık alanında beyin göçü konusuna odaklanan lisansüstü tezleri inceleyerek içerik analizi yapmak ve mevcut literatüre ilişkin genel bir görünüm sunarak gelecekte bu alanda çalışacak araştırmacılara yol göstermektir.

Yöntem: Araştırmada, Yükseköğretim Kurulu Tez Veri Tabanı'nda yer alan lisansüstü tezler sistematik olarak incelenmiştir. Tezler; düzeyi, yazım dili ve yapıldığı üniversite gibi çeşitli değişkenler doğrultusunda değerlendirilmiştir. Tam metnine ulaşılamayan tezler için ilgili araştırmacılarla iletişime geçilerek tam metin temin edilmiş ve analizler bu doğrultuda gerçekleştirilmiştir.

Bulgular: Toplam 163 tez incelenmiş olup bunların 21'inin sağlık alanında beyin göçünü ele aldığı belirlenmiştir. Tezlerin büyük çoğunluğunun son beş yıl içerisinde yazıldığı ve sağlık profesyonellerinin göç etme nedenleri ile bu göçün sağlık sistemine etkilerinin en sık araştırılan konular olduğu görülmüştür.

Sonuç: Sağlık alanında beyin göçü, özellikle hekim ve hemşire göçü çerçevesinde ele alınan önemli bir sorun olmaya devam etmektedir. Gelecekte yapılacak araştırmalarda, bu olgunun politika geliştirme ve önleme stratejileri bağlamında daha derinlemesine incelenmesi gerekmektedir. Bu kapsamda gerçekleştirilecek tez çalışmalarının, multidisipliner yaklaşımlar ve çok paydaşlı iş birlikleriyle desteklenerek beyin göçü olgusunun bütüncül bir şekilde ele alınması önerilmektedir.

Anahtar Kelimeler: Hekimler, hemşireler, sağlık çalışanları, göç, Türkiye

ABSTRACT

Objective: Brain drain refers to the loss of qualified human capital when educated and skilled individuals migrate abroad in pursuit of better employment or educational opportunities. Health professionals are among the occupational groups most affected by this phenomenon. This study aims to conduct a content analysis of postgraduate theses in Türkiye that focus on brain drain in the health sector, providing an overview of existing academic work and guidance for future researchers.

Methods: Postgraduate theses listed in the Council of Higher Education Thesis Database were systematically reviewed. The theses were evaluated based on variables such as degree level, language, and university. For theses with inaccessible full texts, the researchers were contacted directly to obtain the complete versions, and the analyses were conducted accordingly.

Results: A total of 163 theses related to brain drain were reviewed, 21 of which specifically addressed the health sector. The majority were written within the last five years and mainly examined the reasons behind health professionals' migration and its impacts on health systems.

Conclusion: Brain drain in the health sector, particularly physician and nurse migration, remains a critical issue. Future studies should explore this phenomenon more comprehensively within the context of policy development and preventive strategies. Multidisciplinary approaches and multi-stakeholder collaborations are recommended to address brain drain holistically.

Keywords: Physician, nurses, medical staff, emigration, Türkiye

Received/Geliş: 10.10.2025
Accepted/Kabul: 24.11.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Dr. Öğr. Üyesi Orhan ÇAKIR,
İzmir Tınaztepe Üniversitesi, Sağlık Hizmetleri Meslek Yüksekokulu, İzmir, Türkiye
orhan.cakir@tinaztepe.edu.tr
ORCID: 0000-0002-4231-3561

GİRİŞ

Beyin göçü, günümüzde uluslararası iş gücü hareketliliğinin en belirleyici ve stratejik bileşenlerinden biri olarak görülmekte; özellikle yükseköğrenim görmüş ya da mesleki uzmanlık kazanmış bireylerin daha iyi yaşam koşulları, mesleki gelişim olanakları ve güvenli çalışma ortamları arayışıyla başka ülkelere yönelmesiyle şekillenmektedir. Bu süreç, yalnızca bireysel bir tercih olmanın ötesine geçerek ülkelerin beşerî sermayesini, bilgi üretim kapasitesini ve kalkınma potansiyelini doğrudan etkilemektedir.[1-4] Nitelikli insan gücünün kaybı; ekonomik üretkenliğin azalması, inovasyon kapasitesinin zayıflaması ve toplumsal eşitsizliklerin derinleşmesi gibi çok boyutlu sonuçlar doğurmakta, özellikle gelişmekte olan ülkelerde telafisi güç bir katma değer erozyonuna (ülkenin değer üretme kapasitesinin zayıflaması) neden olmaktadır.[1-7]

Göçün nedenlerini açıklamada sık kullanılan kuramsal yaklaşımlar, beyin göçünün yapısal boyutlarına ışık tutmaktadır. Everett Lee'nin geliştirdiği itme-çekme modeli, bireylerin göç kararlarını kaynak ülkelerdeki olumsuz koşullar (düşük ücret, iş güvencesizliği, siyasi belirsizlik) ile hedef ülkelerdeki cazip fırsatlar (yüksek yaşam standartları, mesleki gelişim, sosyal güvence) üzerinden açıklamaktadır.[3-7] Diğer taraftan bağımlılık kuramı ve Wallerstein'ın dünya-sistem yaklaşımı, nitelikli iş gücünün çevre ülkelerden merkez ülkelere yönelmesini, küresel eşitsizliklerin yeniden üretildiği bir süreç olarak ele almaktadır.[6,8,9] Bu çerçevede sağlık çalışanlarının göçü, yalnızca bireysel bir mesleki yönelim değil; aynı zamanda merkez-çevre ülkeler arasındaki asimetrik güç ilişkilerinin bir yansımasıdır.

Son yıllarda literatürde önem kazanan bir diğer yaklaşım olan beyin dolaşımı modeli ise göçün tek yönlü bir kayıp olarak görülmemesi gerektiğini, diaspora aracılığıyla bilgi, deneyim ve teknoloji transferinin kaynak ülkeler için uzun vadeli fırsatlar yaratabileceğini vurgulamaktadır.[6-17] Bu nedenle beyin göçü; istihdam, ekonomik yapı, kamu politikaları, toplumsal eşitsizlikler ve insan hakları gibi farklı eksenlerde değerlendirilen çok boyutlu bir olgu niteliği taşımaktadır.

Sağlık çalışanları, beyin göçünün en yoğun yaşandığı meslek grupları arasında yer almaktadır. Uluslararası literatürde 1970'li yıllardan bu yana sağlık çalışanlarının göç eğilimlerinde belirgin bir artış olduğu bildirilmektedir.[2-6] Türkiye'de ise bu eğilim son yıllarda daha görünür hâle gelmiştir.[3,10,18-21] Türk Tabipleri Birliği verilerine göre, yurtdışında çalışma amacıyla "iyi hâl belgesi" başvurusu yapan hekim sayısı 2012'de 59 iken, 2023 yılında 3025'e ulaşmıştır.[3] Benzer şekilde hemşireler ve sağlık eğitimi alan öğrenciler arasında da göç eğiliminin hızla arttığı görülmektedir. Türk Hemşireler Derneği'nin 2023 verilerine göre hemşirelerin %76,3'ü yurtdışında çalışmak istemekte, ekonomik sıkıntılar, mesleki sorunlar ve şiddet en güçlü itici faktörler olarak öne çıkmaktadır.[10] Üniversite öğrencilerinde yapılan araştırmalar ise, sağlık alanında öğrenim gören gençlerin büyük çoğunluğunun mezuniyet sonrası yurtdışında çalışma eğiliminde olduğunu göstermektedir.[3,10,18,19] Ayrıca göç ağları kuramı, daha önce göç eden sağlık çalışanlarının oluşturduğu sosyal bağların, yeni göçmenlere bilgi ve destek sağlayarak göçün maliyetini düşürdüğünü ve göç kararlarını kolaylaştırdığını ortaya koymaktadır.[6,20,21]

Sağlık çalışanlarının göçü, halk sağlığı ve sağlık sistemi açısından kritik öneme sahiptir. Nitelikli insan gücündeki azalma, hizmet sunumunda aksamalara, sağlık çalışanı başına düşen iş yükünün artmasına, tükenmişlik düzeyinin yükselmesine ve sağlık hizmetlerine erişimde eşitsizliklere yol açabilmektedir. Öte yandan bazı ülkelerde sağlık çalışanlarının uluslararası hareketliliği, insan kaynağı fazlasının dengelenmesi veya uzmanlık açığının kapatılması amacıyla olumlu bir süreç olarak değerlendirilebilmektedir.[12-15]

Bu bağlamda, Türkiye'de sağlık alanında beyin göçüne ilişkin akademik bilgi birikiminin sistematik biçimde değerlendirilmesi önem kazanmaktadır. Ancak mevcut literatür incelendiğinde, bu konuda yapılmış lisansüstü tezlerin kapsamı, tematik dağılımı ve metodolojik özelliklerine yönelik bütüncül bir değerlendirme bulunmadığı görülmektedir. Bu araştırma ile Türkiye'de sağlık alanında beyin göçünü ele alan lisansüstü tezler incelenmiş, alandaki eğilimler ortaya koyulmuş, mevcut boşluklar ifade edilmiş ve gelecekte yapılacak araştırmalara öneriler geliştirilmiştir. Böylece mevcut durumun kapsamlı bir şekilde ortaya konulması ve sağlık alanındaki beyin göçüne ilişkin akademik bilgi birikiminin güçlendirilmesine katkı sağlanması amaçlanmıştır.

YÖNTEM

Araştırmanın amacı ve türü: Araştırma türü bakımından, nitel bir araştırma yaklaşımı benimseyerek, nitel içerik analizi yöntemiyle gerçekleştirilmiştir. Nitel içerik analizi, belirli metinlerin sistematik biçimde incelenmesini amaçlayan bir araştırma yöntemidir. İçerik analizi değişen araştırma amaçlarına göre nitel, nicel ve bazen de karma tarzda yapılabilmektedir.[22] Araştırmadan elde edilen veriler ise nicel desende incelenmiştir.

Araştırma soruları: Araştırma şu sorulara yanıt aramaktadır:

- Sağlık temalı beyin göçü tezlerinin yıllara göre dağılımı nasıldır?
- Araştırmalarda hangi yöntemsel yaklaşımlar tercih edilmiştir?

- Çalışmalarda hangi sağlık meslek grupları ön plana çıkmaktadır?
- Tezlerde beyin göçüne ilişkin hangi itici ve çekici faktörler vurgulanmaktadır?

Araştırmanın evren ve örneklemi: Bu araştırmada, ülkemiz Yükseköğretim Kurulu (YÖK) Ulusal Tez Merkezi'nde yer alan yüksek lisans, doktora ve tıpta uzmanlık tezleri; yazım yılları, türleri, araştırma desenleri ve enstitü alanlarına göre incelenmiştir. Araştırmanın örneklemi, evrende yer alan toplam 163 tezden oluşmaktadır. Çalışmada yalnızca lisansüstü tezlere odaklanılmasının temel nedeni, bu tezlerin içeriklerine dair daha derinlemesine bilgi edinmeyi sağlamak ve aynı zamanda bu içerikleri bilimsel tartışmaya açarak alandaki durumu değerlendirebilmektir.

Verilerin toplanması: Türkiye YÖK Ulusal Tez Merkezi'nde araştırılan konuyla ilişkili belirlenen ölçütler ile 01.08.2025-01.09.2025 tarihleri aralığında "beyin göçü" anahtar kelime grubuyla "Aranacak Alan, İzin Durumu ve Tez Türü" ölçütlerinde "Tümü" seçilerek ve tarama terimi alanına "beyin göçü" kelimeleri yazılarak tarama yapılmış olup; sonucunda 163 araştırmaya ulaşılmıştır. İncelenen tezlerden 21 tanesi sağlık alanında çalışılmış olan tezler olduğu tespit edilmiştir. Araştırma örneklemini oluşturan 21 tezin 2020-2025 yılları arasında yapıldığı tespit edilmiştir. Veri toplama sürecinde, tezlerin giriş, yöntem, bulgular ve tartışma bölümleri ayrıntılı olarak incelenmiştir.

Verilerin değerlendirilmesi: Tezlerden elde edilen veriler tanımlayıcı istatistiklerle değerlendirilmiştir. Verilerin analizinde SPSS 25.0 ve Microsoft Excel kullanılmıştır.

BULGULAR

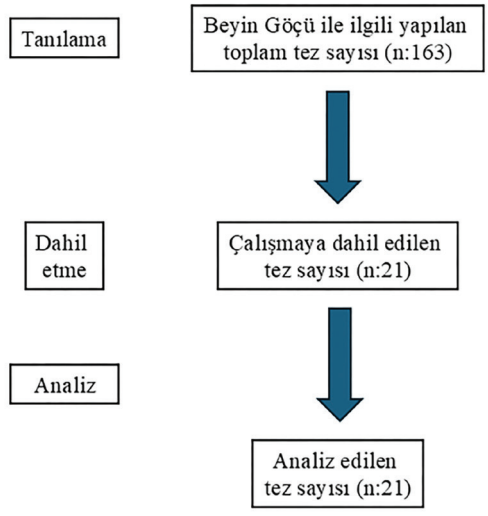
Bu araştırma, YÖK Ulusal Tez Merkezi veri tabanında yer alan tezlerin incelenmesine dayanan tanımlayıcı bir çalışma niteliğindedir. Yapılan taramalar sonucunda toplamda 163 teze ulaşılmıştır.

Beyin göçü ile ilgili yapılan tez taraması sonucunda yapılan 163 tezden, 21 tanesinin sağlık alanında çalışıldığı tespit edilmiştir. Diğer araştırmalar dışlanmıştır ve 21 tez analiz edilmiştir (Şekil 1).

Tezlerin yıllara göre dağılımı incelendiğinde, sağlık alanında beyin göçü konulu çalışmaların 2020 yılından itibaren çalışılmaya başlandığı ve en yoğun şekilde 2024 yılında gerçekleştirildiği görülmektedir (Şekil 2). Ayrıca ilgili konunun trendinin her geçen yıl arttığı da görülmektedir.

Araştırmaya uygun olarak YÖK Ulusal Tez Merkezi veri tabanında yer alan 21 tez, Tablo 1'de yer alan çeşitli değişkenler (tez türü, üniversite adı vb.) doğrultusunda analiz edilmiştir. Tam metnine erişilemeyen tezlerin araştırmacıları iletişim kurulmuştur. Araştırmacılar ile kurulan iletişim sonrasında tezlere ulaşılmıştır ve değerlendirmeye alınmıştır. İncelenen tezlerin düzeylerine göre dağılımlarında %76,2 oranla yüksek lisans düzeyinde çalışıldığı tespit edilmiştir. Konu ile ilgili %52,4 oranında "Doç. Dr." ünvanlı akademisyenlerin danışmanlık yaptığı saptanmıştır. Araştırmada incelenen tezlerin yöntemleri nitel, nicel ve karma olarak gerçekleştirilmiştir. İncelenen tezlerin %71,4'ü nicel, %19,1'i nitel ve %9,5'i karma yöntemle gerçekleşmiştir. Araştırmada incelenen tezlerde veri toplama aracı olarak %71,4 oranında en çok anketlerin kullanıldığı tespit edilmiştir (Tablo 1).

Araştırmaya uygun olarak YÖK Ulusal Tez Merkezi veri tabanında yer alan 21 tez, Tablo 2'de tezlerde vurgulanan itici ve çekici faktörlere göre incelenmiştir. İtici ve çekici faktörler 4 başlık altında sunulmuştur (Tablo 2). İtici faktörler; ekonomik ve mali sorunlar, çalışma koşulları, mesleki ve kurumsal faktörler, sosyo-politik faktörlerdir. Çekici faktörler ise; ekonomik fırsatlar, çalışma ortamı ve koşulları, eğitim, kariyer ve mesleki gelişim, sosyal ve kültürel faktörler olarak kategorize edilmiştir (Tablo 2).



Şekil 1. Araştırma akış şeması

Şekil 2. Tezlerin yıllara göre dağılımı

Etik

Bu araştırmada birincil veri toplama yöntemleri kullanılmamıştır. Veriler kamuoyuna açık YÖK Ulusal Tez Merkezi veri tabanından elde edilmiştir. Çalışmada insan veya hayvan verisi kullanılmamıştır. Bu nedenle çalışma için etik kurul izni alınmasına ihtiyaç duyulmamıştır.

TARTIŞMA

Beyin göçü, özellikle nitelikli iş gücünün, yüksek eğitimli bireylerin, sağlık çalışanlarının ve akademisyenlerin daha iyi yaşam, kariyer ve çalışma koşulları arayışıyla yurt dışına yönelmesi olgusu üzerinden şekillenen çok boyutlu bir toplumsal sorundur. Sağlık alanında faaliyet gösteren bireylerin göç eğilimleri, sadece bireysel bir tercih olmanın ötesinde, sağlık sisteminin işleyişini, insan kaynağı planlamasını ve toplumsal eşitsizlikleri doğrudan etkileyen yapısal bir mesele haline gelmiştir. Bu bağlamda araştırmanın temel amacı, sağlık alanıyla ilişkili lisansüstü tezlerde "beyin göçü" olgusunun nasıl ele alındığını, ne tür temalar ve söylemler çerçevesinde temsil edildiğini ve bu bağlamda literatüre nasıl yansıtıldığını ortaya koymaktır.

Tablo 1. Tezlerin genel özellikleri

Değişkenler	n	%
Düzeyi		
Yüksek lisans	16	76,2
Doktora	2	9,5
Tıpta uzmanlık	3	14,3
Danışman ünvanı		
Dr. Öğr. Üyesi/Yard. Doç. Dr.	3	14,3
Doç. Dr.	11	52,4
Prof. Dr.	7	33,3
Yapıldığı üniversitesi		
Sağlık Bilimleri Üniversitesi	3	14,3
Sivas Cumhuriyet Üniversitesi	3	14,3
Hacettepe Üniversitesi	2	9,5
Diğer üniversiteler	13	61,9
Yapıldığı enstitü/fakülte		
Lisansüstü Eğitim Enstitüsü	5	23,7
Sağlık Bilimleri Enstitüsü	4	19,0
Sosyal Bilimler Enstitüsü	6	28,6
Tıp Fakültesi	3	14,3
Fen Bilimleri Enstitüsü	1	4,8
Nüfus Etütleri Enstitüsü	1	4,8
Göç Enstitüsü	1	4,8
Yapıldığı ABD/bölüm		
Aile Hekimliği Ana Bilim Dalı	3	14,3
Hemşirelik/Halk Sağlığı Ana Bilim Dalı	3	14,3
İşletme Ana Bilim Dalı	4	19,0
Sağlık Yönetimi Ana Bilim Dalı	2	9,5
Sağlık Kuruluşları Yöneticiliği Ana Bilim Dalı	2	9,5
Sosyoloji Ana Bilim Dalı	1	4,8
Diğer	6	28,6
Çalışma grubu		
Hekim	5	23,8
Hemşire	5	23,8
Karma sağlık çalışanı	5	23,8
Öğrenci-sağlık alanında	6	28,6
Araştırma yöntemi		
Nicel	15	71,4
Nitel	4	19,1
Karma	2	9,5
Veri toplama tekniği		
Görüşme	4	19,1
Anket	15	71,4
Görüşme/anket	2	9,5
Toplam	67	100.0

Araştırma bulguları sonucunda toplamda 163 teze ulaşılmıştır. Ulaşılan bu tezlerden 21 tanesi doğrudan sağlık alanı ile ilgili tezlerdir. Bu araştırmaların da son yıllarda yapıldığı görülmektedir. Bu durumda pandemi sonrası süreç, çalışma koşullarındaki memnuniyetsizlikler, küresel talep artış, genç kuşakların değer yönelimi ve politika yapıcıların gündemine girmesi ile açıklanabilir.

Araştırma sırasında yapılan taramalar sonucunda, çalışma grubunun sağlık alanı olduğu ilk tez çalışmasının 2020 yılında Reyhan Kurtbasan tarafından yapılan "Hemşirelik, Okul Öncesi Öğretmenliği ve Psikolojik Danışmanlık ve Rehberlik Bölümlerindeki Öğrencilerin Beyin Göçüne Yönelik Tutumları" başlıklı yüksek lisans tezi olduğu tespit edilmiştir. Tezler yıllara göre değerlendirildiğinde, özellikle 2020 yılı sonrasında tezlerin sayısında önemli bir artışın olduğu tespit edilmiştir. Bu durum, özellikle 2019 koronavirüs hastalığı (COVID-19) pandemisi sürecinde sağlık çalışanlarının toplum tarafından öneminin daha yüksek düzeyde kabul görmesi ve sağlık çalışanlarının koşullarının kamuoyunda daha fazla tartışılmasıyla paralellik göstermektedir. Nitekim pek çok ülkede COVID-19 pandemisi, sağlık çalışanlarının çalışma koşulları üzerindeki baskıyı artırmış, sağlık çalışanı yetersizliği küresel ölçekte görünür hâle gelmiş ve nitelikli iş gücü göçünü hızlandırmıştır.[3,10] Ayrıca, genç kuşakların kariyer yönelimi, küresel hareketlilik isteği ve yurt dışındaki yaşam standartlarına ilişkin beklentilerindeki değişim de bu artışı destekleyen bir diğer faktördür. Literatürde genç sağlık profesyonelleri arasında göç etme isteğinin giderek güçlendiği, özellikle hekimler, hemşireler ve sağlık eğitimi alan öğrenciler arasında göç eğiliminin çok yüksek olduğu bildirilmektedir.[3,18,19]

Hazırlanan tezlerin düzeylerine göre dağılımları incelendiğinde; %76,2 oranla yüksek lisans düzeyinde yapılmış olan tezler daha fazladır. İkinci sırayı %14,3 oranla tıpta uzmanlık tezleri takip ederken, son sırada %9,5 oranla doktora tezlerinin yapıldığı tespit edilmiştir. Yıllara göre tezlerin sayısında artış olduğu gibi, tüm tez düzeylerinde de bir artış söz konusudur. Bu durumu da yine Türkiye'deki sağlık alanında beyin göçü araştırmalarının yıllar içinde gördüğü ilgi ve bunun akademik çalışmalar içinde tezlere yansıması şeklinde açıklanabilir.

Tablo 2. Vurgulanan itici ve çekici faktörler (n=21)	
İtici faktörler	**Çekici faktörler**
Ekonomik ve mali sorunlar • Düşük maaşlar • Geçim zorluğu, ekonomik kriz • Yüksek vergi oranları • Teşvik ve ek ödeme sistemine dair memnuniyetsizlik	**Ekonomik fırsatlar** • Yüksek maaş ve daha iyi gelir • Daha yüksek yaşam standardı • Daha kolay geçim imkânı
Çalışma koşulları • Ağır iş yükü, uzun nöbetler, yoğun tempo • Yetersiz muayene süreleri • Güvencesizleşme ve belirsizlik • Zorunlu hizmet uygulaması • Pandemi süreci ile artan baskılar	**Çalışma ortamı ve koşulları** • Güvenli ve düzenli çalışma ortamı • Daha kısa çalışma saatleri • Sağlıkta şiddetin daha az olması • Malpraktis yasalarının daha esnek olması
Mesleki ve kurumsal faktörler • Sağlıkta şiddet (sözel ve fiziksel) • Mobbing ve destek eksikliği • Mesleki saygınlığın azalması, prestij kaybı • Kariyer belirsizliği • Hukuki/idarî koruma ve düzenlemelerin yetersizliği	**Eğitim, kariyer ve mesleki gelişim** • Araştırma ve akademik fırsatlar • Teknolojik imkanların daha gelişmiş olması • Uzmanlaşmak istenilen branşların bulunması • Bilime verilen değerin daha yüksek olması • Mesleki tatmin ve saygınlığın artması
Sosyo-politik faktörler • Politik baskı ve düşünce özgürlüğü eksikliği • Siyasi istikrarsızlık, güvenlik kaygısı • Sağlık politikaları ve hükümetin söylemleri • Meslek örgütlerinin ve sendikaların zayıflığı	**Sosyal ve kültürel faktörler** • Çocukların eğitim olanakları • Yeni kültürler tanıma isteği, yabancı dil öğrenme • Uluslararası deneyim ve profesyonel saygınlık • Sosyal ve yaşam koşullarının daha iyi olması

Sağlık alanındaki çalışanlar düzeyinde incelenen 21 tezde, hekimlere, hemşirelere ve sağlık alanındaki öğrencilere odaklanıldığı görülmektedir. Ayrıca bu araştırmaların tamamının 2020 yılı ve sonrasında gerçekleştirildiği tespit edilmiştir. Bu durum, pandemi sonrası dönemde sağlık çalışanlarının önemine yönelik artan vurgu ve konunun geniş kitlelerce tartışılır hâle gelmesiyle ilişkilendirilebilir.

Araştırmada analiz edilen tezlerin incelenmesinde tezlerin %71,4'ünün nicel, %19,1'inin nitel ve %9,5'inin karma desende yapıldığı tespit edilmiştir. Araştırma bulgularının değerlendirilmesinde sıklıkla başvurulan kanıt düzeyi piramidi, elde edilen sonuçların güvenilirliğini ve geçerliliğini belirlemede önemli bir araçtır. Bu araştırmanın önemli bir bulgusu, Türkiye'de sağlık alanındaki beyin göçü araştırmalarının büyük ölçüde tanımlayıcı çalışmalar üzerine yoğunlaşmış olmasıdır. Kanıt düzeyi piramidi açısından değerlendirildiğinde, bu araştırmalar piramidin alt basamaklarında yer almakta; yüksek düzeyde kanıt üreten girişimsel, kohort veya karma yöntemli çalışmaların ise oldukça sınırlı olduğu görülmektedir. Bu durum, sağlık çalışanlarının göç eğilimlerinin neden-sonuç ilişkileri bakımından daha derinlemesine incelenmesini güçleştirmektedir. Halk sağlığı perspektifinden bakıldığında, beyin göçü yalnızca bireysel bir kariyer hareketliliği değil; sağlık sisteminin sürdürülebilirliğini tehdit eden kritik bir yapısal sorundur. Sağlık çalışanı yetersizliği, iş yükünün artmasına, tükenmişlik oranlarının yükselmesine, teşhis-tedavi süreçlerinde hata riskinin artmasına ve sağlık hizmetlerine erişimde eşitsizliklere yol açabilmektedir. Bu bağlamda, yüksek düzeyde kanıt üretecek araştırma tasarımlarına ihtiyaç olduğu açıktır.[23]

Vurgulanan İtici ve Çekici Faktörler

Araştırmada ele alınan itici ve çekici faktörlerin analizinde, itici faktörlerin ekonomik ve mali güçlükler, çalışma koşullarına ilişkin olumsuzluklar, mesleki ve kurumsal yapıya bağlı etmenler ile sosyo-politik dinamiklerden oluştuğu saptanmıştır. Çekici faktörler ise ekonomik fırsatlar, çalışma ortamı ve koşulları, eğitim, kariyer ve mesleki gelişim ve sosyal ve kültürel faktörler olarak tespit edilmiştir. Bu durum sağlık alanında beyin göçü ile ilgili yapılan araştırmalarda da vurgulanmıştır.[24,25] Sağlık sektöründe beyin göçünü tetikleyen itici faktörler çoğunlukla kaynak ülkedeki olumsuz koşullardır. Örneğin, Karatuzla tarafından yapılan araştırmada düşük maaş, ağır iş yükü, yetersiz kariyer ve eğitim fırsatları ile sosyal/güvenlik sorunlarının beyin göçünü artıran ana nedenler olduğu belirtilmiştir.[24] Benzer şekilde Köse Tosunöz tarafından yapılan araştırmada sağlık çalışanlarında ekonomik güçlükler, kötü çalışma koşulları ve sağlıkta şiddet gibi etkenlerin sağlık çalışanlarını göçe zorladığı vurgulanmıştır.[25] Bu itici koşullar, bireyleri memnun olmadığı ortamdan ayrılmaya yöneltmektedir. Öte yandan çekici faktörler, nitelikli sağlık çalışanlarını göç edilen ülkelere çeken olumlu unsurlardır. Literatürde bu bağlamda genellikle daha yüksek ücretler, yaşam standardı, ileri eğitim/kariyer fırsatları ve güvenli çalışma koşulları öne çıkmaktadır.[24,25] Literatüre göre, gelişmiş ülkelerde sunulan istikrarlı yaşam kalitesi, güvenli ortam ve mesleki gelişim

olanakları sağlık çalışanlarını cezbetmekte; artırılmış kariyer imkânları, ekonomik olanaklar ve daha konforlu yaşam koşulları ise göç eğilimini güçlendirmektedir. Bu itici ve çekici faktörlerin etkileşimi, nitelikli sağlık iş gücünün yurt dışına yönelmesinde temel belirleyicidir. Literatürdeki araştırmalar sağlık alanındaki beyin göçünde itici faktör olarak düşük ücretleri, olumsuz çalışma ve yaşam koşullarını, politik/sosyal baskıları; çekici faktör olarak ise yüksek kazanç ve yaşam standardı, iyi eğitim-kariyer fırsatları ile güvenli çalışma ortamlarını vurgulamıştır. Araştırma bulguları, alandaki diğer çalışmalarla paralellik göstermektedir.

Sağlık Hizmetlerinde Beyin Göçünün Etkileri

Araştırmada grubunu oluşturan tezlerin analizleri doğrultusunda, Türkiye'de sağlık alanındaki istihdam ve çalışma koşullarının yetersizliği ile yurtdışındaki yüksek yaşam standartları ve mesleki gelişim olanakları, sağlık çalışanları arasında yurtdışında çalışma ve yaşama isteğini güçlü bir eğilim hâline getirmektedir. Beyin göçüne yönelik bu eğilimlerin yalnızca bireysel tercihlerle sınırlı kalmadığı; aynı zamanda sosyo-ekonomik yapının sunduğu kısıtlılıklar, toplumsal değerler ve ulusal düzeydeki politika eksiklikleri gibi makro faktörlerle de yakından ilişkili olduğu görülmektedir. Araştırma bulguları, Türkiye'de sağlık insan gücünün kaybını önlemek için çalışma koşullarının iyileştirilmesi, nitelikli iş fırsatlarının oluşturulması gibi gereklilikleri ortaya koymaktadır. Ayrıca bu gerekliliklerin, acil bir gereklilik olduğu farklı araştırmacılar tarafından da vurgulanmaktadır.[26] Bu kapsamda geliştirilecek stratejilerin yalnızca ekonomik iyileştirmelerle sınırlı kalmaması; aynı zamanda sosyal ve politik reformlarla da desteklenmesi gerektiği anlaşılmaktadır. Özellikle genç sağlık profesyonellerinin mesleki tatmin ve güvencelerini artıracak politikaların hayata geçirilmesi, Türkiye'nin sağlık sektöründe beyin göçüyle mücadelede kritik öneme sahiptir. Bu araştırma, sağlık alanındaki nitelikli insan gücü kaybının önlenmesi için politika yapıcılara ve ilgili kurumlara kapsamlı ve çok boyutlu stratejiler geliştirme sorumluluğunu bir kez daha hatırlatmaktadır. Beyin göçü eğilimlerinin yalnızca ekonomik nedenlerden değil, aynı zamanda Türkiye'nin genel sosyo-ekonomik yapısındaki yapısal sorunlardan beslendiği görülmektedir. Bu nedenle, bireylerin yurtdışına yönelimlerini azaltmak adına yapılacak her türlü iyileştirmenin, sadece bireysel düzeyde değil, aynı zamanda yapısal ve sistem düzeyinde reformlarla desteklenmesi zorunludur. Bu çerçevede hem bireylerin hem de ülkemizin sağlık insan gücünün sürdürülebilirliği açısından gerekli adımların ivedilikle atılması büyük bir gereklilik olarak öne çıkmaktadır.

Sağlık sistemlerinde insan gücü yetersizliğinin halk sağlığına etkileri de göz ardı edilmemelidir. Sağlık iş gücündeki yetersizlik, özellikle halk sağlığı alanında sistemin işleyişini olumsuz etkilemekte ve geride kalan personelin iş yükünü ciddi oranda artırmaktadır. Bu durum yalnızca hizmet sunumunda aksamalara yol açmakla kalmamakta, aynı zamanda tanı ve tedavi süreçlerinde hata oranlarının yükselmesine de neden olmaktadır. Nitekim Harvard Üniversitesi'nde yürütülen bir araştırmada, 1.000 kişi başına düşen hekim sayısının yalnızca bir kişi artırılmasının, bebek ölümlerini azaltmada çok önemli bir etkiye sahip olduğu ve kısa vadede %15, uzun vadede ise %45 oranında düşüş sağladığı ortaya konmuştur.[24] Türkiye'de de beyin göçünün devam etmesi, özellikle koruyucu sağlık hizmetleri ve birinci basamak sağlık hizmetlerinde önemli zorluklara yol açma potansiyeline sahiptir.

Çalışmanın Kısıtlılıkları

Bu araştırma, veri tabanından ulaşılabilen tezlerle sınırlıdır. Farklı tarih aralıkları ve farklı kelimeler ile yapılan taramalarla ulaşılacak çalışmalar için genellenemez.

SONUÇ

Bu araştırma, Türkiye'de sağlık alanında beyin göçünü konu alan lisansüstü tezlerin eğilimlerini ilk kez sistematik biçimde ortaya koymuştur. Bulgular, özellikle 2020 yılından itibaren sağlık çalışanlarının göç eğilimini inceleyen tezlerde belirgin bir artış olduğunu ve hekimler, hemşireler ile sağlık eğitimi alan öğrencilerin başlıca çalışma gruplarını oluşturduğunu göstermektedir.

Sağlık çalışanlarının göç eğilimleri, ekonomik yetersizlikler, çalışma koşullarının olumsuzluğu, sağlıkta şiddet, kariyer sınırlılıkları ve belirsizlikler gibi yapısal sorunlarla yakından ilişkilidir. Bu nedenle sağlık insan gücünün korunmasına yönelik politikaların çok boyutlu olması gerekmektedir. İstihdam olanaklarının güçlendirilmesi, ücret dengesizliklerinin giderilmesi, güvenli ve destekleyici çalışma ortamlarının sağlanması, kariyer gelişim fırsatlarının artırılması ve psikososyal destek mekanizmalarının güçlendirilmesi kritik adımlar arasında yer almaktadır. Halk sağlığı açısından bakıldığında, "sağlık çalışanlarının göç eğilimi, sağlık sisteminde sürdürülebilirliği tehdit eden önemli bir halk sağlığı sorunudur." Bu eğilim devam ettiğinde hizmet sunumu, erişilebilirlik, iş yükü ve hizmet kalitesi üzerinde ciddi olumsuz etkiler ortaya çıkacaktır.

Bu araştırmanın bulgularının, sağlık insan gücü planlamasına yönelik politika yapıcılar için yol gösterici olacağı; lisansüstü öğrencilerin ve akademisyenlerin konu seçimi ve literatür taramalarına katkı sağlayarak gelecekte yürütülecek araştırmalar için sağlam bir temel oluşturacağı düşünülmektedir.

Etik

Etik Kurul Onayı ve Hasta Onayı: Bu araştırmada birincil veri toplama yöntemleri kullanılmamıştır. Veriler kamuoyuna açık YÖK Ulusal Tez Merkezi veri tabanından elde edilmiştir. Çalışmada insan veya hayvan verisi kullanılmamıştır. Bu nedenle çalışma için etik kurul izni alınmasına ihtiyaç duyulmamıştır.

Dipnotlar

Yazarlık Katkıları

Konsept: O.Ç., M.A., Ç.I.Ç., Dizayn: O.Ç., M.A., Ç.I.Ç., Veri Toplama veya İşleme: O.Ç., M.A., A.Ş.G., Analiz veya Yorumlama: O.Ç., M.A., Ç.I.Ç., A.Ş.G., Literatür Arama: O.Ç., M.A., Ç.I.Ç., A.Ş.G., Yazan: O.Ç., M.A., Ç.I.Ç.

Çıkar Çatışması: Yazarlar bu makale ile ilgili olarak herhangi bir çıkar çatışması bildirmemiştir.

Finansal Destek: Çalışmamız için hiçbir kurum ya da kişiden finansal destek alınmamıştır.

KAYNAKLAR

1. Murakami Y. Japan's brain drain: an analysis of Japanese researchers living in the United States. Japanese Economy. 2010;37:23-57.
2. Bongers A, Díaz-Roldán C, Torres JL. Brain drain or brain gain? International labor mobility and human capital formation. J Int Trade Econ Dev. 2021;31:1-25.
3. Akkoç İ, Akkoç M, Çakır O. Development and psychometric analysis of the physician and nurse brain drain attitude scale. J Nurs Meas. 2025;33:443-54.
4. Akyıldız İE. Beyin göçü politikaları: Dünya'dan ve Türkiye'den örnekler. Anadolu Univ İktisadi İdari Bilimler Fakültesi Derg. 2025;26:326-61.
5. Lee ES. A theory of migration. Demography. 1966;3:47-57.
6. Portes A, Celaya A. Modernization for Emigration: determinants and consequences of the brain drain. Daedalus. 2013;142:170-84.
7. Alam GM, Hoque KE. Who gains from "brain and body drain" business-developing/developed world or individuals: a comparative study between skilled and semi/unskilled emigrants. African Journal of Business Management. 2010;4:534-48.
8. Castles S, Miller MJ, Ammendola G. The age of migration: international population movements in the modern world. Am Foreign Policy Interests. 2005;27:537-42.
9. Wallerstein I. The modern world system: capitalist agriculture and the origins of the European world economy in the sixteenth century. Canadian Journal of Political Science. 1977;10:2025-3.
10. Bilgin O, Çelik H, Alan H, Torun S. Hemşirelerin beyin göçüne yönelik tutumları ve çalışma ortamları arasındaki ilişki: tanımlayıcı bir çalışma. Mersin Univ Tıp Fak Lokman Hekim Tıp Tarihi Folklorik Tıp Derg. 2025;5:296-304.
11. Aktürk S, Yanardağ R. Kaygıdan göçe: üniversite lisans öğrencilerinde gelecek kaygısının beyin göçüne yönelik tutumlarına etkisi. Kahramanmaraş Sütçü İmam Univ İktisadi İdari Bilimler Fakültesi Derg. 2025;15:81-99.
12. Ünlü HY, Daşlı Y. Sağlık eğitimi alan öğrencilerin beyin göçüne yönelik tutumları. Sosyal Beşerî İdari Bilimler Derg. 2024;7:606-21.
13. Köklü C, Naldöken Ü. Sağlık çalışanlarının beyin göçüne yönelik tutumlarının incelenmesi: Sivas ili örneği [yüksek lisans tezi]. Sivas Cumhuriyet Üniversitesi Sağlık Bilimleri Enstitüsü. 2024.
14. Öncü E, Selvi H, Vayısoğlu SK, Ceyhan H. Hemşirelik öğrencilerinde beyin göçüne yönelik tutum ölçeği geliştirilmesi: güvenirlik ve geçerlik çalışması. Çukurova Med J. 2018;43:207-15.
15. Yıldırım T. Uluslararası düzeyde sağlık çalışanlarının göçünü yönetme politikaları: genel bir bakış ve Türkiye için bir durum değerlendirmesi. Amme İdaresi Derg. 2010;43:31-65.
16. Saxenian A. From brain drain to brain circulation: transnational communities and regional upgrading. Stud Comp Int Dev. 2005;40:35-61.
17. Tung RL. Brain circulation, diaspora, and international competitiveness. Eur Manag J. 2008;26:298-304.
18. Aydan S. Hekim göçü açısından Türkiye'nin çalışma koşullarının değerlendirilmesi. Hacettepe Sağlık İdaresi Derg. 2023;26:895-920.
19. Erdoğan Kaya A, Erdoğan Aktürk B, Aslan E. Factors predicting the motivation to study abroad in Turkish medical students: a causal investigation into the problem of brain drain. J Health Sci Med. 2023;6:526-31.
20. Boyd M. Family and personal networks in international migration: recent developments and new agendas. Int Migr Rev. 1989;23:638-70.
21. Massey DS, Arango J, Hugo G, Kouaouci A, Pellegrino A, Taylor JE. Theories of international migration: a review and appraisal. Popul Dev Rev. 1993;19:431-66.
22. White MD, Marsh EE. Content analysis: a flexible methodology. Libr Trends. 2006;55:22-45.
23. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evid Based Med. 2016;21:125-7.
24. Karatuzla M. Sağlık çalışanlarında güncel konu incelemesi: beyin göçü. Genel Sağlık Bilimleri Derg. 2024;6:159-71.
25. Köse Tosunöz İ. Brain drain among health professionals: a review study. Dent Med J Rev. 2024;6:97-108.
26. Küçükkendirci H, Yücel M. Evaluation of doctors' predictions of working abroad and their attitudes towards brain drain: a web-based research. Forbes J Med. 2025;6:66-74.

DOI: 10.4274/forbes.galenos.2025.12499
Forbes J Med 2025;6(3):249-258

Original Article / Orijinal Araştırma

Effects of White Noise and Swaddling on Pain and Physiological Parameters During Eye Examination in Healthy Term Infants: A Randomized Controlled Trial

Sağlıklı Term Bebeklerde Göz Muayenesi Sırasında Beyaz Gürültü ve Kundaklamanın Ağrı ve Fizyolojik Parametreler Üzerindeki Etkileri: Randomize Kontrollü Bir Çalışma

Sultan GÜNER BAŞARA[1], Hüsniye ÇALIŞIR[2], Emel BAHADIR ARSU[3], Özdemir ÖZDEMİR[3]

[1]Tokat Gaziosmanpaşa University Erbaa Faculty of Health Sciences, Department of Child Health and Disease Nursing, Tokat, Türkiye
[2]Aydın Adnan Menderes University Faculty of Nursing, Department of Child Health and Diseases Nursing, Aydın, Türkiye
[3]Ankara Bilkent City Hospital, Clinic of Ophthalmology, Neonatal Outpatient Clinic, Ankara, Türkiye

Cite as: Güner Başara S, Çalışır H, Bahadır Arsu E, Özdemir Ö. Effects of white noise and swaddling on pain and physiological parameters during eye examination in healthy term infants: a randomized controlled trial. Forbes J Med. 2025;6(3):249-258

ABSTRACT

Objective: Eye examinations in infants for vision screening can cause pain and stress. The aim of this study was to evaluate the effect of white noise and swaddling on pain, heart rate, and oxygen saturation in healthy term infants undergoing post-discharge eye examinations.

Methods: The study was conducted from September 1, 2022, to March 1, 2023, in the neonatal eye outpatient clinic of a maternity hospital. This randomized controlled trial included 120 term infants randomized to four groups: white noise (n=30), swaddling (n=30), white noise plus swaddling (n=30), and control (n=30). Pain was assessed using the premature infant pain profile scale, and heart rate and peripheral capillary oxygen saturation (SpO_2) were measured by pulse oximetry before, 30 seconds after, and at the end of the examination.

Results: The groups showed comparable pain scores, heart rates, and SpO_2 levels, with no significant differences observed ($p>0.05$). However, all groups showed significant increases in pain scores and heart rates, and a decrease in SpO_2 during the examination compared with baseline ($p<0.001$).

Conclusion: White noise, swaddling, and their combination did not reduce pain or improve physiological parameters during eye examinations in term infants.

Keywords: Eye examination, pain management, swaddling, term infant, white noise

Received/Geliş: 14.10.2025
Accepted/Kabul: 24.11.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Sultan GÜNER BAŞARA, MD,
Tokat Gaziosmanpaşa University Erbaa Faculty of Health Sciences, Department of Child Health and Disease Nursing, Tokat, Türkiye
✉ sgune6377@hotmail.com
ORCID: 0000-0002-0665-5396

ÖZ

Amaç: Yenidoğanlarda görme taraması için yapılan göz muayeneleri ağrı ve stres oluşturabilir. Bu çalışmada, taburculuk sonrası göz muayenesi yapılan sağlıklı term yenidoğanlarda beyaz gürültü ve kundaklamanın ağrı, kalp hızı ve oksijen satürasyonu üzerine etkisi değerlendirildi.

Yöntem: Çalışma, 1 Eylül 2022-1 Mart 2023 tarihleri arasında bir doğum hastanesinin yenidoğan göz polikliniğinde yürütüldü. Randomize kontrollü bu çalışmaya 120 term yenidoğan dahil edildi ve dört gruba ayrıldı: beyaz gürültü (n=30), kundaklama (n=30), beyaz gürültü + kundaklama (n=30) ve kontrol (n=30). Ağrı, kalp hızı ve periferik kapiller oksijen satürasyonu (SpO_2), prematüre bebek ağrı profili ölçeği ve pulse oksimetre kullanılarak muayene öncesi, muayeneden 30 saniye sonra ve muayene sonunda ölçüldü.

Bulgular: Gruplar arasında ağrı skorları, kalp hızı ve SpO_2 seviyeleri açısından anlamlı bir fark gözlenmedi ($p>0,05$). Ancak, tüm gruplarda muayene sırasında bazal değerlere kıyasla ağrı skorları ve kalp hızında anlamlı artış, SpO_2'de ise anlamlı düşüş görüldü ($p<0,001$).

Sonuç: Beyaz gürültü, kundaklama ve bunların kombinasyonu, term yenidoğanlarda göz muayenesi sırasında ağrıyı azaltmadı ve fizyolojik parametreleri iyileştirmedi.

Anahtar Kelimeler: Ağrı yönetimi, beyaz grültü, göz muayenesi, kundaklama, term bebek

INTRODUCTION

Eye problems that arise during early childhood are of critical importance, as they can lead to permanent vision loss.[1] It is recommended that term infants be evaluated for congenital cataracts, congenital glaucoma, retinoblastoma, strabismus, amblyopia, and refractive errors.[2] An estimated 20,00-40,000 infants are born with bilateral cataracts globally each year, leading to blindness in over 14 million children. Congenital glaucoma, a developmental disorder that may be unilateral or bilateral, typically presents at birth or within the first few months of life and can damage the optic nerve, leading to vision loss.[3] Detecting risk factors that may hinder normal visual development through vision screening in infants and providing early treatment help prevent vision loss and blindness. In the country where the study was conducted, healthy term infants undergo detailed eye examinations for conditions such as congenital cataracts and glaucoma, using techniques similar to those employed in retinopathy of prematurity (ROP) screening.

Eye examinations for screening and early diagnosis can be painful and stressful for infants. Exposure to pain in infancy can lead, in the short term, to increased pain sensitivity and altered cortical development, and, in the long-term, to behavioral and learning problems.[4,5] Therefore, it is crucial to minimize pain and stress during these examinations.[6] Although analgesic agents are considered effective for pain control during eye examinations, the most effective and safest combination of interventions that does not increase adverse effects has yet to be determined.[7] Kinoshita et al.[8] Cochrane review indicated that topical anesthesia administered during ROP screenings is insufficient for complete pain relief. The evidence suggests that a single intervention is not sufficient to completely alleviate pain. Systemic pharmacological analgesics used to reduce pain in infants may adversely affect their early brain development, nutrition, socialization, and memory later in life.[8,9] Currently, non-pharmacologic methods to reduce pain in infants are gaining prominence. These methods are more cost-effective, easier to administer, and safer than pharmacologic approaches. They can also be used in combination with pharmacologic treatments.[10]

During eye examinations, Blefastop is used to keep the baby's eyelids open, and a scleral depressor is employed to examine the ocular fundus and to move the eye laterally, which can cause discomfort. The white light source used by the ophthalmologist further unsettles the infant. Studies have explored various non-pharmacologic interventions to reduce pain during eye examinations in preterm infants, including touch, positioning, non-nutritive sucking, breast milk, glucose, sucrose, white noise, and music.[11-19] However, these studies have yielded conflicting results, and the evidence remains insufficient. In our review of the literature, we found no study that definitively supports the claim that swaddling alone significantly reduces pain, nor any research comparing the effectiveness of non-pharmacological methods for reducing pain experienced by term infants during eye examinations. There is a need for studies assessing the efficacy of non-drug interventions in alleviating pain during eye examinations in healthy term infants after hospital discharge. White noise and swaddling were preferred for term infants returning for ophthalmological examination 30-36 days post-discharge due to their ease of use and ready availability.

The purpose of this study was to assess the impact of white noise and swaddling on procedural pain, heart rate, and oxygen saturation (SpO_2) in healthy term infants undergoing hospital eye examinations for post-discharge vision screening.

Hypotheses of the study:

H_1: Playing white noise to healthy term infants during eye examinations affects pain scores, heart rate, or oxygen saturation.

H_2: Swaddling healthy term infants during eye examinations affects pain scores, heart rate, or oxygen saturation.

H_3: Swaddling combined with listening to white noise during eye examinations affects pain scores, heart rate, or oxygen saturation.

METHODS

Design

This study adopted a prospective, randomized, controlled experimental design with four study groups (three intervention groups and one control group), in accordance with CONSORT guidelines.[20]

This randomized controlled trial was registered at ClinicalTrials.gov (registration number:NCT06535984).

Sample and Setting

The study was conducted between September 1, 2022, and March 1, 2023, in the neonatal eye outpatient clinic at a maternity hospital. In this hospital, healthy infants who do not require intensive care after birth receive a single eye screening examination between 30 and 36 days of age. Approximately 400 infants who do not require intensive care attend the neonatal eye outpatient clinic each month. Participants were divided into four groups: three intervention groups and one control group. Infants were assigned to these groups using simple randomization to ensure baseline comparability and to prevent selection bias. Randomization was performed using a computer-

generated random sequence created through the random.org website. To ensure comparability between groups, infants will be assigned to groups by simple random sampling. The allocation list was prepared by an independent researcher who was not involved in data collection or outcome assessment, thereby maintaining allocation concealment. Sample size estimation was performed using G-Power (version 3.0.10). With a 95% confidence interval, α=0.05, effect size =0.159, and power (1-β) =0.80, a total sample size of 27) 108 infants per group) was calculated for the four groups. The effect size (0.159) was not derived from pilot data or a specific previous study; rather, it was determined as a small-to-medium effect size based on general recommendations for behavioral and clinical research, considering the expected variability in pain and physiological parameters among infants. The sample size estimation was based on the primary outcome measure of the study, the premature infant pain profile (PIPP) score. The study included 120 infants divided into the swaddling group (n=30), the white noise group (n=30), the white noise + swaddling group (n=30), and the control group (n=30) (Figure 1).

The sample included infants with gestational age >37 6/7 weeks, birth weight ≥2000 g, and 5-minute Apgar score ≥7, who were aged 30-36 days postpartum and undergoing their first eye screening. The study excluded infants who had received care in the neonatal intensive care unit, who had required, who had a diagnosed or suspected congenital and/or genetic disorder, who had experienced hearing problems, who had undergone any surgical procedure, or who had received systemic analgesics within the last six hours. No infants withdrew during the study period.

Interventions

Healthy term infants who were swaddled during the eye examination formed the swaddling group (intervention group 1). Infants who listened to white noise formed the white-noise group (intervention group 2). Infants who were both swaddled and exposed to white noise during the eye examination constituted the white noise + swaddling group (intervention group 3). The control group received routine care and follow-up. The outpatient clinic nurse (the third researcher) numbered the infants in order of admission and assigned them to groups using simple randomization. All infants were examined by the same ophthalmologist (the fourth investigator). Another outpatient nurse, who was not part of the research team, measured the infants' pain scores, heart rate, and oxygen saturation at three times: before the examination (T1), 30 seconds after the ophthalmologist began examining the first eye (T2), and immediately after finishing the examination of the second eye (T3). No adverse events or complications related to the interventions were observed during the study.

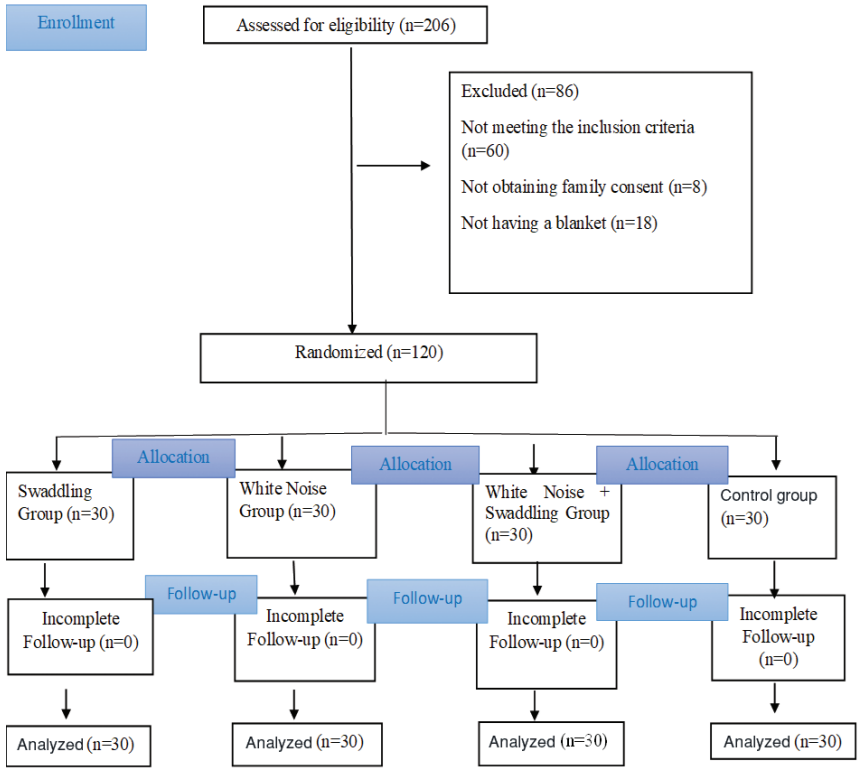


Figure 1. Consort flow diagram

Eye Examination Preparation

All healthy term infants included in the study (intervention groups 1, 2, and 3, and the control group) underwent routine procedures at the outpatient eye clinic prior to the examination. Families with appointments for their baby's eye examination were instructed to feed their infant one hour before the appointment. On the day of the examination, the third investigator inquired about the time of the last feeding. Following this, pupil dilation was initiated in the waiting room. According to the hospital's standard procedure, pupil dilation involved instilling phenylephrine and tropicamide eye drops twice, at five-minute intervals. Approximately 45 minutes later, once the pupils were dilated, each infant was taken to the preparation room. One drop of proparacaine was instilled into each eye of the infants before the examination. Proparacaine-containing drops are used to provide topical anesthesia for rapid and short-term diagnostic or surgical procedures.[21] During these procedures, which occur before the examination, the baby remains with the parent, usually in the mother's arms. Once the infant is taken into the examination room, the parent does not accompany them. Instead, a third researcher brings the infants into the examination room without their parents and places them on the examination table. The left foot of all infants was fitted with a pulse oximetry probe.

Eye Examination Procedure

The third investigator positioned the infant supine and aligned the infant's head in the midline with a slight extension. The eye examination was performed by an ophthalmologist (the fourth researcher). Each eye examination lasted approximately one to one-and-a-half minutes.

Participant Groups

Swaddling group (n=30); The third researcher swaddled each infant in its thin blanket. The swaddling involved bringing the infant's left hand close to the left buttock and wrapping it with the upper right end of the blanket. Similarly, she wrapped the baby's right hand with the blanket's upper-left end by bringing that end closer to the right hip. She wrapped the baby's legs with the lower part of the blanket, ensuring that the baby's head could move freely. All environmental conditions during the procedure were kept constant. The room temperature was maintained at approximately 22-24 °C, and the examination was performed in a dim, quiet environment. No changes in lighting or room temperature were made during the interventions.

White noise group: the third researcher turned on a pre-recorded rain sound file before the start of the eye examination and played it for the infants until the end of the examination. The sound was played on a mobile phone (Samsung Galaxy A32, 2021) at approximately 50-55 dB, as measured with a sound-level meter app, and the speaker was positioned approximately 30-40 cm from the infant's head to avoid discomfort. The same recording was used throughout the study to ensure acoustic consistency.

White noise + swaddling group: the third researcher swaddled the infants in this group as described for the swaddling group, and turned on a pre-recorded rain sound file before the examination began, which was played continuously until the end of the examination.

Control group: The healthy term infants in this group received routine care at the eye outpatient clinic. In routine outpatient clinic practice, infants are examined with their clothes on, and no non-pharmacologic methods are applied during the examination. As in the intervention groups, infants in the control group received proparacaine eye drops and pupil dilation as part of the routine clinical protocol prior to the examination.

A single outpatient clinic nurse who was not part of the research team conducted all data collection activities. This nurse works in the neonatal eye outpatient clinic and has a neonatal intensive care nursing certificate.

Data Collection Tools

"The infant ınformation and monitoring form" and "the PIPP" were used during data collection.

Infant Information and Monitoring Form

Researchers prepared the form, which contained 11 questions.[11,22,23] The form included information about the infant's demographic characteristics, mode of delivery, birth weight, gender, gestational age, Apgar score, postnatal age, diet, oxygen therapy, and phototherapy. Additionally, pain scores (T1, T2, and T3), heart rates, and oxygen saturation levels were recorded.

The Premature Infant Pain Profile

The infants' pain scores were evaluated using the PIPP. The Turkish validity and reliability of the scale developed by Stevens et al.[24] were assessed by Taplak and Bayat.[25] The scale includes seven indicators (heart rate, oxygen saturation, gestational age, behavioral state, nasolabial furrow, brow bulge, and eye squeeze), each scored from 0 to 3. The total score ranges from 0 to 21, with higher scores indicating more severe pain.[24]

Statistical Analysis

Statistical analyses were conducted using IBM's SPSS software (version 27). Descriptive statistics summarized the study findings. The association between two categorical variables was assessed.

Data were assessed using Pearson's chi-square (χ^2) test. For data with a normal distribution, comparisons among three or more independent groups were performed using one-way ANOVA (F-value), while repeated-measures ANOVA (F-value) was applied for comparisons involving three or more related groups. When data were not normally distributed, the Kruskal-Wallis H test (χ^2 statistic) was used to compare three or more independent groups, and the Friedman test (χ^2 statistic) was used to compare three or more dependent groups.

Ethical Approval

All procedures were conducted in compliance with the ethical standards of the Institutional and National Research Committee and within the guidelines of the Declaration of Helsinki. This study was approved by the Ankara Bilkent City Hospital No. 2 Clinical Research Ethics Committee (decision number: E2-22-2419, date: 16.10.2022). Information about the study's aims, procedures, and other relevant details was provided to the infants' parents. They signed a written consent form before the study commenced.

RESULTS

Comparison of The Groups on Terms of Demographic Data

A total of 120 neonates [females n=68 (56.6%) and males n=52 (43.4%)] were included in the study and were randomly assigned to the swaddling (n=30), white noise (n=30), white noise plus swaddling (n=30), and control (n=30) groups. There were no significant differences between the intervention and control groups with respect to mode of delivery, gender, birth weight, gestational age, and 1- and 5-minute Apgar scores. The groups were also similar with respect to feeding method, oxygen therapy, phototherapy, postnatal age, and current weight (p>0.05). The groups were homogeneous in terms of these characteristics. Other demographic data are shown in Table 1.

Comparison of The Groups in Terms of PIPP Score, Heart Rate, and SpO_2

The groups did not differ significantly in PIPP scores before the eye examination (T1), at 30 seconds into the examination (T2), or at the end of the examination (T3) (p>0.05). The PIPP score in the swaddling + white noise group at T3 was lower than in the other groups, although this difference was not statistically significant (Figure 2). The groups did not differ significantly in PIPP scores before the eye examination (T1, p=0.601), at 30 seconds into the examination (T2, p=0.108), and at the end of the examination (T3, p=0.526). The PIPP score in the swaddling + white noise group at T3 was lower than in the other groups, although this difference was not statistically significant (Figure 2). The mean PIPP score of all infants at T1 was significantly lower than that at T2 and T3 (p<0.001). In both the white noise group and the swaddling group, the mean PIPP score at T3 was lower than at T2 (p<0.001). Heart rate values were comparable between groups at T1 (p=0.418), T2 (p=0.892), and T3 (p=0.742). Heart rates at T1, T2, and T3 varied significantly across all infant groups (p<0.001). In all three intervention groups (swaddling + white noise, white noise, and swaddling), heart rate values at T1 were lower than those at T2 and T3, while values at T3 were lower than those at T2 (p<0.001). In the control group, only the heart rates at T1 were lower than those at T2 and T3 (p<0.001). There were no statistically significant differences in SpO_2 between groups at T1 (p=0.710), T2 (p=0.230), and T3 (p=0.854). At T3, although the SpO_2 value in the white noise group was slightly higher than in the other groups and those in the control and swaddling groups were slightly lower, these differences were not statistically significant. SpO_2 values in all groups at T2 and T3 were lower than at T1 (p<0.001). In addition, the SpO_2 values of infants in the white noise group at T3 were higher than at T2 (p<0.001) (Table 2).

Primary Outcomes

The primary endpoints of the study were the total pain score from PIPP, the instantaneous heart rate measured by the monitor, and the oxygen saturation.

DISCUSSION

In this randomized controlled trial, we investigated the effects of swaddling and white noise—alone and in combination—on pain, heart rate, and oxygen saturation in healthy term infants undergoing routine post-discharge eye examinations. Although our study was methodologically robust, our findings indicated that these interventions did not significantly reduce pain scores or improve heart rate and oxygen saturation during the examination. The analysis showed no differences in pain scores, heart rates, or saturation values among the groups before, 30 seconds into, or at the end of the eye examination. However, all groups exhibited an increase in pain and heart rate and a decrease in oxygen saturation during the examination compared with baseline.

Newborns experience varying degrees of pain during invasive procedures, such as eye examinations conducted for diagnosis and treatment.[26] During painful interventions, infants experience increases in respiratory rate, heart rate, and blood pressure, and a decrease in oxygen saturation.[27] Studies have shown that infants experience significant pain even when local anesthetics are used during the procedure.[26,28]

White noise is a constant, uniform sound composed of

Table 1. Descriptive characteristics of infants (n=120)

Characteristics	Swaddling group (n=30)		White noise group (n=30)		Swaddling + white noise group (n=30)		Control group (n=30)		Test* p-value
	n	%	n	%	n	%	n	%	
Mode of delivery Vaginal Caesarean	 12 18	 40.0 60.0	 14 16	 46.7 53.3	 15 15	 50.0 50.0	 13 17	 43.3 56.7	 χ2=0.673 p=0.879
Gender Girl Boy	 15 15	 50.0 50.0	 18 12	 60.0 40.0	 17 13	 56.7 43.3	 18 12	 60.0 40.0	 χ2=0.814 p=0.846
Feeding Breast milk Formula food Both	 21 1 8	 70.0 3.3 26.7	 25 1 4	 83.3 3.4 13.3	 28 - 2	 93.3 - 6.7	 25 2 3	 83.3 6.7 10.0	 χ2=7.882 p=0.247
Oxygen therapy Yes No	 8 22	 26.7 73.3	 8 22	 26.7 73.3	 3 27	 10.0 90.0	 6 24	 20.0 80.0	 χ2=3.385 p=0.336
Phototherapy application Yes No	 6 24	 20.0 80.0	 6 24	 20.0 80.0	 7 23	 23.3 76.7	 5 25	 16.7 83.3	 χ2=0.417 p=0.937
	M ± SD	Mdn (IQR)	M ± SD	Mdn (IQR)	M ± SD	Mdn (IQR)	M ± SD	Mdn (IQR)	Test** p-value
Birth weight	3280.50±431.96	3222.5 (465.0)	3249.83±415.41	3282.5 (498.0)	3224.83±409.1	3297.5 (408.0)	3137.67±550.14	3190.0 (558.0)	χ2=2.230 p=0.526
Week of gestation	39.00±0.98	39.0 (2.0)	39.00±1.08	39.0 (2.0)	38.70±0.99	38.0 (1.0)	38.83±0.91	39.0 (1.0)	χ2=2.682 p=0.443
Apgar 1 minute	7.73±0.69	8.0 (1.0)	7.37±0.56	7.0 (1.0)	7.27±0.74	7.0 (1.0)	7.63±0.62	8.0 (1.0)	χ2=7.265 p=0.064
Apgar 5 minutes	9.17±0.38	9.0 (0.0)	8.97±0.49	9.0 (0.0)	8.90±0.66	9.0 (0.0)	9.07±0.45	9.0 (0.0)	χ2=3.905 p=0.272
Postnatal age	40.63±7.53	40.0 (9.0)	42.07±6.01	40.0 (5.0)	41.9±10.91	40.0 (5.0)	39.53±5.83	40.0 (4.0)	χ2=1.831 p=0.608
Current weight	4444.83±639.58	4400.0 (813.0)	4363.57±597.52	4276.0 (725.0)	4339.00±835.49	4225.0 (900.0)	4263.50±624.29	4245.0 (763.0)	χ2=1.157 p=0.763

*Pearson X2test, **Kruskal-Wallis test, Mdn: Median, M: Mean, IQR: Interquartile range, SD: Standard deviation

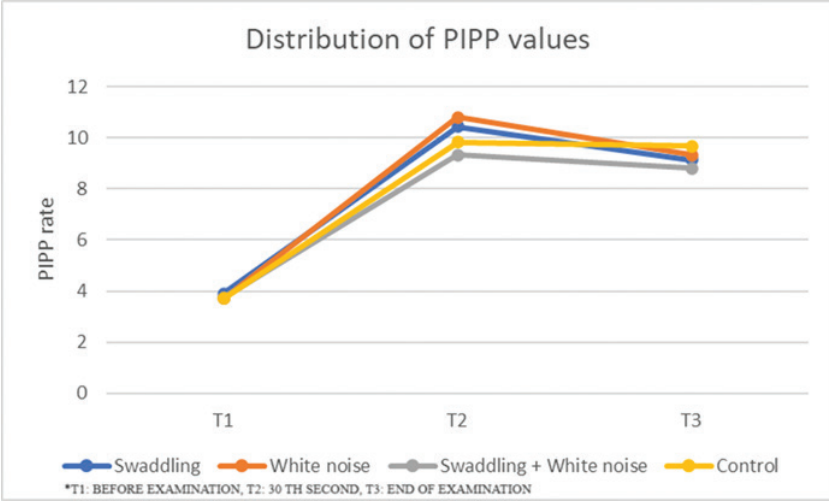


Figure 2. Distribution of PIPP values across groups at different measurement times (T1: before eye examination; T2: 30 seconds of eye examination; T3: End of eye examination).

Error bars represent the variability of the data (standard deviation/standard error as applicable)

PIPP: Premature infant pain profile

many different frequencies (e.g., the howling of the wind, the flow of water, or the running of a fan). The potential mechanism of action of white noise may be explained by desensitization and sensory modulation theories, which suggest that consistent auditory input can help regulate the infant's sensory responses and reduce reactivity to painful stimuli.[29] White noise can have a calming effect on emotional states, reducing pain and anxiety. It also elicits stochastic resonance in hearing, which can positively affect behavioral and physiological responses.[17] Some studies[30,31] have shown that white noise reduces pain scores during invasive procedures, is effective in increasing weight gain and improving comfort, but has no positive effect on pain scores during severely painful procedures such as endotracheal aspiration.[22] Contrary to studies showing that white noise during eye examinations decreased pain scores and positively affected saturation and heart rate[17,18] a studies found no positive effect on pain scores and physiological parameters.[23] Similarly, our findings indicate that the pain scores (PIPP>9-10 points), heart rates, and oxygen saturation of infants exposed to white noise were not improved, suggesting that applying white noise during eye examinations may be ineffective in reducing pain. Considering that infants' crying during the examination may have reduced their ability to perceive the white noise, future studies should explore whether using a slightly higher sound intensity within safe limits or adjusting the timing of the intervention—such as starting the white noise earlier or maintaining it longer—could enhance its effectiveness.

Swaddling is a method that reduces excessive movement by wrapping the baby's arms and legs.[32] Safe swaddling of newborns can reduce the pain experienced during medical interventions, support neuromuscular development, and reduce physiological and behavioral stress.[32,33] Swaddling alone has been reported to relieve pain during invasive procedures, such as aspiration.[34] However, it has been concluded that in premature infants, swaddling alone is not effective in reducing pain during retinopathy examinations, nor is it effective when combined with other non-pharmacologic methods such as sucrose, breast milk, or distilled water.[35] Metreş and Yıldız[12] similarly reported that swaddling combined with a pacifier was more effective at reducing pain during eye examinations in premature infants than a pacifier alone. Another study reported that a modified developmental care bundle—including environmental modifications, swaddling, oxygen supplementation, and cue-based individualized care—can reduce pain and stress in premature infants during ROP examinations.[36] The results of our study indicate that this method did not alleviate pain or positively influence physiological responses during highly painful procedures, such as eye examinations performed on healthy term infants. These findings suggest that the intense pain induced by detailed eye examinations may not be adequately managed by swaddling alone or by swaddling combined with white noise. Although the differences in PIPP scores between the groups were not statistically significant, the lower pain scores observed in infants who received the combined white noise and swaddling intervention and the higher scores observed in the control group suggest that the combined application of these methods may provide partial benefits in pain management. In addition, all infants received proparacaine eye drops as part of

Table 2. Comparison of infants' premature infant pain profile, heart rate, and SpO_2 values at T1, T2, and T3

Variable	Swaddling group (n=30)		White noise group (n=30)		Swaddling + white noise group (n=30)		Control group (n=30)		Test p-value	Effect size (η^2)
PIPP	**M ± SD**	**Mdn (IQR)**	**M ± SD**	**Mdn (IQR)**	**M ± SD**	**Mdn (IQR)**	**M ± SD**	**Mdn (IQR)**		
T1 (1)	3.90±0.61	4.0 (0.0)	3.70±0.92	4.0 (1.0)	3.73±1.14	4.0 (1.0)	3.70±0.95	4.0 (1.0)	$\chi2$=1.866 p=0.601*	0.000
T2 (2)	10.43±1.89	10.0 (2.0)	10.80±2.28	12.0 (2.0)	9.33±2.75	9.5 (4.0)	9.83±2.52	10.0 (4.0)	$\chi2$=6.069 p=0.108*	0.035
T3 (3)	9.13±1.61	9.0 (2.0)	9.33±1.86	10.0 (2.0)	8.80±2.47	8.0 (3.0)	9.67±2.09	10.0 (3.0)	$\chi2$=2.230 p=0.526*	0.002
Test p-value	$\chi2$=51.254 **p<0.001**¥ (1<2,3) (2>3)		$\chi2$=52.667 **p<0.001**¥ (1<2,3) (2>3)		$\chi2$=47.345 **p<0.001**¥ (1<2,3)		$\chi2$=48,828 **p<0.001**¥ (1<2,3)‡			
Heart rate	M ± SD	Mdn (IQR)	M ± SD	Mdn (IQR)	M ± SD	Mdn (IQR)	M ± SD	Mdn (IQR)		
T1 (1)	147.83±14.17	150.0 (22.0)	142.77±13.55	144.0 (25.0)	145.23±14.31	143.0 (23.0)	141.67±18.98	143.0 (31.0)	F=0.953 p=0.418*	0.024
T2 (2)	166.67±14.14	165.5 (16.0)	165.87±19.24	166.5 (22.0)	167.17±18.05	168.0 (30.0)	162.70±19.58	163.5 (23.0)	$\chi2$=0.620 p=0.892*	0.000
T3 (3)	159.77±13.97	161.5 (12.0)	156.63±18.20	158.0 (24.0)	159.77±20.67	163.0 (31.0)	157.23±15.87	159.0 (19.0)	$\chi2$=1.248 p=0.742*	0.002
Test p-value	$\chi2$=38.034 **p<0.001**¥ (1>2,3) (2>3)		F=20.505 **p<0.00**§ (1>2,3) (2>3)		$\chi2$=41.496 **p<0.001**¥ (1>2,3) (2>3)		$\chi2$=21.529 **p<0.001**¥ (1>2,3)‡			
SpO_2	M ± SD	Mdn (IQR)	M ± SD	Mdn (IQR)	M ± SD	Mdn (IQR)	M ± SD	Mdn (IQR)		
T1 (1)	95.53±1.94	95.5 (3.0)	95.43±1.65	95.0 (3.0)	95.10±1.79	95.0 (2.0)	95.10±1.86	95.0 (2.0)	F=0.461 p=0.710*	0.011
T2 (2)	90.10±4.13	91.0 (5.0)	90.40±3.82	90.0 (5.0)	89.70±3.53	90.0 (4.0)	91.37±4.26	92.0 (6.0)	$\chi2$=4.309 p=0.230*	0.019
T3 (3)	91.83±3.59	92.0 (3.0)	92.47±3.06	93.0 (4.0)	91.63±4.23	92.5 (4.0)	92.43±3.24	92.0 (3.0)	$\chi2$=0.781 p=0.854*	0.000
Test p-value	$\chi2$=36.207 **p<0.001**¥ (1>2,3)‡		F=36.727 **p<0.00**§ (1>2,3) (2<3)‡		$\chi2$=34.207 **p<0.001**¥ (1>2,3)‡		$\chi2$=25.304 **p<0.001**¥ (1>2,3)‡			

¥Kruskal-Wallis test, §ANOVA test, **Repeated Measures test, *Friedman test, M: Mean, Mdn: Median, IQR: Interquartile range, SD: Standard deviation

‡Comparison of the groups' scores at T1 (1) before the eye examination, T2 (2) at the 30th second of the eye examination, and T3 (3) at the end of the eye examination

T1: Before eye examination, T2: 30 seconds of eye examination, T3: End of eye examination

the routine clinical protocol prior to the examination. The analgesic effect of this topical anesthetic may have reduced the overall pain response during the procedure, thereby masking the differences between the intervention and control groups. Furthermore, in our study, the second measurement (T2) was taken at the 30th second of the eye examination, a time previous research has identified as when infants typically exhibit the most intense pain and stress responses. This standardized time point ensured comparability between groups. However, the duration of the eye examination varied slightly among infants (approximately 1-1.5 minutes per eye). This variability may have influenced the T3 measurements, as some infants were exposed to the procedure for a longer duration, potentially resulting in differences in cumulative stress or in recovery time. Another limitation is that the temporal recovery of pain and physiological parameters after the procedure could not be evaluated, as no T4 measurement was obtained. This prevented the assessment of whether any intervention provided an advantage during the recovery phase. Additionally, due to the nature of the interventions, it was not possible to blind the nurse performing the assessments, which may have introduced observer bias in the evaluation of pain scores. The findings of the study highlight the need to develop new, more effective strategies to manage pain during detailed eye examinations in healthy term infants.

Study Limitations

Infants who presented for an eye examination were assessed in a quiet, dark environment. However, the infants' crying during the procedure might have reduced their ability to hear the white noise, diminishing its effectiveness. The pain scores and physiological values of the infants returning to normal levels (T4) could not be assessed. Another limitation of the study was that the nurse who evaluated the infants' physiological findings and pain scores could not be blinded because of the study design. Video recordings were not used for the blinded assessment of PIPP scores. Therefore, the possibility of observer bias cannot be completely ruled out; such bias may have influenced the evaluation of pain scores.

CONCLUSION

In this study, healthy term infants experienced procedure-related pain during detailed eye examinations, accompanied by a significant increase in heart rate and a decrease in oxygen saturation. Our findings suggest that non-pharmacological interventions, such as white noise and swaddling, whether applied individually or in combination, may not be sufficient during invasive procedures such as eye examinations, thereby indicating the need to revise pain management protocols. To better understand the effectiveness of different non-pharmacological methods in managing pain and changes in heart rate and oxygen saturation during eye examinations of term infants, larger randomized controlled trials are recommended.

Ethics

Ethics Committee Approval: This study was approved by the Ankara Bilkent City Hospital No. 2 Clinical Research Ethics Committee (decision number: E2-22-2419, date: 16.10.2022).

Informed Consent: Information about the study's aims, procedures, and other relevant details was provided to the infants' parents. They signed a written consent form before the study commenced.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ö.Ö., Concept: S.G.B., H.Ç., E.B.A., Ö.Ö., Design: S.G.B., H.Ç., Data Collection or Processing: S.G.B., E.B.A., Ö.Ö.,Analysis or Interpretation: S.G.B., H.Ç., E.B.A., Literature Search: S.G.B., H.Ç., Writing: S.G.B., H.Ç., E.B.A.

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

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

REFERENCES

1. World Health Organization. Ethics: includes universal newborn eye screening in their first ever postnatal care guidelines [Internet]. 2022 [cited 2025 Feb XX]. Available from: https://www.iapb.org/blog/who-includes-universal-newborn-eye-screening-in-their-first-ever-postnatal-care-guidelines/
2. Malik AN, Evans JR, Gupta S, et al. Universal newborn eye screening: a systematic review of the literature and review of international guidelines. J Glob Health. 2022;12:12003.
3. Bell SJ, Oluonye N, Harding P, Moosajee M. Congenital cataract: a guide to genetic and clinical management. Ther Adv Rare Dis. 2020;1:2633004020938061.
4. Griffiths N, Spence K, Loughran-Fowlds A, Westrup B. Individualised developmental care for babies and parents in the NICU: evidence-based best practice guideline recommendations. Early Hum Dev. 2019;139:104840.
5. Duerden EG, Miller SP. Pain in the newborn brain: a neural signature. Lancet Digit Health. 2020;2:e442-3.
6. Pollaci M, Schlenk EA, Baum C, Godfrey K. Supportive interventions to reduce pain and stress during ophthalmic examinations for retinopathy of prematurity in premature infants. Adv Neonatal Care. 2021;21:274-9.
7. Sindhur M, Balasubramanian H, Srinivasan L, Kabra NS, Agashe P, Doshi A. Intranasal fentanyl for pain management during screening for retinopathy of prematurity in preterm infants: a randomized controlled trial. J Perinatol. 2020;40:881-7.

8. Kinoshita M, Olsson E, Borys F, Bruschettini M. Opioids for procedural pain in neonates. Cochrane Database Syst Rev. 2023;4:CD015056.
9. Mozetic V, Cruz MFSD, Cruz NFSD, Polizelli MU, Moraes NSBD. Analysis of cochrane systematic reviews about retinopathy of prematurity. Revista Brasileira de Oftalmologia. 2021;80:42-8. https://doi.org/10.5935/0034-7280.20210008
10. Thirunavukarasu AJ, Hassan R, Savant SV, Hamilton DL. Analgesia for retinopathy of prematurity screening: a systematic review. Erratum in: Pain Pract. 2023;23:127.
11. Sun Y, Zhang J, Chen X, et al. Effectiveness of gentle human touch for pain control during examination for retinopathy of pre-maturity: a randomized controlled trial. Front Pediatr. 2020;8:608378.
12. Metreş Ö, Yıldız S. Pain Management with ROP position in Turkish preterm infants during eye examinations: a randomized controlled trial. J Pediatr Nurs. 2019;49:e81-9.
13. Özkan TK, Yüksel ED, Akar S. Effect of non- nutritive sucking on pain during the examination of retinopathy of prematurity. Journal Neonatal Nursing. 2022;28:155-8.
14. Naik A, D'Lima A, Sreekumar K, Silveira MP. Efficacy of expressed breast milk alone or in combination with paracetamol in reducing pain during ROP screening: a randomized controlled trial. Journal of Clinical Neonatology.2021;10:73
15. Nayak R, Nagaraj KN, Gururaj G. Prevention of pain during screening for retinopathy of prematurity: a randomized control trial comparing breast milk, 10% dextrose and sterile water. Indian J Pediatr. 2020;87:353-8.
16. Erçelik ZE, Yilmaz D. Nonpharmacological applications during the retinopathy of prematurity examination and their effects on pain control: a systematic review and meta-analysis. J Perinat Neonatal Nurs. 2022;36:297-304.
17. Ren XF, Wang ZZ, Yang M, Li L, Kong XY, Feng ZC. [Clinical effect of white noise combined with glucose in reducing the pain of retinopathy screening in preterm infants]. Zhongguo Dang Dai Er Ke Za Zhi. 2019;21:1159-63.
18. Dur Ş, Ustabaş Yıldız N. The effect of white noise and classical music on pain and physiologic parameters in preterm infants during retinopathy of prematurity examinations: a randomized controlled trial. Early Child Development and Care. 2022;193:60-71.
19. Corrigan MJ, Keeler JR, Miller HD, Ben Khallouq BA, Fowler SB. Music therapy and retinopathy of prematurity screening: using recorded maternal singing and heartbeat for post exam recovery. J Perinatol. 2020;40:1780-8.
20. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c869.
21. Iselaimy R, Al Tawil L, Abouammoh MA. Anesthesia in retinopathy of prematurity. Saudi J Ophthalmol. 2022;36:251-9.
22. Taplak AŞ, Bayat M. Comparison the effect of breast milk smell, white noise and facilitated tucking applied to turkish preterm infants during endotracheal suctioning on pain and physiological parameters. J Pediatr Nurs. 2021;56:e19-e26.
23. Ren X, Li L, Lin S, Zhong C, Wang B. Effects of white noise on procedural pain-related cortical response and pain score in neonates: a randomized controlled trial. Int J Nurs Sci. 2022;9:269-77.
24. Stevens BJ, Gibbins S, Yamada J, et al. The premature infant pain profile-revised (PIPP-R): initial validation and feasibility. Clin J Pain. 2014;30:238-43
25. Taplak AŞ, Bayat M. Psychometric testing of the Turkish version of the premature infant pain profile revised-PIPP-R. J Pediatr Nurs. 2019;48:e49-e55.
26. Dammann O, Hartnett ME, Stahl A. Retinopathy of prematurity. Dev Med Child Neurol. 2023;65:625-31.
27. Ahsan MS, Kalamdani P, Kalathingal T, Patra S, Manerkar S, Mondkar J. Evaluation of pain and physiological stress during targeted neonatal echocardiography. J Neonatal Perinatal Med. 2022;15:89-93.
28. Fajolu IB, Dedeke IOF, Ezenwa BN, Ezeaka VC. Non-pharmacological pain relief interventions in preterm neonates undergoing screening for retinopathy of prematurity: a systematic review. BMJ Open Ophthalmol. 2023;8:e001271.
29. Pickens TA, Khan SP, Berlau DJ. White noise as a possible therapeutic option for children with ADHD. Complement Ther Med. 2019;42:151-5.
30. Cetinkaya S, Yavas Celik M, Ozdemir S. Effect of white noise on alleviating the pain of new-born during invasive procedures. J Matern Fetal Neonatal Med. 2022;35:1426-32.
31. Kahraman A, Gümüş M, Akar M, Sipahi M, Bal Yılmaz H, Başbakkal Z. The effects of auditory interventions on pain and comfort in premature newborns in the neonatal intensive care unit; a randomised controlled trial. Intensive Crit Care Nurs. 2020;61:102904.
32. Melo GMD, Cardoso ML, Almeida PCD, Rodrigues EC. Effect of music combined with swaddling on pain in full-term newborns: Randomized clinical trial. Revista Brasileira de Enfermagem. 2022;75:1-8.
33. Pillai Riddell RR, Bucsea O, Shiff I, et al. Non-pharmacological management of infant and young child procedural pain. Cochrane Database Syst Rev. 2023;6:CD006275.
34. Ayyıldız TK, Tanrıverdi E, Tank DY, Akkoç B, Topan A. The effect of swaddling method applied to preterm infants during the aspiration procedure on pain. J Pediatr Nurs. 2023;70:61-7.
35. Dolgun G, Bozlak Ş. Effect of nonpharmacologic pain control during examination for retinopathy of prematurity. Journal of Obstetric, Gynecologic & Neonatal Nursing. 2017;46:709-15.
36. Chuang LJ, Wang SH, Ma MC, Lin CN, Chen CL, Huang MC. A modified developmental care bundle reduces pain and stress in preterm infants undergoing examinations for retinopathy of prematurity: a randomised controlled trial. J Clin Nurs. 2019;28:545-59.

DOI: 10.4274/forbes.galenos.2025.14880
Forbes J Med 2025;6(3):259-265

Original Article / Orijinal Araştırma

Comparison of Nurses' Digital Literacy Levels: A Study in Digital and Non-Digital Hospitals

Hemşirelerin Dijital Okuryazarlık Düzeylerinin Karşılaştırılması: Dijital ve Dijital Olmayan Hastanelerde Bir Çalışma

Mustafa ORHAN[1], Süleyman MERTOĞLU[2]

[1]İzmir Bakırçay University Faculty of Health Sciences, Department of Health Management, İzmir, Türkiye
[2]İzmir Provincial Health Directorate, İzmir, Türkiye

Cite as: Orhan M, Mertoğlu S. Comparison of nurses' digital literacy levels: a study in digital and non-digital hospitals. Forbes J Med. 2025;6(3):259-265

ABSTRACT

Objective: The aim of this study is to compare the digital literacy levels of nurses working in digital and non-digital hospitals and to examine the effect of demographic and professional variables on their digital competence.

Methods: This study was designed as a descriptive comparative cross-sectional study and included 219 nurses working in one digital and one non-digital hospital in Türkiye. Data were collected using a structured questionnaire comprising demographic information and a 12-item unidimensional Digital Literacy Scale. The reliability of the scale was confirmed (Cronbach's alpha=0.948). Statistical analyses were performed using SPSS, including descriptive statistics, independent samples t-tests, one-way analysis of variance, Games-Howell post-hoc tests, and chi-square tests.

Results: Nurses working in digital hospitals had significantly higher levels of digital literacy than nurses in non-digital hospitals ($p<0.05$). No significant differences were observed based on age, education level, or years of experience. However, digital literacy scores increased significantly with greater frequency of digital tool use. No significant relationship was found between professional experience and the frequency of digital interaction.

Conclusion: Corporate digital infrastructure and frequency of digital tool usage are key factors affecting nurses' digital literacy. Encouraging daily digital interactions and providing access to technological resources may help strengthen digital competencies in healthcare settings.

Keywords: Digital literacy, hospital, health technology, nurse

ÖZ

Amaç: Bu çalışmanın amacı, dijital ve dijital olmayan hastanelerde çalışan hemşirelerin dijital okuryazarlık düzeylerini karşılaştırmak ve demografik ve mesleki değişkenlerin dijital yeterlilikleri üzerindeki etkisini incelemektir.

Yöntem: Tanımlayıcı ve karşılaştırmalı kesitsel bir çalışma olarak tasarlanan araştırmaya, Türkiye'de bir dijital ve bir dijital olmayan hastanede çalışan 219 hemşire katılmıştır. Veriler, demografik bilgileri ve 12 maddelik, tek boyutlu bir Dijital Okuryazarlık Ölçeği'ni içeren yapılandırılmış bir ölçek aracılığıyla toplanmıştır. Ölçeğin güvenilirliği doğrulanmıştır (Cronbach's alpha=0,948). İstatistiksel analizler, tanımlayıcı istatistikler, bağımsız örneklem t-testleri, tek yönlü ANOVA, Games-Howell post-hoc testleri ve ki-kare testleri içeren SPSS kullanılarak gerçekleştirilmiştir.

Bulgular: Dijital hastanelerde çalışan hemşirelerin, dijital olmayan hastanelerdeki hemşirelere kıyasla anlamlı derecede daha yüksek dijital okuryazarlık düzeyleri olduğu bulunmuştur ($p<0,05$). Yaş, eğitim düzeyi veya deneyim yılına göre anlamlı bir fark gözlenmemiştir. Ancak, dijital okuryazarlık puanları, dijital araç kullanım sıklığının artmasıyla önemli ölçüde artmıştır. Mesleki deneyim ile dijital etkileşim sıklığı arasında anlamlı bir ilişki bulunamamıştır.

Received/Geliş: 16.11.2025
Accepted/Kabul: 04.12.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Mustafa ORHAN, PhD,
İzmir Bakırçay University, Department of Health Management, İzmir, Türkiye
m.orhan@bakircay.edu.tr
ORCID: 0000-0002-4598-1717

Sonuç: Kurumsal dijital altyapı ve dijital araç kullanım sıklığı, hemşirelerin dijital okuryazarlığını etkileyen temel faktörlerdendir. Günlük dijital etkileşimi teşvik etmek ve teknolojik kaynaklara erişimi sağlamak, sağlık hizmetleri ortamlarında dijital yeterliliklerin güçlendirilmesine yardımcı olabileceği söylenebilir.

Anahtar Kelimeler: Dijital okuryazarlık, hastane, hemşire, sağlık teknolojisi

INTRODUCTION

The impact of digitalization on healthcare services is undeniable and involves comprehensive change processes that require healthcare professionals to develop new competencies.[1-3] Hospitals, which encompass numerous and complex specialties, possess substantial information and communication technology infrastructure.[4] Digitalization extends from electronic documentation to robot-assisted procedures and can improve care quality, patient safety, and patient outcomes.[5] Digitalization is expected to digitalization will alleviate the challenges facing the healthcare system, which is grappling with demographic change and a shortage of qualified professionals.[5]

With developments in the delivery of healthcare services, the role of nursing professionals can no longer be limited to providing care that addresses only patients' physical needs.[6] The digitization of healthcare delivery has created a new relationship for nurses -between nurses and technology. While technology cannot replace compassion today, it has become an important competency that nurses must acquire.[7] Digital literacy is required to acquire these technological competencies. Digital literacy is generally defined as the knowledge, skills, and attitudes necessary to effectively use information and communication technologies.[8] Digital literacy is seen as a competency related to the ability to access, evaluate, produce, and share information, encompassing technical skills and information literacy.[9,10] To implement this, recent studies have developed assessment tools to measure nurses' knowledge, skills, and attitudes toward digital technologies.[11] These tools emphasize that digital literacy in nursing is not only the ability to use devices but also to integrate data into clinical reasoning and patient care.

The importance of digital literacy in healthcare service delivery is increasing, while its absence poses challenges. Nurses with sufficient digital skills can support accurate documentation, clinical decision-making, and patient education, while the lack of these competencies can jeopardize patient safety and quality of care.[12] For example, a lack of digital competence has been associated with increased error rates, compromised safety, and reduced service quality.[13] Conversely, positive outcomes have been demonstrated in healthcare providers with higher digital capacity. Research shows that electronic medical records and related tools provide benefits such as increased documentation accuracy, compliance, and decision support.[14] Hospitals with a high level of digital maturity (HIMSS EMRAM stages 6-7) achieve significantly better safety and quality outcomes than hospitals with lower digital maturity.[15] Hospitals with higher levels of digital maturity have achieved better safety outcomes and lower infection rates, reflecting more effective use of data in patient care. Similarly, more digitally mature hospitals report higher patient experience scores (e.g., communication and care coordination) than less digitally mature hospitals.[15]

Empirical studies directly examining the relationship between an institution's level of digitalization and nurses' digital literacy are limited. Therefore, this research demonstrates that the digital literacy of frontline personnel, such as nurses, given the institution's level of digital infrastructure, is crucial to realizing the security and quality benefits of health information technology.

METHODS

This study was designed as a descriptive, comparative cross-sectional study. The primary objective is to examine the digital literacy levels of nurses working in both digital and non-digital hospitals. A hospital accredited at level 7 by HIMSS EMRAM was selected as the digital hospital. For the non-digital hospital, a hospital with a similar number of beds was selected. The research examines whether the level of hospital digitalization affects nurses' digital competencies and explores the relationship between digital literacy and various demographic and professional factors. Ethics committee approval has been obtained from the Non-Interventional Research Ethics Committee of İzmir Bakırçay University (decision number: 2106, research number: 2096, date: 07.03.2025). Participants gave their voluntary consent after being informed in detail about the purpose, method, possible risks and rights of the research.

Population and Sample

The study population consists of all nurses working at one digital hospital and one non-digital hospital in Türkiye. The total number of nurses in the two hospitals is reported to be 502. The formula for a known population size was used to determine the required sample size: $n=(N \times t^2 \times p \times q) / [(d^2 \times (N - 1)) + (t^2 \times p \times q)]$. According to this calculation, the minimum sample size was set at 218 nurses. The final sample consisted of 219 participants who met the inclusion criteria.

Hypotheses

The following hypotheses were tested in the study:

H1: The digital literacy levels of nurses working in digital hospitals are significantly different from those of nurses in non-digital hospitals.

H2: Nurses' digital literacy levels vary significantly according to their personal information.

- H2a: Nurses' digital literacy levels differ significantly according to age groups.
- H2b: Nurses' digital literacy levels differ significantly by education level.
- H2c: Nurses' digital literacy levels differ significantly according to years of professional experience.
- H2d: Nurses' digital literacy levels differ significantly according to their unit of work.
- H2e: Nurses' digital literacy levels differ significantly according to daily digital tool usage frequency.

Statistical Analysis

The data used in the study were collected through face-to-face interviews. The Digital Literacy Scale, whose validity and reliability were established by other researchers, was administered alongside a form collecting the participants' demographic characteristics. The scale used to assess participants' digital literacy was developed by Nabiyeva et al.[16] in a study conducted among employees of the Psychiatry Clinic at the Akdeniz University Faculty of Medicine. It is a five-point Likert-type self-report scale, consisting of a single dimension and 12 items, that assesses adults' ability to use various digital technologies correctly and to access, produce, and share accurate information. Statistical analysis was performed using SPSS software. Initially, descriptive statistics (mean, standard deviation, skewness, and kurtosis) were used to summarize the distribution of digital literacy scores.

The reliability of the digital literacy scale was assessed using Cronbach's alpha coefficient, where a value above 0.70 is considered acceptable for internal consistency. Independent-samples t-tests were used to compare mean digital literacy scores between two groups (e.g., hospital or department type). For categorical variables with more than two levels (e.g., age group, education level, professional experience, frequency of digital tool use), one-way analysis of variance (ANOVA) was used to test for significant differences in mean scores across groups. Games-Howell post-hoc tests were applied when ANOVA results were statistically significant. This method was preferred because it does not assume homogeneity of variance and is robust to unequal sample sizes.

Additionally, chi-square tests of independence were used to investigate associations between categorical variables such as education level and frequency of digital technology use. All statistical tests were two-tailed, and results were considered significant at $p<0.05$.

RESULTS

This section presents the results of statistical analyses examining the digital literacy levels of nurses in digital and non-digital hospitals. First, descriptive statistics and reliability analyses are presented to summarize the overall digital literacy scores. Subsequently, inferential statistical tests such as independent-samples t-tests, one-way ANOVAs, post-hoc comparisons, and chi-square tests were used to evaluate the study hypotheses regarding the effects of institution type, professional experience, educational background, and digital usage patterns.

A total of 219 participants volunteered to take part in the study as shown in Table 1. By age group, 43.8% of participants were 34 years of age or younger, 24.7% were 35-45 years of age, and 31.5% were 46 years of age or older. In terms of education, 10% had an associate degree or less, 73.5% had a bachelor's degree, and 16.4% had a graduate degree. More than half of the participants (53.4%) worked in non-digital hospitals (46.6% in digital hospitals), and 56.6% worked in administrative roles (43.4% in clinical roles). Approximately 63.0% had 16 years or more of professional experience; 20.5% had 10 years or less, and 16.4% had 11-15 years. Most participants reported using digital tools daily (66.1%); 13.8% used digital tools several times a day, 13.3% used them several times a week, and only 6.9% reported never using digital tools.

As shown in Table 2, the 12-item Digital Literacy Scale demonstrated excellent internal consistency with a Cronbach's alpha value of 0.948. This high alpha value indicates that the items reliably measure a single underlying construct, reflecting high internal consistency of the scale.

The average digital literacy score is around 4.20 (on a scale of 1-5) and the standard deviation is 0.84. Skewness (-0.04) and kurtosis (-1.27) indicate that the distribution is approximately symmetric and slightly flatter than normal (platykurtic).

To examine the relationship between the daily frequency of digital use among study participants and various demographic factors, a chi-square test of independence was performed. As shown in Table 3, only hospital type showed a statistically significant relationship with the frequency of use [χ^2 (3)=11.733, p=0.008]. This indicates that daily digital usage differs between digital and non-digital hospitals. In contrast, no significant relationship was found between usage frequency and education level (p=0.090),

age group (p=0.282), or years of experience (p=0.591). Although education, age, and experience do not show a significant relationship with the frequency with which staff use digital tools, hospital type (digital or non-digital) does.

The one-way ANOVA results presented in Table 4 indicate no statistically significant differences in mean scores among age, education, and professional experience categories. In practical terms, respondents' age, highest

Table 1. Descriptive statistics

Variable	Category	n	%
Age	34 years and under	96	43.8
	35-45	54	24.7
	46 years and over	69	31.5
Educational status	Associate degree	22	10.0
	Undergraduate degree	161	73.5
	Postgraduate degree	36	16.4
Hospital type	Non-digital	117	53.4
	Digital hospital	102	46.6
Service provided	Administrative	124	56.6
	Clinical	95	43.4
Professional experience	10 years and under	45	20.5
	11-15 years	36	16.4
	16 years and over	138	63.0
Daily digital usage frequency	Never use	15	6.9
	A few times a week	29	13.3
	Every day	144	66.1
	Every few hours	30	13.8

Table 2. Reliability and distribution properties of the Digital Literacy Scale

Cronbach's alpha		Number of items		
0.948		12		
n	**Mean**	**SD**	**Skewness**	**Kurtosis**
219	4.20	0.84	-0.04	-1.27

SD: Standard deviation

Table 3. Chi-square test results

Variable (chi-square)	χ^2 (SD)	p
Hospital type	11.733 (3)	0.008
Educational status	10.947 (6)	0.090
Age	7.442 (6)	0.282
Professional experience	4.636 (6)	0.591

SD: Standard deviation

Table 4. One-way ANOVA by age, education, experience and ınternet usage frequency

Factor	F	p
Age	2.224	0.111
Educational status	2.223	0.111
Professional experience	1.902	0.152
Internet usage frequency	3.214	0.001

ANOVA: One-way analysis of variance

level of education, and length of work experience do not appear to significantly affect mean scores.

The one-way ANOVA comparing groups based on daily digital usage frequency reveals a significant difference in mean scores [F (3, 214)=5.920, p=0.001]. This result indicates that there is a difference between usage frequency groups, with at least one group's mean score being significantly different from the others. Given that the overall test was significant, a post-hoc analysis was conducted to determine which usage groups differed.

Games-Howell post-hoc comparisons are used to identify specific groups that differ in mean scoresas shown in Table 5. The analysis indicates that participants who use digital tools "every few hours" scored significantly higher than those who use them "every day", "several times a week", and "never": the average differences are approximately 0.32 (p=0.010), 0.71 (p=0.013), and 0.89 (p=0.018), respectively. No other statistically significant pairwise differences were found between usage frequency groups. In summary, more frequent digital use (multiple times per day) is associated with higher average scores, while other usage groups do not differ significantly from one another.

The independent-samples t-test comparing the average scores of staff working in non-digital and digital hospitals found a statistically significant difference (t=-2.045, p=0.042), as presented in Table 6. Participants working in digital hospitals reported a higher average score ($\bar{x}$=4.3235) than those working in non-digital hospitals ($\bar{x}$=4.0919). The findings show the positive effect of working in digital hospitals on nurses' digital literacy levels.

An independent-samples t-test comparing administrative and clinical service groups found no significant difference in mean scores (t=-0.623, p=0.534). Employees in administrative roles ($\bar{x}$=4.1687) and employees in clinical roles ($\bar{x}$=4.2404) had very similar mean scores. This indicates that there is no significant difference in the mean score on the scale between service types (administrative and clinical).

When the established hypotheses were examined, H1 was accepted because the digital literacy of nurses working in hospitals differed significantly according to the hospitals' digital level. However, H2 was only partially accepted because a significant difference in internet usage frequency (H2e) was observed across demographic groups.

Table 5. Games-Howell post-hoc analysis

(I) Usage frequency	(J) Usage frequency	Mean difference (I-J)	Std. error	p	95 % CI lower	95 % CI upper
I never use it	A few times a week	-0.18429	0.32625	0.942	-1.0694	0.7009
	Every day	-0.57049	0.26343	0.175	-1.3244	0.1835
	Every few hours	-0.89167*	0.26581	0.018	-1.6497	-0.1336
A few times a week	Never use	0.18429	0.32625	0.942	-0.7009	1.0694
	Every day	-0.38619	0.21406	0.289	-0.9642	0.1918
	Every few hours	-0.70738*	0.21699	0.013	-1.2922	-0.1226
Every day	Never use	0.57049	0.26343	0.175	-0.1835	1.3244
	A few times a week	0.38619	0.21406	0.289	-0.1918	0.9642
	Every few hours	-0.32118*	0.10021	0.010	-0.5840	-0.0584
Every few hours	Never use	0.89167*	0.26581	0.018	0.1336	1.6497
	A few times a week	0.70738*	0.21699	0.013	0.1226	1.2922
	Every day	0.32118*	0.10021	0.010	0.0584	0.5840

*The mean difference is significant at the 0.05 level

Table 6. Independent sample t-test results

Groups mean	± SD	t	p
Non-digital hospital	4.0919±0.8382	-2.045	0.042
Digital hospital	4.3235±0.8336		
Administrative service	4.1687±0.8353	-0.623	0.534
Clinical service	4.2404±0.8538		

SD: Standard deviation

DISCUSSION

Digitalization plays a crucial role in healthcare delivery, just as it does in every other aspect of life. This study aims to compare the digital literacy levels of nurses working in digital and non-digital hospitals and to examine the impact of demographic and professional characteristics on digital literacy. The findings reveal that nurses working in digital hospitals have significantly higher levels of digital literacy than their colleagues in non-digital hospitals. This supports the hypothesis that institutional digital infrastructure can positively influence individuals' interactions with, and proficiency in using, digital tools.

The findings of the current study are consistent with those of Erbir[17] and, Şahin and Seçer,[18] but they also reveal significant differences. Consistent with both studies, the current research shows that more frequent digital interaction (such as daily use of digital tools) is associated with higher levels of digital literacy among nurses. This supports the idea that regular exposure to digital environments develops practical digital competencies. However, contrary to the findings of Şahin and Seçer,[18] who reported significant differences in digital literacy based on age and gender, the current study did not observe statistically significant differences between age groups or between male and female participants. Similarly, while Erbir[17] found that younger nurses and those with less experience had higher digital literacy scores, no such difference emerged in this study. The differing findings may stem from contextual factors such as regional differences in institutional infrastructure or educational practices.

The findings of this study align with existing research indicating that nurses in digitally advanced hospitals tend to have higher digital literacy than those in less digitally advanced hospitals. Comparcini et al.[19] found that nurses exposed to a greater number of digital resources scored significantly higher on digital health literacy assessments, thereby supporting the idea that digitalization promotes digital competence. Similarly, Hariyati et al.[20] reported that nurses in managerial roles and with higher education levels, often associated with institutions with greater digital integration, demonstrated better digital information skills, which further emphasized the impact of professional context and organizational support.

Additionally, these studies emphasize that demographic factors such as age, education, and position in the nursing hierarchy influence digital literacy levels through interaction with the digital environment. For example, Hariyati et al.[20] highlighted the relationship between educational level and computer literacy, underscoring the need to tailor digital literacy programs to these characteristics.[20] Furthermore, it has been argued that infrastructure and continuous professional development opportunities are vital for nurses to develop their digital skills and adapt to changing health technologies.[21]

Study Limitations

The study should be evaluated with certain limitations in mind. The study was conducted among nurses at two hospitals; therefore, the sample is not generalizable. In addition, the cross-sectional design of the study is another limitation.

CONCLUSION

This study provides an important overview of how digitalization shapes nurses' professional practices and, in particular, their digital literacy levels. From this perspective, digital hospitals and high digital literacy among healthcare workers are of substantial importance. The results emphasize the positive effect of working in digitally integrated hospital environments on nurses' digital literacy. Furthermore, frequent use of digital tools plays an important role in strengthening digital competence among healthcare workers. This information underscores the importance of continuous digital interactions and infrastructure investments in healthcare settings. Considering the study results and existing literature demonstrating the contribution of digitalization to nurses' care processes, increasing the number of digital hospitals would be beneficial. Furthermore, strengthening the educational curriculum and organizing in-house training programs aimed at developing nurses' digital skills appear to be important requirements. Further research is recommended to investigate longitudinal effects and assess how digital literacy impacts clinical outcomes and patient care efficiency.

Ethics

Ethics Committee Approval: Ethics committee approval was obtained from the Non-Interventional Research Ethics Committee of İzmir Bakırçay University (decision number: 2106, research number: 2096, date: 07.03.2025).

Informed Consent: Participants gave their voluntary consent after being informed in detail about the purpose, method, possible risks and rights of the research.

Footnotes

Authorship Contributions

Concept: M.O., Design: M.O., S.M., Data Collection or Processing: S.M., Analysis or Interpretation: M.O., S.M., Literature Search: M.O., S.M., Writing: M.O., S.M.

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

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

REFERENCES

1. Krick T, Huter K, Domhoff D, Schmidt A, Rothgang H, Wolf-Ostermann K. Digital technology and nursing care: a scoping review on acceptance, effectiveness and efficiency studies of informal and formal care technologies. BMC Health Serv Res. 2019;19:400.
2. Öberg U, Orre CJ, Isaksson U, Schimmer R, Larsson H, Hörnsten Å. Swedish primary healthcare nurses' perceptions of using digital eHealth services in support of patient self-management. Scand J Caring Sci. 2018;32:961-70.
3. Zeeb H, Pigeot I, Schüz B; Leibniz-WissenschaftsCampus Digital Public Health Bremen. Digital Public Health – ein Überblick [Digital public health-an overview]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2020;63:137-44.
4. Fachinger U, Mähs M. Digitalisierung und Pflege. In: Klauber J, Geraedts M, Friedrich J, Wasem J, editors. Krankenhaus-Report 2019: Das digitale Krankenhaus. Berlin, Heidelberg: Springer; 2019. p. 115-28.
5. Wong BLH, Khurana MP, Smith RD, et al. Harnessing the digital potential of the next generation of health professionals. Hum Resour Health. 2021;19:50.
6. Çelik A, Mertoğlu S. Palyatif bakım veren hemşirelerin hasta merkezli bakım yetkinliğinin bakım kalitesine etkisi. Forbes J M. 2024;5:95-101.
7. Burkoski V, Yoon J, Hutchinson D, Hall TNT, Solomon S, Collins BE. Generational differences in hospital technology adoption: a cross-sectional study. Nurs Leadersh (Tor Ont). 2019;32:86-97.
8. Ferrari A. Digital competence in practice: an analysis of frameworks. vol 10. Publications Office of the European Union Luxembourg. 2012.
9. Kuek A, Hakkennes S. Healthcare staff digital literacy levels and their attitudes towards information systems. Health Informatics J. 2020;26:592-612.
10. Shudayfat T, Hani SB, Al Qadire M. Assessing digital health literacy level among nurses in Jordanian hospitals. Electronic Journal of General Medicine. 2023;20:em525.
11. Bulut A, Kasap Rİ, Yılmaz N. Reliability and validity of the Turkish version of the digital competence questionnaire for nurses. BMC nursing. 2025;24:1225.
12. Salahuddin L, Ismail Z. Classification of antecedents towards safety use of health information technology: a systematic review. Int J Med Inform. 2015;84:877-91.
13. Rahal RM, Mercer J, Kuziemsky C, Yaya S. Factors affecting the mature use of electronic medical records by primary care physicians: a systematic review. BMC Med Inform Decis Mak. 2021;21:67.
14. Hariyati RT, Yani A, Eryando T, Hasibuan Z, Milanti A. The Effectiveness and efficiency of nursing care documentation using the SIMPRO model. Int J Nurs Knowl. 2016;27:136-42.
15. Snowdon A, Hussein A, Danforth M, Wright A, Oakes R. Digital maturity as a predictor of quality and safety outcomes in US hospitals: cross-sectional observational study. J Med Internet Res. 2024;26:e56316.
16. Nabiyeva N, Yapar D, Yardımcı A, Metin Ö. COVID-19 pandemisi sürecinde telepsikiyatri hizmetleri kullanımı ve dijital okuryazarlık ilişkisi. In: 6. Uluslararası 24. Ulusal Halk Sağlığı Kongresi Bildiri Kitabı. Presented at: 6. Uluslararası 24. Ulusal Halk Sağlığı Kongresi; 2022; Antalya. 2022.
17. Erbir M. Digital literacy in the nursing profession: the case of Kayseri province. Ekonomi İşletme Siyaset ve Uluslararası İlişkiler Dergisi. 2021;7:336-52.
18. Şahin G, Seçer Ş. An Analysis On Digital Literacy Levels Of Health Employees: The Case Of İzmir. Kırklareli Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi. 2024;13:93-126.
19. Comparcini D, Simonetti V, Tomietto M, et al. The relationship between nurses' digital health literacy and their educational levels, professional roles, and digital attitudes: a cluster analysis based on a cross-sectional study. J Clin Nurs. 2025;34:2885-97.
20. Hariyati RTS, Handiyani H, Wildani AA, Afriani T, Nuraini T, Amiruddin MH. Disparate digital literacy levels of nursing manager and staff, specifically in nursing informatics competencies and their causes: a cross-sectional study. J Healthc Leadersh. 2024;16:415-25.
21. Seki EA. Digital Skills in the 21st century: understanding the impact of digital literacy on nurses. Research on Education and Psychology. 2024;8:360-72.
22. Baumann LA, Baker J, Elshaug AG. The impact of electronic health record systems on clinical documentation times: a systematic review. Health Policy. 2018;122:827-36.

DOI: 10.4274/forbes.galenos.2025.90267
Forbes J Med 2025;6(3):266-272

Original Article / Orijinal Araştırma

The Relationship of Clinical and Molecular Factors with Prognosis in Metastatic Colorectal Cancer

Metastatik Kolorektal Kanserde Klinik ve Moleküler Faktörlerin Prognoz ile İlişkisi

Evrim ATACA[1], Merve KESKİNKILIÇ[2], Sulen SARIOGLU[3], Yasemin BASBINAR[4], Tuğba YAVUZŞEN[2]

[1]Dokuz Eylül University Faculty of Medicine, Department of Internal Medicine, İzmir, Türkiye
[2]Dokuz Eylül University Faculty of Medicine, Department of Internal Medicine, Division of Medical Oncology, İzmir, Türkiye
[3]Dokuz Eylül University Faculty of Medicine, Department of Pathology, İzmir, Türkiye
[4]Dokuz Eylül University Faculty of Medicine, Department of Translational Oncology, Institute of Oncology, İzmir, Türkiye

Cite as: Ataca E, Keskinkılıç M, Sarioglu S, Basbinar Y, Yavuzşen T. The relationship of clinical and molecular factors with prognosis in metastatic colorectal cancer. Forbes J Med. 2025;6(3):266-272

ABSTRACT

Objective: Metastatic colorectal cancer (mCRC) remains a major global health challenge with highly variable survival outcomes. Both tumor-related and host-related factors contribute to prognosis, yet the combined influence of comorbidity burden, rat sarcoma (RAS) mutation status, and age has not been clearly established. This study aimed to evaluate the prognostic impact of the Charlson Comorbidity index (CCI), RAS mutation status, and age on overall survival (OS) and progression-free survival (PFS) in patients with mCRC.

Methods: A retrospective analysis was conducted on 446 patients with histologically confirmed mCRC treated at a tertiary oncology center between 2010 and 2016. Demographic, clinical, and molecular data were collected. OS and PFS were analyzed using Kaplan-Meier estimates and compared with log-rank tests. Independent prognostic factors were determined using multivariate Cox proportional hazards models.

Results: Among 446 patients, 56.7% had a CCI ≥7 and 38.6% were aged ≥65 years; 48.7% harbored RAS mutations. Higher comorbidity burden (CCI ≥7) was significantly associated with shorter OS and PFS (OS: 27.4 vs. 46.8 months, p<0.001; PFS: 31.2 vs. 41.6 months, p=0.002). Multivariate analysis confirmed CCI ≥7 [hazard ratio (HR) =1.82, 95% confidence interval (CI) 1.31-2.52, p<0.001] and RAS mutation (HR =1.47, 95% CI 1.08-2.00, p=0.014) as independent predictors of poorer OS, whereas age lost significance after adjustment.

Conclusion: CCI and RAS mutation status independently predict worse OS and PFS in mCRC. Incorporating comorbidity assessment into prognostic models may enhance personalized treatment strategies in real-world practice.

Keywords: Metastatic colorectal cancer, Charlson Comorbidity index, RAS mutation, overall survival, progression-free survival

ÖZ

Amaç: Metastatik kolorektal kanser (mKRK), dünya genelinde yüksek mortaliteye sahip önemli bir sağlık sorunudur ve hastalar arasında sağkalım sonuçları değişkendir. Tümörle ve hasta ile ilişkili faktörlerin her ikisi de prognozu etkilemektedir; ancak komorbidite yükü, RAS mutasyon durumu ve yaşın birlikte sağkalım üzerindeki etkisi yeterince aydınlatılamamıştır. Bu çalışmanın amacı, mKRK hastalarda Charlson Komorbidite indeksi (CCI), RAS mutasyon durumu ve yaşın genel sağkalım (OS) ve progresyonsuz sağkalım (PFS) üzerindeki prognostik etkisini değerlendirmektir.

Yöntem: 2010-2016 yılları arasında üçüncü basamak onkoloji merkezinde tedavi gören, histolojik olarak doğrulanmış mKRK tanılı 446 hastanın retrospektif analizi yapılmıştır. Demografik, klinik ve moleküler

Received/Geliş: 09.11.2025
Accepted/Kabul: 27.11.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Evrim ATACA, MD,
Dokuz Eylül University Faculty of Medicine, Department of Internal Medicine, İzmir, Türkiye
evrimbayrakdar@gmail.com
ORCID: 0000-0001-6572-0277

veriler toplanmıştır. OS ve PFS, Kaplan-Meier yöntemleriyle analiz edilerek log-rank testleri ile karşılaştırılmış; bağımsız prognostik faktörler çok değişkenli Cox regresyon modeli ile belirlenmiştir.

Bulgular: Dört yüz kırk altı hastanın %56,7'sinde CCI ≥7, %38,6'sında yaş ≥65 yıl ve %48,7'sinde RAS mutasyonu saptanmıştır. Yüksek komorbidite yükü (CCI ≥7), anlamlı olarak daha kısa GS (27,4 vs. 46,8 ay; p<0,001) ve PFS (31,2 vs. 41,6 ay; p=0,002) ile ilişkili bulunmuştur. Çok değişkenli analizde CCI ≥7 [risk oranı (HR) =1,82; %95 güven aralığı (GA): 1,31-2,52; p<0,001] ve RAS mutasyonu (HR =1,47; %95 GA: 1,08-2,00; p=0,014) bağımsız olarak daha kötü OS ile ilişkili bulunmuş; yaş ayarlama sonrası anlamlılığını kaybetmiştir.

Sonuç: CCI ve RAS mutasyon durumu, mKRK'de bağımsız olarak daha kötü genel ve PFS öngörmektedir. Komorbidite değerlendirmesinin prognostik modellerde yer alması, gerçek yaşam pratiğinde kişiselleştirilmiş tedavi stratejilerinin geliştirilmesine katkı sağlayabilir.

Anahtar Kelimeler: Metastatik kolorektal kanser, Charlson Komorbidite indeksi, RAS mutasyonu, genel sağkalım, progresyonsuz sağkalım

INTRODUCTION

Colorectal cancer (CRC) remains one of the leading causes of cancer-related morbidity and mortality worldwide, ranking third in incidence and second in cancer-related deaths.[1] Advances in molecular profiling and targeted therapies have improved outcomes for patients with metastatic CRC (mCRC); however, survival remains highly variable and is influenced by both tumor-related and host-related factors.[2,3] Among these factors, patients' comorbidity burden and molecular alterations, such as RAS mutations, are increasingly recognized as key determinants of prognosis and treatment response.[4,5]

The Charlson Comorbidity index (CCI) is a widely validated tool that quantifies comorbidity burden by assigning weighted scores to various chronic diseases, providing an estimate of 10-year mortality risk.[6] Several studies have shown that a higher CCI score is associated with poorer survival across multiple malignancies, including CRC.[7-9] However, limited data exist regarding its prognostic relevance in patients with metastatic disease, especially in real-world settings where comorbidity burden can strongly influence treatment selection and tolerability.

In addition to comorbidities, RAS mutations are among the most clinically relevant molecular markers in CRC.[10,11] These mutations are known predictors of resistance to anti-epidermal growth factor receptor (EGFR) therapies and are often associated with distinct tumor biology and outcomes.[12] While both CCI and RAS mutation status have been individually linked to survival, their combined and age-adjusted impact on real-world survival outcomes, including overall survival (OS) and progression-free survival (PFS), has not been well characterized.

Therefore, this study aimed to evaluate the prognostic significance of CCI, RAS mutation status, and age in patients with mCRC treated at a tertiary oncology center. By integrating clinical, molecular, and comorbidity data, this analysis provides a comprehensive assessment of host- and tumor-related factors influencing both overall and PFS in a real-world mCRC cohort.

METHODS

Study Design and Population

This retrospective, single-center study included adult patients (≥18 years) diagnosed with CRC and followed at a tertiary medical oncology center between January 2010 and December 2016.

Patients were eligible if they had a confirmed diagnosis of mCRC at presentation or during follow-up, underwent RAS mutation analysis, received systemic treatment and follow-up care at our tertiary medical oncology center, and had complete clinical, pathological, and survival data.

Patients were excluded if they lacked confirmed metastatic disease, had incomplete records, were lost to follow-up, or were treated exclusively at external centers. After applying these criteria, 446 patients were included in the final analysis.

Data Collection and Variables

Demographic, clinical, and pathological data were retrospectively collected from hospital records. Collected variables included age at diagnosis, sex, tumor stage at presentation, tumor location (right- vs. left-sided), primary site (colon vs. rectum), comorbidity burden (CCI), and RAS mutation status. Molecular analyses included RAS [Kirsten rat sarcoma (KRAS), neuroblastoma rat sarcoma (NRAS)] and B-raf proto-oncogene (BRAF) mutation status with codon-specific evaluation (codons 12, 13, and 61 for KRAS/NRAS; codons 600 and 464 for BRAF). Mutation detection was conducted using validated PCR-based assays, including nested PCR, minisequencing, and allele-specific real-time PCR performed on the COBAS z480 platform.

The CCI was calculated according to the original scoring system described by Charlson et al.[6] age adjustment was performed using the Quan modification, in which 1 point is added for each decade above 50 years (i.e., 50–59 = +1, 60–69 = +2, 70–79 = +3, ≥80 = +4).[13] For the age-adjusted version, one additional point was assigned for each decade above 50 years. The index includes major comorbid conditions such as cardiovascular, pulmonary, hepatic, renal, and metabolic diseases, as well as malignancy and AIDS, each weighted according to its impact on mortality risk.

The total score represents the patient's overall comorbidity burden.

Tumor stage was determined according to the 8th edition of the American Joint Committee on Cancer staging system.[14] Baseline laboratory data included serum carcinoembryonic antigen and carbohydrate antigen 19-9 levels.

Survival Definitions and Follow-Up

OS was defined as the interval between the date of CRC diagnosis and either death from any cause or the date of last follow-up. The last follow-up date for survival ascertainment was March 2022. Survival outcomes were obtained from institutional medical records and verified against the national death registry to ensure their accuracy.

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were used to summarize patient characteristics. Continuous variables were expressed as mean ± standard deviation (SD) or median (range), depending on the distribution, while categorical variables were summarized as counts and percentages.

The CCI was analyzed both as a continuous variable and as a dichotomized variable according to the median value (CCI <7 vs. ≥7) to evaluate its prognostic impact. Group comparisons for categorical variables (e.g., RAS mutation status, tumor location, age group) were performed using the chi-square (χ^2) test or Fisher's exact test, as appropriate. Pearson's and Spearman's correlation coefficients were calculated to assess associations between ordinal or interval variables.

PFS was defined as the time from diagnosis of metastatic disease to documented disease progression or death from any cause, whichever occurred first. OS was defined as the time from the date of CRC diagnosis to death from any cause or to the last follow-up. PFS analysis was performed in 416 patients with available follow-up data, whereas OS was assessed in the entire cohort of 446 patients. Survival distributions were estimated using the Kaplan-Meier method, and comparisons between groups were performed using the Log-rank (Mantel-Cox) and Breslow (Generalized Wilcoxon) tests.

To examine the independent effect of clinical and pathological variables on survival, Cox proportional hazards regression analyses were conducted. Variables with a p value <0.10 in univariate analysis were included in the multivariate model. Hazard ratios and 95% confidence intervals were reported. A two-tailed p value 0.05 was considered statistically significant for all analyses.

Ethical Considerations

Ethics committee approval was obtained from the Non-Interventional Research Ethics Committee of Dokuz Eylül University (decision number: 2021/01-07, date: 04.01.2021).

RESULTS

Patient Characteristics and Demographic Data

A total of 446 patients were included in the analysis, of whom 165 (37.0%) were female and 281 (63.0%) were male. The mean age was 59.7±11.6 years, with 65.7% (n=293) of patients aged ≥65 years. The majority of patients were diagnosed with colon cancer (85.2%; n=380), while 14.8% (n=66) were diagnosed with rectal cancer.

The primary tumor was located in the left colon in 77.6% of patients (n=346) and in the right colon in 22.4% (n=100). At the time of diagnosis, 60.5% (n=270) of patients presented with stage IV disease.

At initial diagnosis, the most frequent site of metastasis was the liver (38.6%, n=172), followed by the lung (9.6%, n=43). During follow-up, secondary metastases most commonly involved the liver (72.2%, n=322) and lung (59.0%, n=263).

Charlson Comorbidity Index Distribution

CCI ranged from 0 to 14 (mean ± SD, 6.66±2.61; median: 7). The majority of patients had a moderate-to-high comorbidity burden (CCI ≥6). The most common scores were 7 (17.5%) and 8 (16.8%). Only a small proportion of patients had minimal comorbidity (CCI ≤2, 7.6%).

Mutation Status

In the whole cohort, 200 patients (44.8%) harbored RAS mutations. Among RAS mutations, KRAS alterations were most frequent (42.1%, n=188), followed by BRAF alterations (3.8%, n=17) and NRAS alterations (3.6%, n=16). Within the KRAS-mutant subgroup, codon 12 mutations predominated (30.3%, n=135), followed by codon 13 mutations (9.2%, n=41) and codon 61 mutations (2.7%, n=12). NRAS mutations occurred mainly at codon 61 (1.6%), while BRAF mutations were primarily at codon 600 (3.4%).

Progression-Free Survival According to Comorbidity Burden

Kaplan-Meier analysis demonstrated a significant association between CCI and PFS in patients with mCRC. Patients with lower comorbidity scores (CCI <7) had a median PFS of 18.2 months (95% CI: 15.4-21.0), whereas those with higher scores (CCI ≥7) had a median PFS of 9.9 months (95% CI: 8.1-11.6). The mean estimated PFS was 28.2±2.1 months for the low-CCI group and 17.2±1.3 months for the high-CCI group. The difference between survival distributions was statistically significant according to both

the Log-Rank test (χ^2 =26.48, p<0.001) and the Breslow test (χ^2 =37.48, p<0.001) (Table 1). These findings indicate that patients with a higher comorbidity burden experience significantly shorter PFS.

Overall Survival According to Comorbidity Burden

Kaplan-Meier analysis revealed that CCI was significantly associated with OS in patients with mCRC. Patients with CCI <7 had a median OS of 52.9 months (95% CI: 46.2-59.6), whereas those with CCI ≥7 had a markedly shorter median OS of 27.8 months (95% CI: 23.9-31.7) (Figure 1). The mean estimated OS was 64.7±3.4 months for patients with low CCI and 41.9±31.1 months for those with high CCI (The mean OS estimate in the high-CCI group was affected by censoring, resulting in a large standard error).

The difference in OS between the two groups was statistically significant according to both the Log-Rank test (χ^2 =25.51, p<0.001) and the Breslow test (χ^2 =41.91, p<0.001). Censoring rates were 24.9% in the CCI <7 group and 15.0% in the CCI ≥7 group (Table 2). These findings indicate that a higher comorbidity burden is associated with substantially poorer OS.

Association Between RAS Mutation Status and Comorbidity Burden

A significant association was observed between RAS mutation status and CCI group (Table 3). Among patients with RAS wild-type (RAS WT) tumors, 50.8% had CCI <7, whereas 49.2% had CCI ≥7.

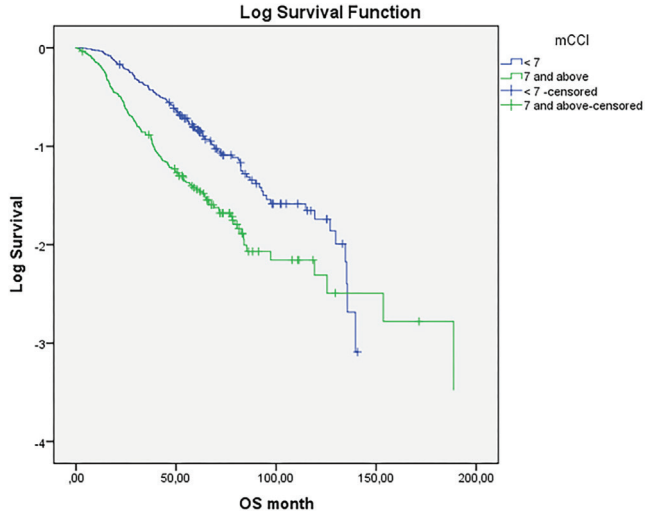


Figure 1. Kaplan-Meier curves for overall survival (OS) according to Charlson Comorbidity index (CCI) group

Table 1. Comparison of PFS between comorbidity groups based on median CCI (n=416)

CCI group	Mean PFS (months) ± SE	95% CI (mean)	Median PFS (months) ± SE	95% CI (median)	Events (n)	Log-rank p
<7	28.23±2.06	24.20-32.27	18.20±1.45	15.36-21.05	181	**<0.001**
≥7	17.17±1.28	14.67-19.68	9.87±0.89	8.11-11.62	235	
Overall	21.99±1.18	19.67-24.30	13.77±0.77	12.26-15.27	416	

CCI: Charlson Comorbidity index, OS: Overall survival, PFS: Progression-free survival, CI: Confidence interval, SE: Standard error

Table 2. OS according to CCI group (n=446)

CCI group	Mean OS (months) ± SE	95% CI (mean)	Median OS (months) ± SE	95% CI (median)	Events (n)	Censored (%)	Log-rank p
<7	64.69±3.36	58.11-71.27	52.90±3.40	46.23-59.57	145	24.9 %	**<0.001**
≥7	41.86±31.06	0-1027.3*	27.80±1.98	23.91-31.69	215	15.0 %	
Overall	38.10±2.28	33.63-42.57	—	—	360	19.3 %	

*Wide CI due to right censoring at longest follow-up time.
CCI: Charlson Comorbidity index, OS: Overall survival, PFS: Progression-free survival, CI: Confidence interval, SE: Standard error

Table 3. Association between RAS mutation status and CCI group (n=446)

RAS mutation status	CCI <7 (n, %)	CCI ≥7 (n, %)	Total (n, %)
RAS wild-type	125 (50.8%)	121 (49.2%)	246 (55.2%)
RAS mutant	68 (34.0%)	132 (66.0%)	200 (44.8%)
Total	193 (43.3%)	253 (56.7%)	446 (100%)

Pearson χ^2 =12.70, df =1, p<0.001 , Spearman's p=0.169, p<0.001
CCI: Charlson Comorbidity index, RAS: Rat sarcoma

In contrast, patients with RAS-mutant tumors had a higher comorbidity burden: 66.0% had CCI ≥7. The difference in distribution between RAS mutation groups was statistically significant (χ^2 =12.70, p<0.001). A weak but significant positive correlation was found between RAS mutation status and comorbidity level (Spearman's ρ=0.169, p<0.001), indicating that RAS-mutant patients tended to present with higher CCI scores.

Association Between Comorbidity Burden and Primary Tumor Sidedness

No statistically significant association was found between CCI group and primary tumor sidedness. Among patients with CCI <7, 81.3% had left-sided tumors and 18.7% had right-sided tumors. In those with CCI ≥7, 74.7% had left-sided tumors and 25.3% had right-sided tumors. Although right-sided tumors were numerically more frequent in patients with higher comorbidity burden, this difference did not reach statistical significance (χ^2 =2.78, p=0.096). The weak negative correlation between CCI and tumor sidedness (Spearman's ρ=-0.079, p=0.096) also did not indicate a meaningful relationship.

Progression-Free Survival According to Age Group

The mean age at diagnosis was 61.1±11.8 years. Patients were dichotomized into two age groups: <65 years (n=259) and ≥65 years (n=157). Kaplan-Meier analysis revealed no significant difference in PFS between the two age groups.

The median PFS was 13.9 months (95% CI: 12.1-15.7) for patients aged <65 years and 13.0 months (95% CI: 10.2-15.8) for those aged ≥65 years (Log-rank test: χ^2 =2.60, p=0.107). Similarly, the Breslow (Wilcoxon) test did not indicate a significant difference (p=0.641). These findings suggest that age at diagnosis was not a determining factor for PFS in this cohort.

Overall Survival According to Age Group

When stratified by age, 274 patients (61.4%) were younger than 65 years and 172 patients (38.6%) were 65 years or older. Kaplan-Meier survival analysis demonstrated a significant difference in OS between the two groups (Log-rank χ^2 =7.62, p=0.006; Breslow χ^2 =7.20, p=0.007) (Figure 2). The median OS was 40.7 months (95% CI: 34.6-46.8) in patients aged <65 years, compared with 30.2 months (95% CI: 22.1-38.3) in those aged ≥65 years (Table 4). Although the mean OS estimate was numerically higher in the younger group, this estimate was limited by censoring. These results indicate that older age was associated with shorter OS in this cohort.

Overall Survival by Age Group Stratified According to Charlson Comorbidity Index

When analyses were stratified by CCI level, the impact of age on OS was not statistically significant after adjustment. Among patients with CCI below the median (<7), the median OS was 50.7 months (95% CI: 42.6-58.8) for patients aged <65 years and 58.5 months (95% CI: 39.6-77.3) for those aged ≥65 years. In contrast, among patients with CCI ≥7, the median OS was 30.7 months (95% CI: 21.6-39.8) for younger patients and 24.2 months (95% CI: 19.5-28.9) for older patients. The log-rank test, adjusted for CCI, did not show a significant difference in OS between age groups (χ^2 =2.77, p=0.096; Breslow χ^2 =1.58, p=0.209) (Table 5).

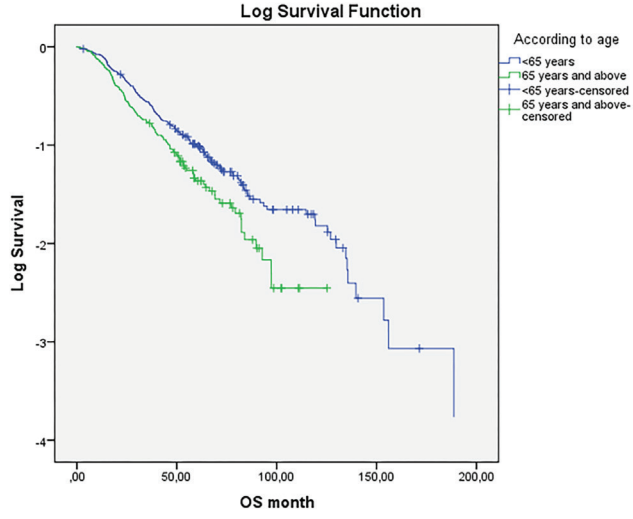


Figure 2. Kaplan-Meier curves for OS according to age group

OS: Overall survival

Table 4. OS according to age group

Age group	n	Events (n)	Censored (%)	Median OS (months)	95% CI	Log-rank p	Breslow p
<65 years	**274**	**218**	**20.4**	**40.7**	**34.6-46.8**	**0.006**	**0.007**
≥65 years	**172**	**142**	**17.4**	**30.2**	**22.1-38.3**		
Overall	**446**	**360**	19.3	38.1	33.6-42.6		

OS: Overall survival, CI: Confidence interval

Table 5. OS by age group, stratified according to CCI								
CCI group	**Age group**	**n**	**Events (n)**	**Censored (%)**	**Median OS (months)**	**95% CI**	**Log-rank p (adjusted)**	**Breslow p**
CCI <7	**<65 years**	**143**	**112**	21.7	50.7	42.6-58.8		
	≥65 years	50	33	34.0	58.5	39.6-77.3		
CCI ≥7	**<65 years**	**131**	**106**	**19.1**	**30.7**	**21.6-39.8**		
	≥65 years	122	109	10.7	24.2	19.5-28.9		
Overall		**446**	**360**	**19.3**	**38.1**	**33.6-42.6**	**0.096**	**0.209**
CCI: Charlson Comorbidity index, OS: Overall survival, CI: Confidence interval								

DISCUSSION

In our single-centre retrospective cohort of 446 patients with mCRC, we found that a higher comorbidity burden, older age, and RAS mutation status were each associated with poorer survival outcomes. Specifically, patients with CCI ≥7 had significantly shorter PFS and OS than those with CCI <7; RAS-mutant patients were more likely to have a high CCI; and older age was strongly associated with worse OS, although the effect on PFS was less pronounced. These findings highlight the importance of integrating host-related factors (age, comorbidities) with tumour molecular features in prognostic modelling of mCRC.

The demographic and clinical characteristics observed in this study are consistent with previously reported epidemiologic patterns of CRC, including a predominance of older patients and males, and a higher frequency of left-sided tumors.[1] The high proportion of patients presenting with stage IV disease could be related our selection criteria , while the liver and lung remained the most common sites of metastasis, as documented in prior studies.[1] The overall RAS mutation frequency (44.8%) aligns with large multicenter analyses reporting mutation rates of approximately 40-50% in mCRC.[15]

Our findings reinforce the prognostic importance of comorbidity in mCRC. Patients with higher CCI experienced nearly twofold reductions in both PFS and OS compared with patients with lower comorbidity burden. This is consistent with prior studies demonstrating that comorbidities negatively influence treatment tolerance, clinical decision-making, and long-term outcomes in CRC.[16-18] Population-based analyses have also shown that comorbidity independently predicts cancer-specific mortality regardless of stage or treatment modality.[19]

The mechanisms underlying this association are likely multifactorial. Chronic comorbidities such as cardiovascular, pulmonary, or metabolic disorders can impair physiological reserve, limit the ability to undergo intensive systemic therapies, and exacerbate treatment-related toxicities. Furthermore, comorbidity-associated inflammation and immune dysregulation may promote tumor progression and reduce therapeutic efficacy. Therefore, comorbidity assessment should be an integral component of baseline evaluation and prognostic modeling in patients with mCRC.

An intriguing observation in our study was the positive association between RAS mutation and a higher comorbidity burden. Patients with RAS-mutant tumors were significantly more likely to have a CCI ≥7 compared with RAS WT counterparts. Similar patterns have been reported in population-based datasets, where RAS-mutant mCRC patients demonstrated increased comorbidity and poorer outcomes.[20] The biological basis for this relationship remains speculative, but may involve shared molecular mechanisms linking metabolic dysregulation, inflammation, and RAS-driven carcinogenesis. These findings suggest that host systemic health may influence or reflect tumor biology, emphasizing the need for integrated clinical-molecular risk stratification.

Age alone was not associated with differences in PFS, implying comparable disease control under treatment in both younger and older patients. However, OS was significantly shorter in older patients, likely reflecting the cumulative effects of frailty, comorbidities, and reduced tolerance to systemic therapy, rather than intrinsic tumor aggressiveness.[3,17] Importantly, after adjustment for comorbidity (CCI), the prognostic effect of age on OS was no longer significant, suggesting that comorbidity burden, rather than chronological age, is the dominant determinant of long-term survival. This finding aligns with prior reports that the adverse effect of aging on CRC outcomes is largely mediated through comorbid conditions and functional decline.[21]

Our results highlight the prognostic value of systematically incorporating comorbidity assessment into the management of mCRC. Relying solely on age or tumor stage may underestimate the heterogeneity in treatment tolerance and outcomes among mCRC patients. Because of its simplicity and reproducibility, CCI remains a practical

and informative tool for prognostic stratification and individualized therapy planning, particularly in older or medically complex patients.

Study Limitations

Our study has several limitations. Its retrospective, single-center design may limit external validity. Moreover, treatment heterogeneity (chemotherapy regimens, anti-EGFR use, or treatment lines) was not analyzed in detail and may partly explain survival differences. Nevertheless, our cohort represents a real-world population, enhancing the generalizability of these findings to clinical practice.

CONCLUSION

In summary, this study demonstrates that a higher comorbidity burden, older age, and RAS mutation status are significant determinants of survival in mCRC. CCI provides clinically relevant prognostic information beyond traditional tumor features. Incorporating comorbidities and molecular data into prognostic models may improve risk stratification and guide individualized management strategies for patients with mCRC.

Ethics

Ethics Committee Approval: Ethics committee approval was obtained from the Non-Interventional Research Ethics Committee of Dokuz Eylül University (decision number: 2021/01-07, date: 04.01.2021).

Informed Consent: Retrospective study.

This study was derived from the medical specialty thesis of Evrim Ataca, completed in 2022.

Footnotes

Authorship Contributions

Concept: E.A., M.K., S.S., Y.B., T.Y., Design: E.A., M.K., S.S., Y.B., T.Y., Data Collection or Processing: E.A., M.K., Analysis or Interpretation: E.A., M.K., Literature Search: E.A., M.K., T.Y., Writing: E.A.

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

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

REFERENCES

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin. 2025;75:10-45.
2. Biller LH, Schrag D. Diagnosis and treatment of metastatic colorectal cancer: a review. JAMA. 2021;325:669-85.
3. McCleary NJ, Zhang S, Ma C, et al. Age and comorbidity association with survival outcomes in metastatic colorectal cancer: CALGB 80405 analysis. J Geriatr Oncol. 2022;13:469-79.
4. Boakye D, Rillmann B, Walter V, Jansen L, Hoffmeister M, Brenner H. Impact of comorbidity and frailty on prognosis in colorectal cancer patients: a systematic review and meta-analysis. Cancer Treat Rev. 2018;64:30-9.
5. Osumi H, Shinozaki E, Suenaga M, et al. RAS mutation is a prognostic biomarker in colorectal cancer patients with metastasectomy. Int J Cancer. 2016;139:803-11.
6. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40:373-83.
7. Morishima T, Matsumoto Y, Koeda N, et al. Impact of comorbidities on survival in gastric, colorectal, and lung cancer patients. J Epidemiol. 2019;29:110-5.
8. Braithwaite D, Moore DH, Satariano WA, et al. Prognostic impact of comorbidity among long-term breast cancer survivors: results from the LACE study. Cancer Epidemiol Biomarkers Prev. 2012;21:1115-25.
9. Zhang X, Wang X, Wang M, et al. Effect of comorbidity assessed by the Charlson Comorbidity index on the length of stay, costs, and mortality among colorectal cancer patients undergoing colorectal surgery. Curr Med Res Opin. 2023;39:187-95.
10. Winder T, Mündlein A, Rhomberg S, et al. Different types of K-Ras mutations are conversely associated with overall survival in patients with colorectal cancer. Oncol Rep. 2009;21:1283-7.
11. Rimbert J, Tachon G, Junca A, Villalva C, Karayan-Tapon L, Tougeron D. Association between clinicopathological characteristics and RAS mutation in colorectal cancer. Mod Pathol. 2018;31:517-26.
12. Wang X, Wu W, Zheng Z, Chi P. Exploring better strategies for RAS mutation-associated EGFR-targeted resistance in colorectal cancer: from the perspective of cancer community ecology. Front Oncol. 2021;11:754220.
13. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. Am J Epidemiol. 2011;173:676-82.
14. Amin MB, Greene FL, Edge SB, et al. The eighth edition AJCC cancer staging manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin. 2017;67:93-9.
15. Kavgaci G, Akiva I, Ozon YH, et al. Evaluation of KRAS and NRAS mutations in metastatic colorectal cancer: an 8-year study of 10754 patients in Turkey. Mol Oncol. 2025;19:2967-77.
16. Rieker RJ, Hammer E, Eisele R, Schmid E, Högel J. The impact of comorbidity on the overall survival and the cause of death in patients after colorectal cancer resection. Langenbecks Arch Surg. 2002;387:72-6.
17. Tominaga T, Nonaka T, Takeshita H, et al. The Charlson Comorbidity index as an ındependent prognostic factor in older colorectal cancer patients. Indian J Surg. 2018;80:54-60.
18. Cuthbert CA, Hemmelgarn BR, Xu Y, Cheung WY. The effect of comorbidities on outcomes in colorectal cancer survivors: a population-based cohort study. J Cancer Surviv. 2018;12:733-43.
19. Ostenfeld EB, Nørgaard M, Thomsen RW, Iversen LH, Jacobsen JB, Søgaard M. Comorbidity and survival of Danish patients with colon and rectal cancer from 2000-2011: a population-based cohort study. Clin Epidemiol. 2013;5:65-74.
20. Ording AG, Öztürk B, Spindler KG, Sørensen HT, McCusker M, Ehrenstein V. KRAS mutation status, comorbidity, and mortality in patients with metastatic colorectal cancer in Denmark. Acta Oncol. 2018;57:1727-9.
21. Wu CC, Hsu TW, Chang CM, Yu CH, Lee CC. Age-adjusted Charlson comorbidity index scores as predictor of survival in colorectal cancer patients who underwent surgical resection and chemoradiation. Medicine (Baltimore). 2015;94:e431.

DOI: 10.4274/forbes.galenos.2025.94940
Forbes J Med 2025;6(3):273-282

Original Article / Orijinal Araştırma

The Role of Diffusion Tensor Magnetic Resonance Imaging in Understanding Neurological and Ocular Outcomes in Preterm Infants with Periventricular Leukomalacia

Periventriküler Lökomalazili Prematüre İnfantlarda Nörolojik ve Oküler Sonuçların Değerlendirilmesinde Difüzyon Tensor Manyetik Rezonans Görüntülemenin Rolü

Yıldız EKEMEN KELEŞ[1], Elif DEMİRKILINÇ BİLER[2], Filiz AFRASHİ[2], Mehmet Cem ÇALLI[3], Mehmet YALAZ[4], Sema TANRIVERDİ[5], Sanem YILMAZ[6], Gül AKTAN[6], Sarenur GÖKBEN[6]

[1]Çiğli Training and Research Hospital, Clinic of Pediatric Infectious Diseases, İzmir, Türkiye
[2]Ege University Faculty of Medicine, Department of Ophthalmology, İzmir, Türkiye
[3]Ege University Faculty of Medicine, Department of Radiology, İzmir, Türkiye
[4]Ege University Faculty of Medicine, Department of Neonatology, İzmir, Türkiye
[5]Manisa Celal Bayar University Faculty of Medicine, Department of Neonatology, Manisa, Türkiye
[6]Ege University Faculty of Medicine, Department of Pediatric Neurology, İzmir, Türkiye

Cite as: Ekemen Keleş Y, Demirkılınç Biler E, Afrashi F, et al. The role of diffusion tensor magnetic resonance imaging in understanding neurological and ocular outcomes in preterm infants with periventricular leukomalacia. Forbes J Med. 2025;6(3):273-282

ABSTRACT

Objective: This study aimed to compare neurological and ophthalmological findings with diffusion tensor magnetic resonance imaging (DTMRI) results in patients with periventricular leukomalacia (PVL).

Methods: This prospective study included 24 premature infants with PVL diagnosed by cranial ultrasonography (USG) or magnetic resonance imaging (MRI). Neurological and comprehensive ophthalmological evaluations were conducted at 3, 6, and 12 months of corrected age. MRI was performed on a 3-Tesla scanner, and DTMRI data were analyzed for fractional anisotropy (FA) and other diffusion parameters. Statistical analyses were performed using SPSS software, with $p<0.05$ considered significant.

Results: Our study included 11 preterm patients with PVL, of whom 72.7% developed cerebral palsy (CP) by 6 months' corrected age. Cranial USG detected PVL in 63.3% of patients, whereas MRI identified PVL of varying severity: mild (54.5%), moderate (18.2%), and severe (27.3%). Neurological examinations revealed increased muscle tone and brisk deep tendon reflexes (DTRs), findings consistent with early indicators of CP. Ophthalmologic assessments indicated normal light fixation and tracking in most patients, though visual evoked potentials (VEPs) revealed abnormalities in 63.6% of patients. In the DTMRI results, right optic tract, FA values were found to be significantly lower in the patient group compared to the control group ($p=0.010$).

Conclusion: Early neurological signs, such as increased muscle tone and brisk DTRs, are predictive of subsequent CP. Despite improvements in visual tracking, VEP abnormalities highlight the need for comprehensive neuro-ophthalmological evaluations in preterm infants. DTMRI provides valuable insights into white matter microstructure and may help predict neurological and ophthalmological outcomes in infants with PVL.

Keywords: Periventricular leukomalacia, premature, cerebral visual impairment, visual evoked potentials, diffusion tensor magnetic resonance imaging

Received/Geliş: 08.11.2025
Accepted/Kabul: 27.11.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Yıldız EKEMEN KELEŞ, MD,
Çiğli Training and Research Hospital, Clinic of Pediatric Infectious Diseases, İzmir, Türkiye
✉ kutupylz@hotmail.com
ORCID: 0000-0002-6122-1726

ÖZ

Amaç: Bu çalışma, periventriküler lökomalazili (PVL) hastalarda nörolojik ve oftalmolojik bulguların difüzyon tensor manyetik rezonans görüntüleme (DTMRG) sonuçlarıyla karşılaştırılmasını amaçlamıştır.

Yöntem: Prospektif olarak yürütülen bu çalışmaya, kranial ultrasonografi (USG) veya manyetik rezonans görüntüleme (MRG) ile PVL tanısı konulan 24 prematüre bebek dahil edilmiştir. Nörolojik ve kapsamlı oftalmolojik değerlendirmeler düzeltilmiş yaşın 3., 6. ve 12. aylarında yapılmıştır. Görüntülemeler 3 Tesla cihazla gerçekleştirilmiş, DTMRG verileri fraksiyonel anizotropi (FA) ve diğer difüzyon parametreleri açısından analiz edilmiştir. İstatistiksel analizler SPSS yazılımı kullanılarak yapılmış, p<0,05 değeri anlamlı kabul edilmiştir.

Bulgular: Çalışmaya dahil edilen 11 prematüre PVL tanılı hastanın %72,7'sinde düzeltilmiş altıncı ayda serebral palsi (SP) gelişmiştir. Kranial USG ile PVL olgularının %63,3'ünde saptanırken, MRG bulgularına göre hastaların %54,5'i hafif, %18,2'si orta, %27,3'ü ağır düzeyde PVL saptanmıştır. Nörolojik muayenelerde artmış kas tonusu ve canlı derin tendon refleksleri (DTR) erken SP göstergeleriyle uyumlu bulunmuştur. Oftalmolojik değerlendirmelerde çoğu olguda ışık fiksasyonu ve takibi normal olmasına rağmen, görsel uyarılmış potansiyellerin (VEPs) %63,6'sında anormallikler saptanmıştır. DTMRG analizinde hasta grubunda sağ optik trakt FA değerleri kontrol grubuna göre anlamlı derecede düşük bulunmuştur (p=0,010).

Sonuç: Erken dönemde saptanan artmış kas tonusu ve canlı DTR gibi nörolojik bulgular ileride gelişebilecek SP için öngörücü olabilir. Oftalmolojik anormal bulgular izlemde düzelmesine rağmen devam eden VEP anormallikleri, prematüre bebeklerde kapsamlı nöro-oftalmolojik değerlendirmenin gerekliliğini vurgulamaktadır. Difüzyon tensor MRG görüntüleme, PVL'ye eşlik eden motor ve duysal bozuklukların patofizyolojisini anlamada değerli bilgiler sunmaktadır.

Anahtar Kelimeler: Periventriküler lökomalazi, prematüre, serebral görme bozukluğu, görsel uyarılmış potansiyeller, difüzyon tensor manyetik rezonans görüntüleme

INTRODUCTION

Periventricular leukomalacia (PVL) is a condition characterized by ischemic injury to the white matter, especially in premature infants. It is the leading cause of cerebral palsy (CP) in preterm infants.[1] Today, the incidence of PVL is increasing in parallel with the survival rates of premature infants.[2] With the development of more sensitive imaging techniques, such as magnetic resonance imaging (MRI), the detection rate of PVL has increased. Despite these advances, not all cases can be detected by conventional radiological examinations. Some cases can be diagnosed at school age, when they cause neurocognitive and behavioral problems.[3] The pathogenesis of PVL remains unclear, though ischemia, hypoperfusion, and hypoxia are considered key initiating factors leading to oligodendrocyte precursor cell damage, axonal injury, and necrosis. Damage occurs in areas reperfused after ischemia. Destruction of oligodendrocyte precursor cells in the periventricular area leads to axonal damage and necrosis.[4]

The consequences of PVL can vary significantly; ocular problems are frequently encountered in affected infants because the periventricular area is adjacent to the optic nerve fibers.[5] The consequences of PVL may range from severe cerebral visual impairment (CVI) accompanied by CP to milder conditions such as strabismus and learning disabilities.[5,6] Visual acuity in patients with CVI due to damage to the periventricular areas or to generalized hypoxic encephalopathy can vary widely. This condition has become the leading non-ocular cause of visual impairment in the developed world.[7] Although visual acuity is often reduced, it can range from no light perception to 20/20 vision, with some individuals exhibiting symptoms solely attributable to CVI.[6]

In the past, early-stage PVL was diagnosed based on repeated neurological evaluations. Today, MRI can still be used to diagnose PVL before clinical findings are fully developed.[8,9] It can provide information about PVL stage and neuromotor status by assessing lateral ventricular volume, white matter volume loss, and myelination status. However, the fact that clinical findings in affected patients do not always correspond to imaging results constitutes a limitation of this technique.[8,9]

Diffusion tensor imaging (DTI) is a recently developed technique that reveals the microstructural integrity of the brain by imaging the motion of water molecules.[10,11] In white matter, the movement of water molecules varies with axonal structure and myelination status.[10] Water molecules move along the axon. Thus, the localization and spread of tracts can be visualized. Objective data from DTI can be obtained as the apparent diffusion coefficient (ADC), fractional anisotropy (FA), axial diffusion [(AD), parallel to white matter fibers], and radial diffusion [(RD), perpendicular to white matter fibers].[12] The FA value varies with axonal myelination, orientation, number, density, and other cellular components.[13] In studies of newborns, increased FA and decreased ADC and RD values are observed with increasing age, myelination, and axonal growth. Increased myelin production reduces water diffusion and ADC values. In addition, increased axon diameter and greater longitudinal movement of water molecules along the axon will increase FA.[12] Decreased FA values and increased ADC values are observed in white matter damage.[12,14] This technique has recently been used, especially in cases of PVL, and has provided more detailed information about the extent of brain damage.[8,11,15,16]

This study aimed to compare neurological and ophthalmological findings with DTI findings in patients

with PVL and to investigate whether DTI findings contribute to these patients' neurological and ophthalmological prognoses.

METHODS

In this prospective study, patients with a diagnosis of PVL who were evaluated at three months of age or younger were enrolled. The study was conducted at Ege University Hospital. Demographic data, gestational age, gender, birth weight, history of birth asphyxia or chorioamnionitis, the need for and duration of mechanical ventilation, history of sepsis, and maternal age were collected from all participants.

Inclusion criteria encompassed the detection of PVL via cranial ultrasonography (USG) and/or cranial MRI; the presence or absence of Grade 1 intraventricular hemorrhage (IVH); and stage 1 premature retinopathy (ROP). Exclusion criteria included Grade 2 or higher IVH, stages 2-3 ROP, optic nerve atrophy, severe refractive errors, significant genetic or cardiac anomalies, neurometabolic disorders, hydrocephalus, and a ventriculoperitoneal shunt.

A total of 24 patients with a corrected age of 3 months or less and PVL detected on cranial USG and/or cranial MRI were included in the study. Five patients were excluded from the study because stages 2-3 ROP was detected. Four patients who did not return for follow-up and two who relocated to other cities were excluded from the study. Two patients with PVL detected on USG were excluded from the study because their cranial MRI and neurological examinations were normal. The study included eleven patients.

The study received ethical approval from the Ege University Clinical Research Ethics Committee (decision number: 12-11.1/9, date: 27.12.2012). Participation was contingent on obtaining informed consent from the patients' parents.

Neurological Evaluation

The same physicians conducted neurological examinations at three and six months of corrected age. Evaluations included assessments of newborn reflexes, cranial nerve function, light and sound responses (opticofacial reflex and acoustic facial reflexes), gross motor function, and psychomotor development.

Ophthalmological Examination

The same physician performed a thorough ophthalmological evaluation at the third, fifth, and sixth months of corrected age. In four patients (patient nos. 1, 3, 4, 5), eye examinations were repeated at the twelfth month. The evaluations included cycloplegic refractive error assessment, oculomotor status (strabismus assessment), anterior segment and fundus examinations, and flash VEP recordings. Flash VEP responses were analyzed for amplitude, latency, and configuration, with interpretation based on our institution's typical values for the same age group. Given that the development of visual fixation and object tracking typically occurs by six months of age, examinations were targeted for this time frame.[17]

Radiological Techniques

Cranial USG was performed on all participants during the neonatal period. Cranial MRI was performed if USG detected PVL or if any suspicious findings appeared before three months of corrected age. Only one patient, a twin whose sibling had a PVL diagnosis, did not undergo an MRI. In addition to early MRI, a second cranial MRI was performed as part of the study protocol at 6 months corrected age, and DTI acquisition, allowing both structural PVL grading and microstructural diffusion assessment at the same time point. All 6-month MRI and DTI scans were acquired with a 3 Tesla MRI scanner (Siemens Magnetom Verio 3 Tesla MRI and 16-channel head coil), 30 minutes post-sedation. The radiological images were assessed by a single neuroradiologist blinded to patient identities, using the Flodmark criteria for grading PVL.[18] DTI data were evaluated on diffusion-weighted images using FSL software and eddy-current correction. Color-coded maps were prepared using twenty-four directions. FA was performed semi-automatically with the tract-based spatial statistics (TBSS) method by the same operator, blinded to the patients. FA values of all vectors of the determined neuronal fibers were evaluated for both hemispheres. Images containing FA values were arranged until they reached a standard space of 1 x 1 x 1 mm^3. The average FA value and its skeleton were created. All images containing the arranged FA value were transferred to the skeleton. The average threshold value was 0.2. In this way, the results were purified from the diversity in the peripheral neuronal fibers in the cases and the partial effect of the gray matter. A voxel-based statistical analysis was conducted to compare FA, AD, RD, and ADC values between the patient and control groups.

The DTI free-hand region-of-interest (ROI) method was used to determine the patients' FA values from the images, and a single radiologist, blinded to clinical information, manually delineated the ROIs for each patient. FA values were measured in the genu and splenium of the corpus callosum, the right and left internal capsules, the right optic tract, and multiple regions of the left optic tract. They were calculated by averaging the values.

Control Group Imaging Protocol

Seven infants were included in the control group. All control patients underwent cranial MRI, diffusion MRI, and DTI at a corrected age of 6 months for clinical indications (including

nystagmus, macrocephaly, microcephaly, epilepsy, or widespread cutaneous hemangiomas); therefore, no healthy infants were imaged solely for research purposes. MRI and DTI acquisitions in both the PVL and control groups were performed on the same scanner (Siemens Magnetom Verio 3T), using identical parameters (b-value =1000 s/mm², 24 diffusion directions, voxel size of 1 × 1 × 1 mm³, standard EPI sequence, identical scan duration) and the same sedation protocol. Imaging was evaluated by the same neuroradiologist, who was blinded to group allocation. Gestational age and corrected age at the time of imaging were matched between groups.

Statistical Analysis

Descriptive statistics for continuous variables were reported as means or medians, depending on the data distribution. Categorical variables were presented as counts and percentages. Depending on sample size and variable distribution, the chi-square test or Fisher's exact test was applied to compare categorical variables. A p value of less than 0.05 was deemed statistically significant. Statistical analyses were conducted using SPSS version 15.

RESULTS

Patient's Demographics

The gestational ages of the 11 patients with PVL ranged from 26 to 34 weeks, with a mean gestational age of 30.9±3.0 weeks; seven were boys (63.6%). The mean birth weight of the patients was 1727±718 g (range, 830-2790 g). When the patients were evaluated according to their gestational weeks, three patients were detected between 26 and 28 weeks, three cases were detected between 28 and 32 weeks, and five cases were detected between 32 and 34 weeks (Table 1).

Cranial USG was performed on each patient during their intensive care stay. PVL was detected on cranial USG in seven patients (63.3%). Cranial MRI was performed on all patients, except one, before their corrected age of 3 months. The only case without cranial MRI (patient no. 4) was the twin of case no. 3, and since case no. 3 had PVL confirmed by MRI, only cranial USG was performed on this patient.

Neurological Examination

Neurological and developmental assessments of the patients were performed at corrected ages of 3 and 6 months (Table 2). All patients exhibited social smiles; only patients 5 and 8 lacked head control. Increased muscle tone was detected in the third month, except for two patients (patient nos. 10 and 11), and brisk deep-tendon reflexes (DTRs) were detected in the lower extremities, except for three patients (patient nos. 1, 2, and 7).

At a corrected age of six months, neurological examinations revealed spastic CP in eight patients (72.7%) (Table 2). In addition, five of the patients with CP (62.5%) had spastic diparesis, while three (37.5%) had spastic tetraparesis. Regarding psychomotor development, eight patients rolled over. However, they were not acquired in two patients (patients nos. 8 and 10) and were partially acquired in one patient (patient no. 3). Supported sitting was acquired by all but five patients by the sixth month.

Radiological Examination

Cranial MRI revealed mild PVL in six patients (54.5%), moderate PVL in two patients (18.1%), and severe PVL in three patients (27.2%) (Table 2, Figure 1). The DTI results

Table 1. Perinatal and postnatal characteristics of the patients

Patient number	Gender	Birth week	Birth weight*	Perinatal asphyxia	Chorioamnionitis	Need for MV	Number of days of ICU stay	History of septicemia
1	Boy	32	830	-	-	+	64	+
2	Girl	33	1850	-	-	+	17	-
3	Boy	33	2140	+	-	+	18	-
4	Boy	33	2240	-	-	-	10	-
5	Boy	33	2500	+	-	+	8	-
6	Boy	26	900	-	-	+	72	+
7	Boy	27	1160	+	-	+	90	+
8	Girl	32	1320	-	-	+	21	+
9	Girl	34	2790	-	-	-	9	-
10	Boy	26	930	-	+	+	38	+
11	Girl	31	2340	-	-	+	42	-

*Gram
MV: Mechanical ventilation, ICU: Intensive care unit

Table 2. Comparison of psychomotor development at three and six months and cranial magnetic resonance imaging results at six months of patients

Patient no	Third month				Sixth month			
	Head control	Social smile	Muscle tone	DTR	Rolling over	Supported sitting	Neurological examination	PVL grade on cranial MRI
1	+	+	Inc (LE)	Normal	+	-	Spastic diparesis with right Achilles contracture	Moderate
2	+	+	Inc (LE)	Normal	+	+	Normal	Mild
3	± (5. mo)	+ (5. mo)	Spastic tetraparesis		±	-	Spastic tetraparesis	Severe
4	+ (5. mo)	+ (5. mo)	Inc	Brisk (LE)	+	+	Spastic diparesis	Mild
5	-	+	Inc	Brisk (LE)	+	-	Spastic diparesis, right extremity was more affected	Moderate
6	+	+	Inc	Brisk (LE)	+	+	Spastik diparesis	Mild
7	+	+	Inc	Normal	+	+	Normal	Mild
8	-	+	Inc	Brisk (right LE)	-	-	Spastic tetraparesis, right extremity was more affected	Severe
9	+	+	Inc	Brisk (LE)	+	+	Spastic diparesis	Mild
10	+	+	Normal	Brisk (LE)	-	-	Spastic tetraparesis, left extremity more affected	Severe
11	+	+	Normal	Brisk (LE)	+	+	Normal	Mild

DTR: Deep tendon reflex, PVL: Periventricular leukomalacia, MRI: Magnetic resonance imaging, Inc: Increase, LE: Lower extremity, mo: Month

of the patients at six months were evaluated using TBSS and freehand ROI methods. The TBSS method found no statistically significant differences between the FA, AD, RD, and ADC values in the whole white matter between the patient and control groups (p>0.05), as shown in Figure 2. The average FA values of the patients, calculated using the free-hand ROI method, are given in Table 3 (shown in Figure 3). Right optic tract FA values were significantly lower in the patient group compared with the control group (p=0.010). Mean right optic tract FA was 0.244±0.98 in patients and 0.42±0.32 in controls. No statistically significant differences were found in FA values of the genu and splenium of the corpus callosum, right internal capsule, left internal capsule, and left optic tracts between the two groups (p>0.05).

Ophthalmologic Examination

Ophthalmologic evaluation was performed in seven cases at the corrected age of three months. One patient had no fixation to light. Only three patients were able to track objects. At the corrected age of 6 months, all patients demonstrated light fixation, light tracking, and object tracking, except patient 3, who showed none of these (Table 4). The optic discs were pale, and the patient had cortical blindness. Flash VEP results were normal in four patients; prolonged latencies were detected in five patients, decreased amplitude in one patient, and wave pattern disorder in four patients.

DISCUSSION

Normal cerebral blood flow is crucial for brain development, and even minor stress during the prenatal and perinatal periods can disrupt this flow, leading to cerebral ischemia and the formation of free oxygen radicals.[19] Consequently, many infants who survive hypoxic-ischemic encephalopathy develop sequelae; with approximately 90% experience spasticity and minor neurological impairments (4). In our study, we found CP in 72.7% of preterm infants. The literature reports CP rates resulting from PVL ranging from 60% to 100%.[20,21] While CP may not be clinically evident in patients with PVL, perceptual and behavioral issues can emerge during the school-age years.[22]

Although a combination of clinical history, standardized neuromotor assessment, and MRI findings can identify, at an early age, infants with CP and those at risk of developing CP, in clinical practice CP is diagnosed more accurately by age 2.[23]

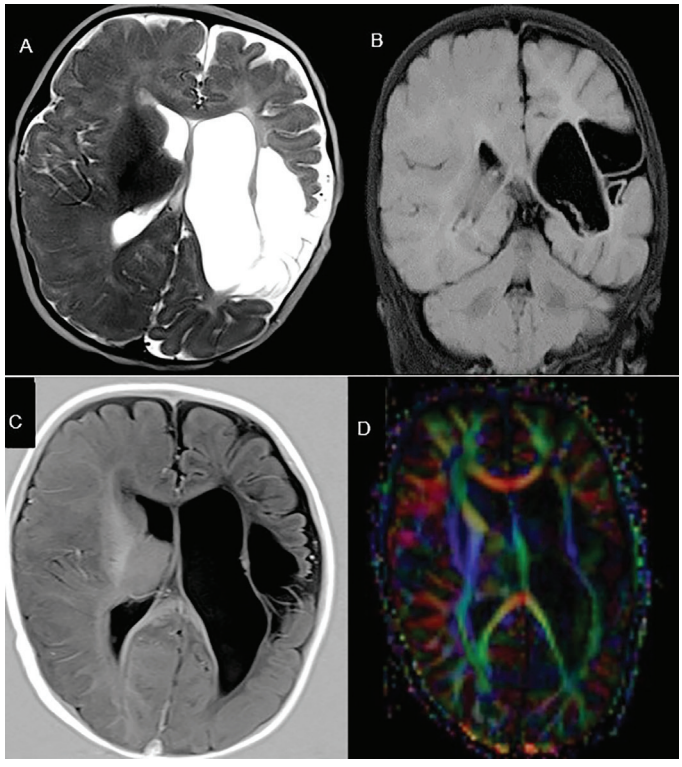


Figure 1. Magnetic resonance imaging of patient number eight. A) Severe periventricular leukomalacia on the left in T2 axial imaging. Retraction and dilatation in the same ventricle. Marked tissue loss in the cortex and white matter. Encephalomalacia area and porencephalic cyst. B) Coronal FLAIR C) T1 axial image. D) Color-code fractional anisotropy map. Red indicates transverse (x-axis), green indicates anterior-posterior (y-axis), and blue indicates superior-inferior (z-axis) direction

In newborns with CP, only brisk DTRs and increased muscle tone are detected on neurological examination.[24] In our study, neurological evaluations of patients at the corrected age of three months revealed increased muscle tone in 81% of patients and brisk DTR in 72%. Additionally, 27% of patients exhibited changes in muscle tone or DTR at three months, which resolved by six months; mild PVL was observed on MRI. At 3 months, 75% of those with CP showed increased muscle tone and brisk DTRs. These clinical findings are significant for early CP development.

Serial cranial USG provides essential information about the preterm infant's overt brain injury and how it changes over time.[25] Cranial USG is safe, relatively inexpensive, and readily available at the bedside for serial imaging of brain injury and its evolution.[15,25] In our study, cranial USG detected PVL in 63.3% of patients during their stay in the intensive care unit. Among the eight infants who developed CP, PVL was detected on cranial USG in five (62.5%). The positive predictive value of USG for detecting CP was 71.4%,

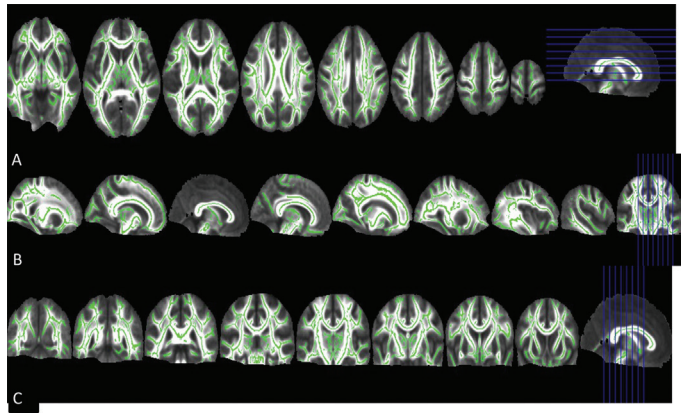


Figure 2. Fractional anisotropy (FA) skeleton. A) Axial B) Sagittal C) Coronal sections. Comparison of mean FA values in the white matter of the patient and control groups using the Tract-Based Spatial Statistics method was performed at each voxel level on the FA skeleton (green color)

while the negative predictive value was only 25%. Previous studies indicate that USG can diagnose severe PVL and cystic changes in only about one-third of patients, making it less effective at detecting diffuse PVL.[16,26] In our study, cranial MRI was performed in all cases except one before their corrected age reached 3 months. PVL was observed on cranial MRI in all patients except one. On neurological examination at three months, DTR vitality was noted in the lower extremities, whereas the examination at six months was normal. Mild PVL was detected on the six-month MRI. In our study, the positive predictive value of cranial MRI for diagnosing PVL was 77.8%, while the negative predictive value was 100%. As expected, MRI is both more sensitive and more specific than USG in identifying PVL. Numerous studies have directly compared MRI and cranial USG in predicting neurodevelopmental outcomes within the same population, including a recent systematic review of preterm infants.[9,26] MRI was found to detect more abnormalities and offer more detailed insights into the severity and extent of brain injuries related to prematurity, especially concerning white matter injury and cerebellar hemorrhage. While MRI demonstrated a high negative predictive value for outcomes, it showed a relatively low positive predictive value for the same outcomes. The prognostic value of MRI varied across studies, primarily influenced by the evaluation tools used and the comprehensiveness of brain abnormality assessments.[9] Based on MRI results at six months, we found mild PVL in 54.5% of cases, moderate PVL in 18.2%, and severe PVL in 27.3%. Neuromotor functions corresponded with MRI findings: normal neurological examinations correlated with mild PVL, whereas all patients with tetraparesis had severe PVL.

Table 3. The mean fractional anisotropy values of the patients with the diffusion tensor imaging results with the free-hand region of interest method

FA values	Genu	Splenium	Right internal capsule	Left internal capsule	Right optic tract	Left optic tract
PVL group (n=11)	0.375±0.93	0.435±0.16	0.378±0.10	0.558±0.87	0.244±0.98	0.306±0.21
Control group (n=7)	0.391±0.79	0.543±0.11	0.388±0.95	0.587±0.69	0.420±0.32	0.318±0.44
p value	0.710	0.140	0.860	0.510	0.010	0.770
FA depreciation	5%	20%	2.6%	5%	42%	3.2%

FA: Fractional anisotropy, PVL: Periventricular leukomalacia

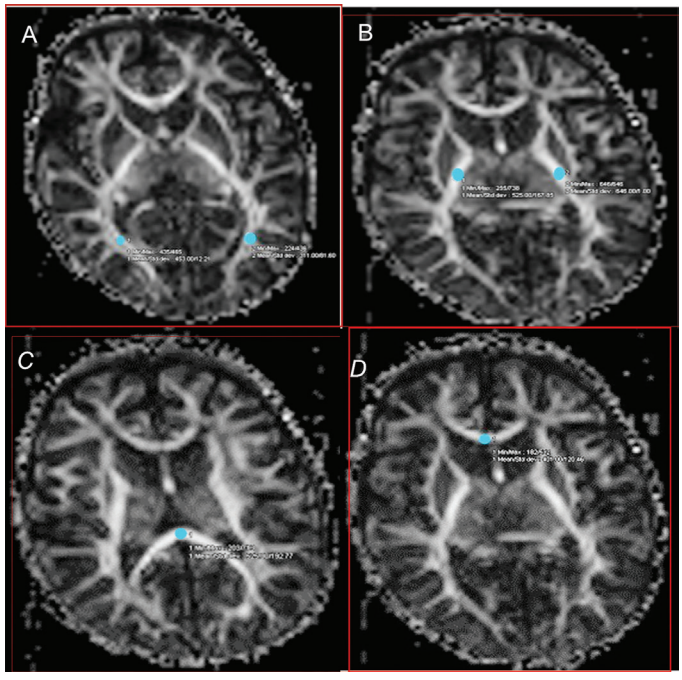


Figure 3. Fractional anisotropy calculation with free-hand region of interest method. The diffusion tensor imaging of the 6-month-old patient was appeared normal. A) Right and left optic tracts. B) Right and left internal capsule. C) Genu section of the corpus callosum. D) Splenium section of the corpus callosum

Inspection, assessment of optokinetic nystagmus, light-response testing, object tracking, and forced preferential gaze can be used in the ophthalmologic evaluation of newborn patients. Eye tracking shows promise for visual assessment in clinical and research settings because it is objective, quantitative, and capable of assessing diverse visual parameters.[27] For ophthalmologic evaluations, we assessed object fixation and tracking. Six of seven patients with a corrected age of three months showed light fixation and tracking, but only three were able to track objects. By six months, all but one case achieved normal light fixation and tracking. This development aligns with the maturation of the optic nerve and visual cortex, which continues until approximately two years of age.[28]

Visual evoked potentials (VEPs) are used in patients unable or unwilling to consistently complete subjective or behavioral tests, as well as in individuals with difficulties in perception and recognition, to assist in localizing visual defects. VEPs can also be used to measure thresholds as a proxy for visual acuity; this technique has been used for many years.[29,30] In our study, VEPs revealed abnormalities in 63.6% of the patients, all of whom developed spastic diplegia or tetraplegia by six months. Six of 10 cases with good vision showed a VEP abnormality (60%). No significant correlation was observed between neuromotor development and VEP responses. This may be due to the small sample size in the normal group. In one study, visual impairment was identified in 357 of 383 (93.2%) patients with CP who had abnormal VEP waveforms.[31] Howes et al.[32] investigated whether pattern reversal VEPs could predict future visual acuity in infants with CVI. VEPs may help predict future visual acuity in young children with CVI.

MRI has revolutionized the early diagnosis of PVL and enabled the assessment of neuromotor status. While traditional imaging techniques have limitations, advancements such as DTI enable the measurement of brain microstructure and connectivity. In our study, we used DTI at a corrected age of 6 months to compare diffusion parameters between 11 PVL patients and 7 controls. FA values were measured multiple locations within the genu and splenium of the corpus callosum, in the right and left internal capsules, and in the right and left optic tracts using the free-hand ROI method; mean values were obtained. FA values in all tracts were lower in the patient group than in the control group. The FA values were lower in the patient group, with statistical significance noted only in the right optic tract. While FA values measured via TBSS did not reveal significant differences, a higher frequency of mild PVL may have influenced the results. In the study conducted by Murakami et al.,[8] DTI was used to evaluate sensorimotor fibers in patients with PVL. The analysis included ROI measurements and TBSS in a cohort of 10 patients with a mean age of 19±9.5 months. The results showed that the mean FA values of the motor tract were

Table 4. Ophthalmological evaluations and visual evoked potential results at corrected age of three and six months of the patients

	Third month			Sixth month				
Patient no	**Light fixation**	**Light tracking**	**Object tracking**	**Light fixation**	**Light tracking**	**Object tracking**	**Additional**	**VEP**
1	+	+	+	+	+	+	Retina diffuses thin	Prolongation of latency
2	+	+	-	+	+	+	None	Normal
3	None			-	-	-	The optic discs are very pale, and exotropia in one eye	The wave pattern is disrupted.
4	None			+	+	+	None	Prolongation of latency
5	-	-	-	+	+	+	Intermittent exotropia	Prolongation of latency breaks the wave pattern.
6	+	+	+	+	+	+	None	Prolongation of latency, the wave pattern is broken
7	+	-/+	-	+	+	+	Intermittent exotropia, optic disc tilt, peripapillary atrophy	Normal
8	None			+	+	+	High astigmatism	Prolongation of latency, the wave pattern is broken
9	+	+	-	+	+	+	None	Amplitude low
10	+	+	±	+	+	+	None	Normal
11	None			+	+	+	Optic disc tilt and slight pallor	Normal

VEP: Visual evoked potential

significantly higher in patients with mild PVL than in patients with severe PVL. Additionally, ROI measurements were less sensitive than tractography-based measurements.[32] In their study, Fan et al.[10] examined 12 patients of PVL, with patients aged between 3 and 10 years and a mean age of 6.5 years. The researchers found a significant reduction in mean FA across several brain structures compared to those in the ipsilateral regions of healthy controls. Specifically, the affected areas included the posterior limb of the internal capsule, arcuate fasciculus, posterior thalamic radiation, corona radiata, cingulum, superior longitudinal fasciculus, and both the splenium and genu of the corpus callosum. Similarly, Madhavan et al.[33] showed that reduced white matter integrity (i.e., greater white matter damage), as reflected by the decreased FA values, in the middle third of the posterior limb of the internal capsule was the most descending; white matter motor pathways converged, which was highly indicative of poor motor function and the diagnosis of CP at 12 months. These advanced imaging methods can now investigate how brain connectivity networks are altered by white matter injury and other influencing factors.

Study Limitations

There are several limitations to our study. Firstly, the application of DTI MRI in young children can be challenging, particularly because sedation is often required.[11] Additionally, the small size of brain structures in this age group, together with a high prevalence of artifacts and rapid myelination, complicates interpretation of results.[11] Diffusion parameters are known to vary significantly during the first two years of life, especially within the first six months, due to ongoing myelination processes, which may influence the comparability of measurements across patients.

Furthermore, as noted by Plaisier et al.,[34] structural brain abnormalities identified on MRI were strongly associated with long-term neurodevelopmental outcomes, suggesting that imaging at a term-equivalent age may yield more stable and interpretable findings. However, early MRI scans hold significant promise, as they can inform early intervention strategies and advance research on preterm brain injuries. Differences in FA values at an early age in patients with PVL have been reported to correlate with motor function loss in later years.[32,33]

In addition to these methodological considerations, our study is limited by its relatively small sample size, short follow-up duration, and incomplete ophthalmic assessments in a subset of patients. These constraints reflect the inherent challenges of conducting longitudinal, multimodal evaluations in premature infants, they but

should be considered when interpreting the generalizability of the findings. Larger, multicenter studies with extended follow-up are needed to elucidate long-term neuro-ophthalmologic and diffusion-based outcomes in this population.

CONCLUSION

In conclusion, DTI is a powerful tool for assessing the microstructural integrity of neural tissues. Our study demonstrates that DTI provides critical insights into the orientation and integrity of white matter tracts; these insights are essential for understanding various neurological conditions in patients with PVL at a corrected age of six months. Moreover, developmental assessments indicated that early neurological signs, such as increased muscle tone and brisk DTRs, predicted subsequent CP outcomes. While ophthalmologic evaluations showed promising improvement in visual tracking, VEP abnormalities were prevalent, underscoring the need for comprehensive assessments beyond conventional methods.

Ethics

Ethics Committee Approval: The study received ethical approval from the Ege University Clinical Research Ethics Committee (decision number: 12-11.1/9, date: 27.12.2012).

Informed Consent: Participation was contingent on obtaining informed consent from the patients' parents.

Acknowledgments

We wish to express our gratitude to all the radiology department technicians at our hospital.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. Zupan V, Gonzalez P, Lacaze-Masmonteil T, et al. Periventricular leukomalacia: risk factors revisited. Dev Med Child Neurol. 1996;38:1061-7.
2. Hayakawa M. Neurological diseases. In: Kusuda S, Nakanishi H, Isayama T, editors. Neonatal intensive care for extremely preterm infants. Amsterdam: Elsevier / Academic Press; 2024. p. 123-45.
3. Choi JY, Rha DW, Park ES. The Effects of the severity of periventricular leukomalacia on the neuropsychological outcomes of preterm children. J Child Neurol. 2016;31:603-12.
4. Kadhim H, Khalifa M, Deltenre P, Casimir G, Sébire G. Molecular mechanisms of cell death in periventricular leukomalacia. Neurology. 2006;67:293-9.
5. Khanna S, Sharma A, Ghasia F, Tychsen L. Prevalence of the infantile strabismus complex in premature children with and without periventricular leukomalacia. Am J Ophthalmol. 2022;240:342-51.
6. Robitaille JM. Long-term visual outcomes in prematurely born children. J Binocul Vis Ocul Motil. 2024;74:1-8.
7. Pehere N, Chougule P, Dutton GN. Cerebral visual impairment in children: causes and associated ophthalmological problems. Indian J Ophthalmol. 2018;66:812-5.
8. Murakami A, Morimoto M, Yamada K, et al. Fiber-tracking techniques can predict the degree of neurologic impairment for periventricular leukomalacia. Pediatrics. 2008;122:500-6.
9. Burkitt K, Kang O, Jyoti R, Mohamed AL, Chaudhari T. Comparison of cranial ultrasound and MRI for detecting BRAIN injury in extremely preterm infants and correlation with neurological outcomes at 1 and 3 years. Eur J Pediatr. 2019;178:1053-61.
10. Fan GG, Yu B, Quan SM, Sun BH, Guo QY. Potential of diffusion tensor MRI in the assessment of periventricular leukomalacia. Clin Radiol. 2006;61:358-64.
11. Guillot M, Sebastianski M, Lemyre B. Comparative performance of head ultrasound and MRI in detecting preterm brain injury and predicting outcomes: a systematic review. Acta Paediatr. 2021;110:1425-32.
12. Gulani V, Sundgren PC. Diffusion tensor magnetic resonance imaging. J Neuro-Ophthalmol. 2006;26:51-60.
13. Yoshida S, Hayakawa K, Yamamoto A, et al. Quantitative diffusion tensor tractography of the motor and sensory tract in children with cerebral palsy. Dev Med Child Neurol. 2010;52:935-40.
14. van Kooij BJ, de Vries LS, Ball G, et al. Neonatal tract-based spatial statistics findings and outcome in preterm infants. AJNR Am J Neuroradiol. 2012;33:188-94.
15. Berman JI, Glass HC, Miller SP, et al. Quantitative fiber tracking analysis of the optic radiation correlated with visual performance in premature newborns. AJNR Am J Neuroradiol. 2009;30:120-4.
16. Dubois J, Dehaene-Lambertz G, Soarès C, Cointepas Y, Le Bihan D, Hertz-Pannier L. Microstructural correlates of infant functional development: example of the visual pathways. J Neurosci. 2008;28:1943-8.
17. Inder TE, de Vries LS, Ferriero DM, et al. Neuroimaging of the preterm brain: review and recommendations. J Pediatr. 2021;237:276-87.e4.
18. Hinojosa-Rodríguez M, Harmony T, Carrillo-Prado C, et al. Clinical neuroimaging in the preterm infant: Diagnosis and prognosis. Neuroimage Clin. 2017;16:355-68.
19. Cioni G, Fazzi B, Coluccini M, Bartalena L, Boldrini A, van Hof-van Duin J. Cerebral visual impairment in preterm infants with periventricular leukomalacia. Pediatr Neurol. 1997;17:331-8.
20. Flodmark O, Lupton B, Li D, et al. MR imaging of periventricular leukomalacia in childhood. AJR Am J Roentgenol. 1989;152:583-90.
21. Van Dyken P, Lacoste B. Impact of metabolic syndrome on neuroinflammation and the blood-brain barrier. Front Neurosci. 2018;12:930.

22. Shang Q, Ma CY, Lv N, et al. Clinical study of cerebral palsy in 408 children with periventricular leukomalacia. Exp Ther Med. 2015;9:1336-44.
23. Song J, Yue Y, Sun H, et al. Clinical characteristics and long-term neurodevelopmental outcomes of leukomalacia in preterm infants and term infants: a cohort study. J Neurodev Disord. 2023;15:24.
24. Jongmans M, Mercuri E, de Vries L, Dubowitz L, Henderson SE. Minor neurological signs and perceptual-motor difficulties in prematurely born children. Arch Dis Child Fetal Neonatal Ed. 1997;76:F9-14.
25. Patel DR, Bovid KM, Rausch R, Ergun-Longmire B, Goetting M, Merrick J. Cerebral palsy in children: A clinical practice review. Curr Probl Pediatr Adolesc Health Care. 2024;54:101673.
26. Straathof EJM, Hamer EG, Hensens KJ, La Bastide-van Gemert S, Heineman KR, Hadders-Algra M. Development of muscle tone impairments in high-risk infants: associations with cerebral palsy and cystic periventricular leukomalacia. Eur J Paediatr Neurol. 2022;37:12-8.
27. Parodi A, Morana G, Severino MS, et al. Low-grade intraventricular hemorrhage: is ultrasound good enough? J Matern Fetal Neonatal Med. 2015;28(Suppl 1):2261-4.
28. Chang MY, Borchert MS. Methods of visual assessment in children with cortical visual impairment. Curr Opin Neurol. 2021;34:89-96.
29. Caffarra S, Joo SJ, Bloom D, Kruper J, Rokem A, Yeatman JD. Development of the visual white matter pathways mediates development of electrophysiological responses in visual cortex. Hum Brain Mapp. 2021;42:5785-97.
30. Viswanath M, Jha R, Gambhirao AD, et al. Comorbidities in children with cerebral palsy: a single-centre cross-sectional hospital-based study from India. BMJ Open. 2023;13:e072365.
31. Regan D. Rapid objective refraction using evoked brain potentials. Invest Ophthalmol. 1973;12:669-79.
32. Howes J, Thompson D, Liasis A, Oluonye N, Dale N, Bowman R. Prognostic value of transient pattern visual evoked potentials in children with cerebral visual impairment. Dev Med Child Neurol. 2022;64:618-24.
33. Madhavan S, Campbell SK, Campise-Luther R, et al. Correlation between fractional anisotropy and motor outcomes in one-year-old infants with periventricular brain injury. J Magn Reson Imaging. 2014;39:949-57.
34. Plaisier A, Govaert P, Lequin MH, Dudink J. Optimal timing of cerebral MRI in preterm infants to predict long-term neurodevelopmental outcome: a systematic review. AJNR Am J Neuroradiol. 2014;35:841-7.

DOI: 10.4274/forbes.galenos.2025.40374
Forbes J Med 2025;6(3):283-292

Original Article / Orijinal Araştırma

Retrospective Evaluation of the Relationship Between Serum Inflammatory Markers and the Efficacy of Methotrexate in the Treatment of Ectopic Pregnancy

Serum Enflamasyon Belirteçleri ile Ektopik Gebelik Tedavisinde Kullanılan Metotreksat'ın Etkinlik İlişkisinin Retrospektif Olarak Değerlendirilmesi

Aslı AKDÖNER[1], Onur YAVUZ[1], Alper MANKAN[2], Elif Aleyna KÜÇÜK[1], Erkan ÇAĞLIYAN[1]

[1]Dokuz Eylül University Faculty of Medicine, Department of Obstetrics and Gynecology, İzmir, Türkiye
[2]Erzurum City Hospital, Clinic of Obstetrics and Gynecology, Erzurum, Türkiye

Cite as: Akdöner A, Yavuz O, Mankan A, Küçük EA, Çağlıyan E. Retrospective evaluation of the relationship between serum inflammatory markers and the efficacy of methotrexate in the treatment of ectopic pregnancy. Forbes J Med. 2025;6(3):283-292

ABSTRACT

Objective: Inflammatory processes play a significant role in the pathogenesis of ectopic pregnancy. Therefore, certain serum inflammatory markers are considered to have potential value in predicting the treatment outcome. The objective of this study is to investigate the relationship between methotrexate (MTX) efficacy in the management of ectopic pregnancy and serum inflammatory markers.

Methods: The study was conducted retrospectively at Dokuz Eylül University Hospital. Women aged 18 years or older who were diagnosed with an ectopic pregnancy and underwent treatment between 2015 and 2025 were included. The study population was initially stratified into two groups according to the success or failure of MTX treatment. Patients with successful treatment were further subdivided into single-dose and double-dose MTX cohorts. The groups were compared with respect to characteristic features and ultrasonographic, laboratory, and surgical findings.

Results: No significant differences were found in systemic inflammation markers between the groups (p=0.96, p=0.58, p=0.47, p=0.80). When groups with successful and unsuccessful MTX treatment were compared with respect to baseline, fourth-day, and seventh-day β-hCG levels and changes between these measurements, a significant difference was found (p<0.001). β-hCG was an effective predictor of treatment success (p=0.001, p=0.001, p=0.006).

Conclusion: No significant correlation was found between serum inflammatory markers and the success of MTX treatment. β HCG levels and their rates of change are thought to be more powerful predictors of MTX treatment success.

Keywords: Ectopic pregnancy, serum inflammatory markers, methotrexate

ÖZ

Amaç: Ektopik gebelik patogenezinde enflamatuar süreçlerin önemli bir etkisi bulunmaktadır. Bu nedenle, bazı serum enflamasyon belirteçlerinin tedavi sürecinin öngörüsünde etkin olabileceği düşünülmektedir. Amacımız, ektopik gebelik tedavisinde kullanılan metotreksatın etkinliği ile serum enflamasyon belirteçlerinin ilişkisini araştırmaktır.

Yöntem: Çalışma, Dokuz Eylül Üniversitesi Hastanesi' nde retrospektif olarak gerçekleştirilmiştir. 2015-2025 yılları arasında, ektopik gebelik tanısı alarak tedavi edilen 18 yaş ve üzeri kadın olgular çalışmaya dahil edilmiştir. Olgular, öncelikle metotreksat tedavisinin başarılı olup olmamasına göre iki gruba ayrılmıştır. Metotreksat tedavisi başarılı olan olgular kendi içerisinde tek doz ve çift doz tedavi uygulanan olgular olarak ayrıca iki gruba ayrılmıştır. Gruplar, karakteristik özellikleri, ultrason, laboratuvar ve cerrahi bulguları açısından karşılaştırılmıştır.

Received/Geliş: 30.09.2025
Accepted/Kabul: 29.11.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Aslı AKDÖNER, MD,
Dokuz Eylül University Faculty of Medicine, Department of Obstetrics and Gynecology, İzmir, Türkiye
✉ drasliakdoner@gmail.com
ORCID: 0000-0002-9269-0859

Bulgular: Çalışmamızın sonucunda, metotreksat tedavisinin başarılı olduğu grupla başarısız olduğu grup arasında, sistemik enflamasyon belirteçleri açısından anlamlı fark saptanmamıştır (p=0,96, p=0,58, p=0,47, p=0,80). Metotreksat tedavisinin başarılı ve başarısız olduğu gruplar, başlangıç, 4. ve 7. gün β HCG düzeyleri ve aralarındaki değişim açısından karşılaştırıldığında anlamlı fark saptanmıştır (p<0,001). β HCG'' nin tedavi başarısını öngörmede etkin bir belirteç olduğu saptanmıştır (p=0,001, p=0,001, p=0,006).

Sonuç: Sonuç olarak, serum enflamasyon belirteçleriyle metotreksat tedavisinin başarısı arasında anlamlı bir ilişki saptanmamıştır. β HCG düzeyi ve β HCG değişim oranlarının metotreksat tedavisinin başarısını öngörmede daha güçlü bir belirteç olduğu düşünülmektedir.

Anahtar Kelimeler: Ektopik gebelik, serum enflamasyon belirteçleri, metotreksat

INTRODUCTION

Ectopic pregnancies account for approximately 1% of all pregnancies and for about 4.7% of patients who present to the emergency department in early pregnancy.[1,2] Despite declining mortality rates from ectopic pregnancy owing to advances in diagnosis and surgical interventions, the financial burden of numerous diagnostic and follow-up procedures remains substantial, and the psychological impact on patients is considerable.[3]

In ectopic pregnancies, implantation most frequently occurs in the ampullary portion of the fallopian tube. Recognized risk factors include a previous ectopic pregnancy, tubal injury resulting from infection or surgical procedures, smoking, and use of assisted reproductive techniques.[4] Certain serum inflammatory mediators and cytokines are implicated in the pathogenesis of ectopic pregnancy by contributing to impaired embryo transport to the uterus and to premature implantation.[5] Given that ectopic pregnancy is a prominent cause of maternal mortality during the first trimester, prompt diagnosis and appropriate management are essential.[6]

There is no single definitive treatment for ectopic pregnancy. Management varies depending on the site of implantation and the clinical status of the patient. Treatment options include both pharmacological and surgical approaches. In addition, given the possibility of spontaneous resorption in a considerable proportion of cases, expectant management with serial β-hCG monitoring may also be applied. As a pharmacological option, single-dose or multidose methotrexate (MTX) regimens are used.[7] The most commonly employed protocol involves monitoring serum β-hCG levels on days 4 and 7 after MTX administration. If β-hCG decreases by more than 15%, weekly β-hCG monitoring is continued, whereas a decrease of less than 15% warrants a second MTX dose.[8] Reports indicate that approximately 20% of patients require an additional dose of MTX.[6,9]

With increasing evidence demonstrating the important role of inflammatory processes in the pathogenesis of ectopic pregnancy, several studies have suggested that certain serum inflammatory markers may also serve as predictors of treatment outcomes.[5,10-13] While some studies have reported that an increased neutrophil-to-lymphocyte ratio (NLR) is associated with a higher risk of ectopic pregnancy rupture and with reduced MTX treatment success, other studies have found no significant association.[11,12,14] This inconsistency in the literature prompted us to investigate whether systemic inflammatory markers could reliably predict MTX treatment outcomes in patients with ectopic pregnancy. Our aim, in light of the current literature, is to evaluate the predictive value of serum inflammatory markers for the effectiveness of MTX treatment, a pharmacological treatment option for ectopic pregnancy, and to protect these patients from the risks of prolonged hospitalization, MTX side effects, and delayed surgical intervention.

METHODS

This research was conducted retrospectively in the Department of Obstetrics and Gynecology. This study was approved by the Dokuz Eylül University Non-interventional Research Ethics Committee (decision number: 2025/14-08, date: 30.04.2025). The research was performed following the ethical guidelines outlined in the 2008 Declaration of Helsinki. Women aged 18 years and older who were admitted to our gynecology inpatient clinic or outpatient department between January 2015 and April 2025 and who were diagnosed with ectopic pregnancy and subsequently treated were included in the study. Women with systemic inflammatory diseases or concomitant malignancies were excluded. Informed consent was collected from every patient before treatment initiation. Data retrieved retrospectively from medical records included: body surface area (BSA), age, body mass index (BMI), gravida, parity, ultrasonographic (USG) findings at admission (presence of gestational sac, gestational sac diameter, and location of ectopic pregnancy), history of pelvic surgery, history of ectopic pregnancy, history of assisted reproductive treatment, gestational age (weeks), pre-treatment and post-MTX β-hCG levels, pre-treatment complete blood count parameters [hemoglobin (Hb), hematocrit, platelet (Plt) count, white blood cell (WBC) count, red cell distribution width (RDW), mean platelet volume (MPV), NLR, monocyte-to-lymphocyte ratio (MLR), platelet-to-lymphocyte ratio (PLR), pan-immun inflamation value (PIV), systemic immune-inflammation index (SII), systemic immune-inflammation response index (SIRI)] and type of surgery performed (salpingectomy, salpingostomy, diagnostic laparoscopy).

The hematological parameters included in the analysis were collected in a standardized manner as part of routine clinical care and were consistently measured at the time of hospital admission and prior to MTX administration for all patients. Serum β-hCG samples were collected on the day of MTX initiation and on days 4 and 7 thereafter. In patients requiring additional MTX doses, serum β-hCG measurements were repeated. USG findings were recorded during the initial gynecological examination on admission. BSA was calculated as: [height (cm) × weight (kg)] / 3600)$^{1/2}$. BMI was calculated as weight (kg) / height (m)^2. NLR was calculated as neutrophil count (×10³/μL) / lymphocyte count (×10³/μL). PLR was calculated as the Plt count (×10³/μL) divided by the lymphocyte count (×10³/μL). MLR was calculated as the ratio of monocyte count (×10³/μL) to lymphocyte count (×10³/μL). SII was calculated as (Plt count × neutrophil count) / lymphocyte count. SIRI was calculated as [neutrophil count (×10³/μL) x monocyte count (×10³/μL)] / lymphocyte count (×10³/μL) and PIV was calculated as [neutrophil count (×10³/μL) x monocyte count (×10³/μL) x Plt count (×10³/μL)] / lymphocyte count (×10³/μL).

Initially, patients were categorized into two groups based on the outcome of MTX therapy. Group I: patients with successful MTX treatment, and Group II: patients with unsuccessful MTX treatment. Patients in Group I were further subdivided into those who received a single dose of MTX and those who required two doses of MTX; these subgroups were then compared. Patients who underwent surgery during follow-up due to hemodynamic instability or rupture of the ectopic pregnancy, and those who required surgery after two MTX doses because of an insufficient decline in β-hCG levels, were classified as treatment failures. Patients with ectopic pregnancies who demonstrated positive fetal cardiac activity at initial presentation and were scheduled for primary surgery; patients who required emergency surgery due to rupture at admission; patients with extra-tubal or extra-ovarian ectopic pregnancies; and patients with missing data were excluded from the study. The groups were compared with respect to characteristic features, clinical, surgical, and laboratory findings. The study flowchart is summarized in Figure 1.

Statistical Analysis

SPSS v. 26 was used to perform statistical analyses. Comparisons of clinical and demographic characteristics between groups were performed using an independent t-test for parametric distributions and a Mann-Whitney U test for non-parametric distributions. The Shapiro-Wilk and Kolmogorov-Smirnov tests were used to assess normality. Numeric variables were expressed as mean and standard deviation for normally distributed data, and as median (minimum-maximum) for skewed distributions. Categorical data were analyzed using chi-square test or the two-sided Fisher's exact test, as appropriate. The odds ratio was calculated where applicable. To determine the

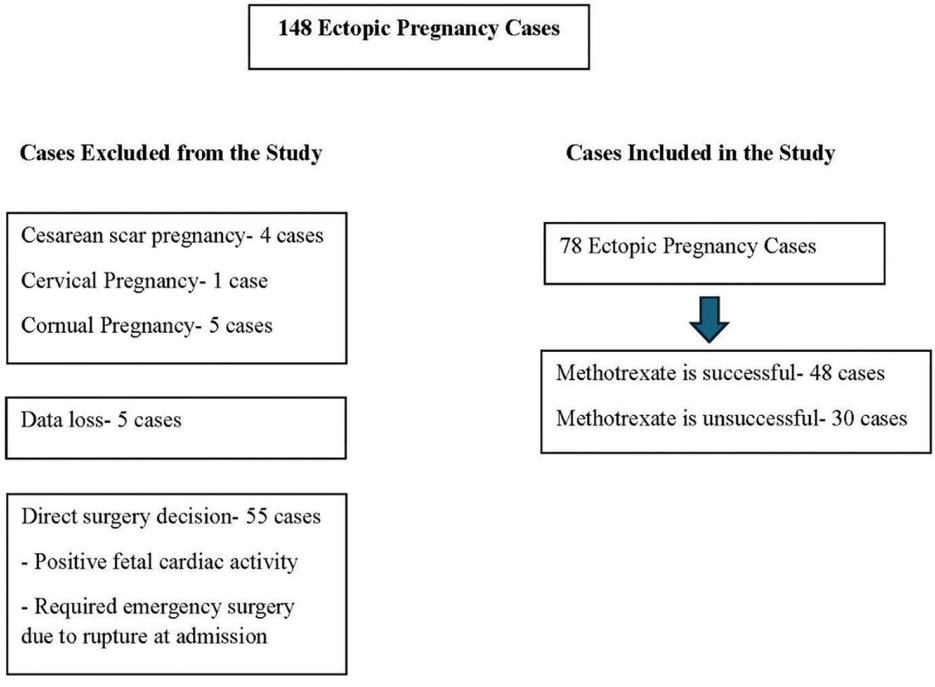


Figure 1. Study flow chart

predictive values of significant variables, logistic regression analyses were performed. In the statistical analysis, the mean values for normally distributed data and the median values for non-normally distributed data are presented in the table. Data with normal and non-normal distributions are indicated in footnotes below the table. p≤0.05 was considered significant.

RESULTS

When the groups with successful and unsuccessful MTX treatment were compared in terms of characteristics, no significant differences were observed in age (31.4 vs. 31.47 years, p=0.95) or in other characteristics. Evaluation of laboratory findings revealed no significant differences among groups in NLR, PLR, MLR, SII, PIV, or SIRI values (p=0.96, p=0.58, p=0.47, p=0.80, p=0.99, p=0.80) (Table 1).

Among patients with successful MTX treatment, a subgroup analysis comparing single-dose and double-dose MTX revealed no statistically significant differences in characteristics. Laboratory evaluation showed that only Hb levels were significantly higher in patients who received a single MTX dose than in those who received two doses (12.45 vs. 11.55 g/dL; p=0.02). No significant differences were found among the groups with respect to NLR, PLR, MLR, SII, PIV, or SIRI (NLR: p=0.24; PLR: p=0.95; MLR: p=0.93; SII: p=0.40; PIV: p=0.717; SIRI: p=0.785) (Table 2).

When patients with unsuccessful versus successful MTX treatment were compared regarding changes in β-hCG levels following MTX administration, the successful treatment group showed a significant decline in β-hCG on days 4 and 7 relative to baseline, whereas the unsuccessful group showed minimal change or an increase. These differences were significant (p<0.001 for all comparisons). Furthermore, a comparison of mean β-hCG levels on the day of MTX initiation, day 4, and day 7 revealed that levels were significantly lower in the MTX-successful group (1469 vs. 3254.5, p<0.001; 1052.4 vs. 5711.5, p<0.001; 619.26 vs. 4062, p<0.001) (Table 3).

A total of 59 patients received a single dose of MTX. Treatment was successful in 34 patients, whereas 25 required surgical intervention due to treatment failure. Evaluation of these patients to assess the predictive role of β-hCG levels for treatment success demonstrated that greater reductions in β-hCG were associated with an increased likelihood of successful treatment; this relationship was significant (p=0.001, p=0.001, p=0.006) (Table 4).

Nineteen patients received a double-dose MTX regimen. Treatment was successful in 14 patients, whereas 5 experienced treatment failure and required surgical intervention. Evaluation of the predictive role of β-hCG levels on the day of the first MTX dose, on day 7, and of the change between these values for the success of the second MTX dose revealed that neither baseline nor day-7 β-hCG levels were predictive of second-dose treatment success. The change in β-hCG levels between baseline and day 7 appeared to be associated with second-dose treatment outcomes. However, these relationships were not statistically significant (p=0.089, p=0.066, and p=0.09) (Table 5).

Table 1. Comparison of groups in terms of characteristic features and laboratory findings

	Group I (n=48)	Group II (n=30)	p
Age	31.4±4.95	31.47±4.33	0.95
BMI (kg/m²)*	23.30 (18-35.20)	22.65 (18.50-33.20)	0.17
BSA (m²)	1.68±0.16	1.65±0.11	0.37
Gravida*	2 (1-6)	2 (1-5)	0.52
Parity*	0.50 (0-3)	1 (0-3)	0.55
Ultrasound findings **Presence of gestational sac** **Gestational sac diameter** **Localization** **Right** **Left**	 22 19.64±5.74 18 30	 17 18.3±6.61 13 17	 0.485 0.35 0.641
History of previous ectopic pregnancy	4	2	1
History of previous pelvic surgery	27	10	0.064
Assisted reproductive therapy **Yes** **No**	 3 45	 2 28	1
Gestational age (weeks)	6.21±1.43	6.50±1.24	0.36

Table 1. Continued

	Group I (n=48)	Group II (n=30)	p
Serum parameters			
Hb	12.0±1.25	12.09±1.12	0.75
Hct	35.98±3.35	36.28±3.31	0.70
WBC	8.40±2.28	8.45±2.96	0.94
Plt	268.05±68.95	267.67±79.15	0.98
MPV	8.72±1	8.52±1.02	0.41
RDW	14.25±1.61	13.80±1.02	0.18
NLR*	2.79 (1.05-12)	2.85 (0.72-15.11)	0.96
PLR*	131.77 (73.64-590)	137.02 (42.63-300)	0.58
MLR	0.3±0.13	0.28±0.09	0.47
SII*	649.32 (156-3540)	713.17 (93.79-3264)	0.80
MCV	83.57±6.71	86.12±5.08	0.08
MCH	27.88±2.65	28.59±1.97	0.21
MCHC	33.31±0.93	33.24±0.60	0.73
RBC	4.31±0.34	4.22±0.42	0.29
Neutrophil	5.56±1.83	5.56±2.42	1.0
Eosinophil*	0.1 (0-0.50)	0.1 (0-1)	0.14
Monocyte*	0.50 (0.10-1.5)	0.50 (0.30-1.70)	0.36
Lymphocyte	2.10±0.73	2.16±1.39	0.81
PCT	0.23±0.05	0.22±0.06	0.56
PIV*	376.2 (15.6-1705.6)	380.1 (28.1-979.2)	0.99
SIRI*	1.4 (0.31-5.59)	1.5 (0.35-4.5)	0.80

BMI: Body mass index (kg/cm^2)

BSA: Body surface area (m^2)

*: Data with non-normal distribution

MCV: Mean corpuscular volume, MCH: Mean corpuscular hemoglobin, MCHC: Mean corpuscular hemoglobin concentration, RBC: Red blood cell, PCT: Plateletcrit, Hb: Hemoglobin, Plt: Platelet, MPV: Mean platelet volume, RDW: Red cell distribution width, NLR: Neutrophil-to-lymphocyte ratio, MLR: Monocyte-to-lymphocyte ratio, SII: Systemic immune-inflammation index, PIV: Pan-immun inflamation value, SIRI: Systemic immune-inflammation response index

Table 2. Comparison of single- and double-dose methotrexate regimens in the successful treatment group in terms of characteristics features and laboratory findings

	Single-dose MTX (n=34)	Double-dose MTX (n=14)	p
Age	32±5.20	29.93±4.08	0.19
BMI (kg/m^2)	25.18±4.87	24.08±3.6	0.45
BSA (m^2)	1.68±0.17	1.68±0.12	0.91
Gravida*	2 (1-6)	2 (1-4)	0.14
Parity*	0.50 (0-3)	0.50 (0-3)	0.9
Ultrasound findings **Presence of gestational sac** **Gestational sac diameter** **Localization** **Right** **Left**	 14 18.96±5.26 10 24	 8 21.29±6.68 8 6	 0.35 0.20 0.103
History of previous ectopic pregnancy	3	1	1
History of previous pelvic surgery	20	7	0.75
Assisted reproductive therapy **Yes** **No**	 2 32	 1 13	1

Table 2. Continued			
	Single-dose MTX (n=34)	**Double-dose MTX (n=14)**	**p**
Gestational age (weeks)	6.41±1.26	5.71±1.73	0.13
Serum parameters			
Hb*	12.45 (7-13.9)	11.55 (10.1-13.3)	0.02**
Hct	36.55±3.46	34.58±2.67	0.063
WBC	8.52±2.42	8.12±1.96	0.59
Plt	269.76±78.48	263.86±39.07	0.73
MPV	8.67±1.15	8.84±0.53	0.60
RDW	14.12±1.62	14.55±1.63	0.41
NLR*	2.90 (1.05-12)	2.37 (1.23-7.36)	0.24
PLR*	130.19 (73.64-590)	133.55 (79.67-278.90)	0.95
MLR	0.3±0.15	0.3±0.10	0.93
SII*	660.92 (156-3540)	554.54 (294.77-1957)	0.40
MCV	84.3±6.16	81.78±7.86	0.24
MCH	28.14±2.41	27.21±3.16	0.27
MCHC	33.34±0.94	33.22±0.94	0.69
RBC	4.34±0.28	4.26±0.46	0.50
Neutrophil	5.65±1.75	5.32±2.06	0.57
Eosinophil*	0.1 (0-0.50)	0.1 (0-0.20)	0.43
Monocyte*	0.50 (0.10-1.50)	0.6 (0.4-0.8)	0.51
Lymphocyte	2.11±0.80	2.08±0.53	0.88
PCT*	0.23 (0.04-0.35)	0.22 (0.18-0.30)	0.89
PIV*	392.6 (15.6-1705.6)	311 (160-978.5)	0.717
SIRI*	1.48 (0.31-5.6)	1.35 (0.55-3.68)	0.785

MTX: Methotrexate

BMI: Body mass index (kg/cm^2)

BSA:Body surface area (m^2)

*: Data with non-normal distribution

**: p<0.05

MCV: Mean corpuscular volume, MCH: Mean corpuscular hemoglobin, MCHC: Mean corpuscular hemoglobin concentration, RBC: Red blood cell, PCT: Plateletcrit, Hb: Hemoglobin, Plt: Platelet, MPV: Mean platelet volume, RDW: Red cell distribution width, NLR: Neutrophil-to-lymphocyte ratio, MLR: Monocyte-to-lymphocyte ratio, SII: Systemic immune-inflammation index, PIV: Pan-immun inflamation value, SIRI: Systemic immune-inflammation response index

Table 3. Comparison of changes in β hCG levels following the first dose of methotrexate between the groups			
	MTX successful group (n=48)	**MTX unsuccessful group (n=30)**	**p**
β hCG levels on the MTX day	1469 (36.71-17.641)	3254.50 (478.77-8060)	<0.001*
β hCG levels on 4th day after MTX	1052.40 (27.91-13586)	5711.5 (527-14819)	<0.001*
β hCG levels on 7th day after MTX	619.26 (3.58-7875)	4062 (519-14454)	<0.001*
β hCG change between MTX day and 4th day	-12.36% (-85.22-66%)	45.01% (-6.1-501.51%)	<0.001*
β hCG change between MTX day and 7th Day	-54.5% (-97.29-59.85%)	42% (-14.68-298.33%)	<0.001*
β hCG change between 4th day and 7th day	-38.19% (-88.93-38.27%)	0.04% (-31.42-164.85%)	<0.001*

MTX: Methotrexate

*: p<0.05

Table 4. Predictive role of β hCG levels in treatment success among cases receiving single-dose methotrexate			
	Odds ratio	**95% confidence interval**	**p**
β hCG change between MTX day and 4th day	0.87	0.80-0.95	0.001*
β hCG change between MTX day and 7th Day	0.96	0.94-0.98	0.001*
β hCG change between 4th day and 7th day	0.85	0.76-0.96	0.006*
MTX: Methotrexate *: p<0.05			

Table 5. Predictive value of β hCG changes between methotrexate day and day 7 for second-dose methotrexate treatment success			
	Odds ratio	**95% confidence interval**	**p**
MTX day β hCG levels	1	1-1.001	0.089
β hCG level on the 7th day after MTX	1.001	1-1.001	0.066
β hCG change between MTX day and 7th Day	0.964	0.923-1.006	0.093
MTX: Methotrexate			

DISCUSSION

Ectopic pregnancy is one of the common causes of maternal morbidity during the first trimester. With the increasing use of assisted reproductive treatments and the rise in cesarean deliveries, it is considered a frequent cause of maternal mortality.[4,15] Delays in the diagnosis and treatment of ectopic pregnancy may result in damage to the fallopian tubes that is sufficient to adversely affect future fertility, as well as in hemorrhage and mortality associated with rupture.[16] MTX is a safe option for the treatment of ectopic pregnancy; however, in some cases, it may result in tubal rupture. Factors such as gestational age, size of the ectopic pregnancy, and pre-treatment β-hCG levels can influence the success of MTX therapy.[17-20] In current practice, β-hCG remains the only biomarker used in ectopic pregnancy, and baseline β-hCG levels are considered the best prognostic indicator of follow-up response and of the success of single-dose MTX treatment.[21]

In the recent study, the predictive value of serum inflammatory markers for the success of MTX treatment in ectopic pregnancy was evaluated; no significant relationship was found between these markers and treatment outcome. Dereli et al.[22] reported that when comparing ectopic pregnancy patients treated with MTX versus surgery, lower NLR and SII levels were observed in the group with successful MTX treatment. Furthermore, patients with prior abdominal surgery or repeated cesarean sections were found to have higher rates of MTX treatment failure, which have been attributed to intra-abdominal adhesions, increased scar tissue, and repeated transverse uterine incisions, all of which lead to compromised adnexal vascularization and reduced MTX delivery to the target tissue. However, in our study, we did not find a relationship between prior abdominal surgery and MTX treatment success (p=0.064). Therefore, our findings are not consistent with the existing literature in this regard. This discrepancy may be attributable to the relatively small sample size of our study, which may not have provided sufficient statistical power to detect this association. In another study evaluating patients treated with MTX for ectopic pregnancy, the decline in NLR from treatment initiation to day 4 was assessed; patients with a decrease of more than 23% in NLR had a greater likelihood of successful MTX treatment.[12] Another study reported higher baseline NLR, MPV, and RDW values in the MTX-treated group, with the underlying mechanism attributed to increased inflammation.[13] A study investigating the association between NLR and PLR and ectopic pregnancy rupture found that ruptured cases exhibited higher NLR and PLR, and that NLR and PLR positively correlated with β-hCG levels and tubal lumen diameter.[14] In our study, no significant association was found between USG findings and treatment modality. Another study indicated that elevated levels of NLR and PDW could serve as markers for tubal rupture, and the association between high NLR and MTX treatment failure was hypothesized to result from neutrophil-mediated suppression of cytotoxic T cells via cytokines and chemokines during inflammatory processes. Accordingly, increased NLR is an indirect measure of the host immune response, and in that study, NLR was higher in the ruptured group.[23] Reis et al.[24] also reported higher WBC levels in the non-ruptured ectopic pregnancy group, suggesting that WBC may play a role in suppressing the maternal immune response against the fetus and in limiting trophoblast invasion. In our study, however, no relationship was observed between MTX treatment success and WBC levels.

Several studies have reported that systemic inflammatory markers may be associated with embryo implantation and the course of ectopic pregnancies.[25-27] The rationale for considering inflammatory processes in these cases is that fetal growth within the fallopian tube can irritate the tube and surrounding structures, such as the peritoneum, leading to inflammation. This process may trigger the release of inflammatory cytokines, and assessing their levels could serve as a useful diagnostic tool for tubal ectopic pregnancy.[27] Supporting this notion, Rajendiran et al.[27] reported increased interleukin-6 levels and decreased interleukin-8 levels in patients with tubal ectopic pregnancy. NLR serves as a strong indicator of systemic immune inflammation and has been identified as an independent predictor of mortality in cardiovascular disease, and as a prognostic marker in patients with primary central nervous system lymphoma undergoing high-dose MTX-based therapy.[28-31] The mechanism linking NLR and MTX treatment is not fully understood, but is thought to be related to the immune response. An elevation in NLR serves as an indirect indicator of immune activity. Neutrophils, as components of the innate immune system, can suppress the maternal immune response toward the fetus and limit trophoblast invasion.[12,23,24,32]

In our study, to add to the existing literature, we evaluated PIV and SIRI -recently highlighted as promising markers for certain cancer types and chronic diseases- with respect to their potential impact on the prognosis of MTX treatment. Previous studies have suggested that PIV may serve as a non-invasive marker for prognosis and remission assessment in some malignancies.[33] Similarly, PIV and SIRI have been reported to be important inflammatory indicators in conditions such as preeclampsia, rheumatoid arthritis, and psoriasis.[34-36] Since these markers are derived from serum measurements and reflect chronic systemic inflammation, it was hypothesized that they might serve as effective predictors of prognosis in ectopic pregnancy. For this purpose, PIV and SIRI values were compared between patients with successful and unsuccessful MTX treatment outcomes. However, no significant associations were identified between MTX treatment success and either PIV or SIRI (p=0.99 and p=0.80, respectively).

Reviewing the literature, Soykan Sert and Bertizlioğlu,[11] similarly to our study, reported no significant relationship between systemic inflammatory markers and ectopic pregnancy treatment outcomes, highlighting β-hCG as a more reliable marker. Tubal ectopic pregnancy can lead to elevated serum inflammatory markers. Soykan Sert and Bertizlioğlu[11] suggested that in cases of ruptured ectopic pregnancy, higher levels of these markers may obscure the relationship with MTX treatment outcomes. MTX exerts an antimetabolic effect, inhibiting cell proliferation; its efficacy is greater at lower β-hCG levels. Soykan Sert and Bertizlioğlu[11] proposed that this could overshadow the association with serum inflammatory markers, emphasizing β-hCG as a stronger prognostic indicator. The findings of our study are consistent with the observations reported by Soykan Sert and Bertizlioğlu.[11]

In our study, evaluation of changes in β-hCG levels after the first MTX dose revealed a significantly greater decrease in the MTX-successful group than in the unsuccessful group. Furthermore, in patients treated with single-dose MTX, β-hCG was a strong predictor of treatment outcome. Consistent with our findings, previous studies have reported that changes in β-hCG levels on days 1 and 4 significantly influence the success of MTX therapy.[11,37,38] Another study demonstrated that both lower baseline β-hCG levels and the rate of decline during the first four days serve as predictors of treatment success.[39] In a study evaluating ruptured and resolved ectopic pregnancies after MTX treatment, cases that ruptured despite therapy were more frequently associated with vaginal spotting and smaller changes in β-hCG levels.[40] In addition to previous studies, our study specifically investigated the predictive role of baseline and day-7 β-hCG levels, as well as the change between these two values, for the success of second-dose MTX treatment. While baseline and day 7 β-hCG levels were not associated with second-dose treatment outcomes, the change in β-hCG levels between baseline and day 7 appeared to reduce the risk of treatment failure. However, this relationship was not statistically significant, which may be attributable to the small sample size. This small sample size (19 cases) represents a major limitation of our study, potentially restricting the ability to detect clinically meaningful differences. Therefore, studies with larger sample sizes may better clarify the nature of this association.

In our study, a comparison of baseline Hb levels between patients with successful and unsuccessful MTX treatment revealed no significant difference. However, within the successful MTX group, a comparison of single- vs. double-dose treatment showed that patients requiring a second MTX dose had significantly lower Hb levels than those treated with a single dose (p=0.02). Nonetheless, this finding does not have a clinical implications.

CONCLUSION

The literature presents conflicting evidence regarding the relationship between serum inflammatory markers and MTX treatment success in ectopic pregnancy. In our study, no significant relationship was found between serum inflammatory markers and MTX treatment outcomes. In our study, as an original contribution to the literature, chronic inflammation markers, including PIV and SIRI, were

evaluated. However, no significant association was found between these markers and MTX treatment outcomes. Based on current literature, β-hCG levels and the rate of β-hCG change are considered the most reliable indicators of MTX treatment success in ectopic pregnancy, a conclusion supported by our findings. As previously noted, one of the major limitations of our study is the relatively small sample size. Therefore, we believe that prospective randomized studies with larger cohorts are needed to better elucidate the relationship between these markers and treatment outcomes.

Ethics

Ethics Committee Approval: This study was approved by the Dokuz Eylül University Non-interventional Research Ethics Committee (decision number: 2025/14-08, date: 30.04.2025).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

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

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

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

REFERENCES

1. O'Herlihy C, Centre for Maternal and Child Enquiries (CMACE). Deaths in early pregnancy. Saving mother's lives: reviewing maternal deaths to make motherhood safer: 2006-2008. The Eighth Report of the Confidential Enquiries into Maternal Deaths in the United Kingdom. BJOG. 2011;118: 81-4.
2. Dooley WM, Chaggar P, De Braud LV, Bottomley C, Jauniaux E, Jurkovic D. Effect of morphological type of extrauterine ectopic pregnancy on accuracy of preoperative ultrasound diagnosis. Ultrasound Obstet Gynecol. 2019;54:538-44.
3. Solangon SA, Van Wely M, Van Mello N, Mol BW, Ross JA, Jurkovic D. Methotrexate vs expectant management for treatment of tubal ectopic pregnancy: an individual participant data meta-analysis. Acta Obstet Gynecol Scand. 2023;102:1159-75.
4. Bouyer J, Coste J, Shojaei T, et al. Risk factors for ectopic pregnancy: a comprehensive analysis based on a large case-control, population-based study in France. Am J Epidemiol. 2003;157:185-94.
5. Attar E. Endocrinology of ectopic pregnancy. Obstet Gynecol Clin North Am. 2004;31:779-94.
6. Marion LL, Meeks GR. Ectopic pregnancy: history, incidence, epidemiology, and risk factors. Clin Obstet Gynecol. 2012;55:376-86.
7. Alkatout I, Honemeyer U, Strauss A, et al. Clinical diagnosis and treatment of ectopic pregnancy. Obstet Gynecol Surv. 2013;68:571-81.
8. Stovall TG, Ling FW, Gray LA. Single-dose methotrexate for treatment of ectopic pregnancy. Obstet Gynecol. 1991;77:754-7.
9. Lipscomb GH, Bran D, McCord ML, Portera JC, Ling FW. Analysis of three hundred fifteen ectopic pregnancies treated with single-dose methotrexate. Am J Obstet Gynecol. 1998;178:1354-8.
10. Lyons RA, Saridogan E, Djahanbakhch O. The reproductive significance of human Fallopian tube cilia. Hum Reprod Update. 2006;12:363-72.
11. Soykan Sert Z, Bertizlioğlu M. The role of inflammatory markers and βhCG levels in predicting the success of single-dose methotrexate treatment in tubal ectopic pregnancy. Int J Gynaecol Obstet. 2025;169:639-44.
12. Abuduxukuer R, Chen X, Ni J, Li S, Lu W. Day 4 and day 0 neutrophil-to-lymphocyte ratios as predictors of treatment failure with single-dose methotrexate for ectopic pregnancies. Int J Gynaecol Obstet. 2024;165:131-7.
13. Akkaya H, Uysal G. Can hematologic parameters predict treatment of ectopic pregnancy? Pak J Med Sci. 2017;33:937-42.
14. Kan Ö, Gemici A, Alkilic A, et al. The effect of preoperative neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio on predicting rupture risk in tubal ectopic pregnancies. Gynecol Obstet Invest. 2019;84:378-82.
15. Stulberg DB, Cain LR, Dahlquist I, Lauderdale DS. Ectopic pregnancy rates in the Medicaid population. Am J Obstet Gynecol. 2013;208:274.e1-7.
16. Barnhart KT. Clinical practice. Ectopic pregnancy. N Engl J Med. 2009 ;361:379-87.
17. Dudley PS, Heard MJ, Sangi-Haghpeykar H, Carson SA, Buster JE. Characterizing ectopic pregnancies that rupture despite treatment with methotrexate. Fertil Steril. 2004;82:1374-8.
18. Nowak-Markwitz E, Michalak M, Olejnik M, Spaczynski M. Cutoff value of human chorionic gonadotropin in relation to the number of methotrexate cycles in the successful treatment of ectopic pregnancy. Fertil Steril. 2009;92:1203-7.
19. Gnisci A, Stefani L, Bottin P, Ohannessian A, Gamerre M, Agostini A. Predictive value of hemoperitoneum for outcome of methotrexate treatment in ectopic pregnancy: an observational comparative study. Ultrasound Obstet Gynecol. 2014;43:698-701.
20. Mirbolouk F, Yousefnezhad A, Ghanbari A. Predicting factors of medical treatment success with single dose methotrexate in tubal ectopic pregnancy: a retrospective study. Iran J Reprod Med. 2015;13:351-4.
21. Menon S, Colins J, Barnhart KT. Establishing a human chorionic gonadotropin cutoff to guide methotrexate treatment of ectopic pregnancy: a systematic review. Fertil Steril. 2007;87:481-4.
22. Dereli ML, Savran Üçok B, Özkan S, et al. The importance of blood-count-derived inflammatory markers in predicting methotrexate success in patients with tubal ectopic pregnancy. Int J Gynaecol Obstet. 2024;167:789-96.
23. King A, Loke YW. On the nature and function of human uterine granular lymphocytes. Immunol Today. 1991;12:432-5.
24. Reis YA, Akay A, Diktaş EG, et al. Prediction of Rupture by Complete Blood Count in Tubal Ectopic Pregnancies Treated with a Single-Dose Methotrexate Protocol. Rev Bras Ginecol Obstet. 2023;45:e503-10.

25. Garlanda C, Maina V, Martinez de la Torre Y, Nebuloni M, Locati M. Inflammatory reaction and implantation: the new entries PTX3 and D6. Placenta. 2008;29 Suppl B:129-34.
26. Mukaida N, Matsumoto T, Yokoi K, Harada A, Matsushima K. Inhibition of neutrophil-mediated acute inflammation injury by an antibody against interleukin-8 (IL-8). Inflamm Res. 1998;47 Suppl 3:S151-7.
27. Rajendiran S, Senthil Kumar GP, Nimesh A, Dhiman P, Shivaraman K, Soundararaghavan S. Diagnostic significance of IL-6 and IL-8 in tubal ectopic pregnancy. J Obstet Gynaecol. 2016;36:909-11.
28. Ham SY, Yoon HJ, Nam SB, Yun BH, Eum D, Shin CS. Prognostic value of neutrophil/lymphocyte ratio and mean platelet volume/platelet ratio for 1-year mortality in critically ill patients. Sci Rep. 2020;10:21513.
29. Luo Q, Yang C, Fu C, Wu W, Wei Y, Zou L. Prognostic role of blood markers in primary central nervous system lymphoma patients treated with high-dose methotrexate-based therapy. Front Oncol. 2021;11:639644.
30. Tamhane UU, Aneja S, Montgomery D, Rogers EK, Eagle KA, Gurm HS. Association between admission neutrophil to lymphocyte ratio and outcomes in patients with acute coronary syndrome. Am J Cardiol. 2008;102:653-7.
31. Wang Q, Ma J, Jiang Z, Ming L. Prognostic value of neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio in acute pulmonary embolism: a systematic review and meta-analysis. Int Angiol. 2018;37:4-11.
32. Faria SS, Fernandes PC Jr, Silva MJ, et al. The neutrophil-to-lymphocyte ratio: a narrative review. Ecancermedicalscience. 2016;10:702.
33. Guven DC, Sahin TK, Erul E, Kilickap S, Gambichler T, Aksoy S. The association between the pan-ımmune-ınflammation value and cancer prognosis: a systematic review and meta-analysis. Cancers (Basel). 2022;14:2675.
34. Tutan D, Doğan AG. Pan-immune-inflammation index as a biomarker for rheumatoid arthritis progression and diagnosis. Cureus. 2023;15:e46609.
35. Özkan S, Dereli ML, Firatligil FB, et al. Role of systemic immune-inflammation index, systemic ınflammation response index, and pan-immune inflammation value in the prediction of preeclampsia: a retrospective cohort study. Am J Reprod Immunol. 2024;92:e70029.
36. Basar Kilic S, Erdal H. Pan-immune inflammation value and systemic inflammatory index as a measure of systemic inflammation in patients with psoriasis: a retrospective study. Medicine (Baltimore). 2025;104:e41715.
37. Mackenzie SC, Moakes CA, Doust AM, et al. Early (Days 1-4) post-treatment serum hCG level changes predict single-dose methotrexate treatment success in tubal ectopic pregnancy. Hum Reprod. 2023;38:1261-7.
38. Girija S, Manjunath AP, Salahudin A, et al. Role of day 4 HCG as an early predictor of success after methotrexate therapy for ectopic pregnancies. Eur J Obstet Gynecol Reprod Biol. 2017;215:230-3.
39. Zhang J, Zhang Y, Gan L, Liu XY, Du SP. Predictors and clinical features of methotrexate (MTX) therapy for ectopic pregnancy. BMC Pregnancy Childbirth. 2020;20:654.
40. Hutchinson AP, Pereira N, Chung ER, et al. Risk factors and human chorionic gonadotropin trends in patients with ruptured tubal ectopic pregnancies despite methotrexate treatment. Gynecol Endocrinol. 2019;35:49-52.

DOI: 10.4274/forbes.galenos.2025.85579
Forbes J Med 2025;6(3):293-298

Orijinal Araştırma / Original Article

Konjenital Diyafragma Hernisinde Mortalite Risk Faktörleri: Tek Merkez Deneyimi

Mortality Risk Factors in Congenital Diaphragmatic Hernia: A Single Centre Experience

Duygu TUNÇEL[1], Leyla ŞERO[2], Suat ÇAL[3], Nilüfer OKUR[1]

[1]Sağlık Bilimleri Üniversitesi, Diyarbakır Gazi Yaşargil Eğitim ve Araştırma Hastanesi, Çocuk Sağlığı ve Hastalıkları Kliniği, Neonatoloji, Diyarbakır, Türkiye
[2]Sağlık Bilimleri Üniversitesi, Diyarbakır Gazi Yaşargil Eğitim ve Araştırma Hastanesi, Çocuk Sağlığı ve Hastalıkları Kliniği, Diyarbakır, Türkiye
[3]Sağlık Bilimleri Üniversitesi, Diyarbakır Gazi Yaşargil Eğitim ve Araştırma Hastanesi, Çocuk Cerrahisi Kliniği, Diyarbakır, Türkiye

Atıf: Tunçel D, Şero L, Çal S, Okur N. Mortality risk factors in congenital diaphragmatic hernia: a single centre experience. Forbes J Med. 2025;6(3):293-298

ÖZ

Amaç: Konjenital diyafragma hernisi (KDH) diyaframın gelişimsel kusuruna bağlı olarak batın içi organların toraksa yerleşmesine neden olan mortalitesi yüksek doğumsal bir defekttir. Sağ taraflı herni, erkek cinsiyet, karaciğer ve mide herniasyonu ve pulmoner hipoplazinin derecesi mortalite ile ilişkilidir. Çalışmamızda diyafram hernilerinde mortalite ile ilgili faktörleri değerlendirmeyi amaçladık.

Yöntem: Aralık 2021-Mart 2025 tarihleri arasında KDH tanısı ile izlenen hastaları retrospektif olarak inceledik. Hastaların demografik özellikleri ile postnatal bulgularını karşılaştırarak mortalite üzerine etkili faktörleri inceledik.

Bulgular: Toplam 30 KDH değerlendirildi. Hastaların 21'i opere edildi, 9 hasta kardiyovaskuler stabilite sağlanamadığı için ilk 24 saat içerisinde eksitus oldu. Hastaların %70'i (n=21) yaşamını kaybederken, %30'u (n=9) sağ kalmıştır. Yenidoğan yoğunbakım yatışında alınan ilk kan gazındaki laktat değerinin 2.3 mmol/L [eğri altındaki alan (EAA): 0,79, p<0,01], ve pCO_2 (EAA: 0,98, p<0,01) değerinin 60 mmHg üzerinde olması mortalite ile ilişkili bulunmuştur.

Sonuç: KDH mortalite ve morbiditesi yüksek bir hastalıktır. Bizim çalışmamızda düşük doğum ağırlığı, kan gazı laktat ve parsiyel karbondioksit değerlerinin mortalite ilişkisini saptadık. Çok merkezli prospektif çalışmalar yapılarak sağ kalımı artıracak izlem rehberlerinin oluşturulması gerekmektedir.

Anahtar Kelimeler: Konjenital diyafragma hernisi, mortalite, yenidoğan

ABSTRACT

Objective: Congenital diaphragmatic hernia is a congenital defect with a high mortality rate due to developmental defect of the diaphragm which causes intra-abdominal organs to be located in the thorax. Right-sided hernia, male gender, liver and stomach herniation, and degree of pulmonary hypoplasia are associated with mortality. In our study, we aimed to evaluate the factors related to mortality in diaphragmatic hernias.

Methods: We retrospectively analysed the patients who were followed up with a diagnosis of congenital diaphragmatic hernia between December 2021 and March 2025. We examined the factors affecting mortality by comparing the demographic characteristics and postnatal findings of the patients.

Results: A total of 30 congenital diaphragmatic hernias were evaluated. Twenty-one of the patients were operated, 9 patients died within the first 24 hours due to lack of cardiovascular stability. While 70% (n=21) of the patients died, 30% (n=9) survived. A lactate value of 2.3 mmol/L [area under the curve (AUC): 0.79, p<0.01] and a pCO_2 (AUC: 0.98, p<0.01) value above 60 mmHg in the first blood gas obtained during neonatal intensive care unit hospitalisation were found to be associated with mortality.

Conclusion: Congenital diaphragmatic hernia is a disease with high mortality and morbidity. In our study, we found that low birth weight, blood gas lactate and partial carbon dioxide values were associated with mortality. Multicentre prospective studies should be conducted to establish follow-up guidelines that will increase survival.

Keywords: Congenital diaphragmatic hernia, mortality, newborn

Received/Geliş: 19.07.2025
Accepted/Kabul: 01.12.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Dr. Duygu TUNÇEL,
Sağlık Bilimleri Üniversitesi, Diyarbakır Gazi Yaşargil Eğitim ve Araştırma Hastanesi, Çocuk Sağlığı ve Hastalıkları Kliniği, Neonatoloji, Diyarbakır, Türkiye
✉ tncldyg@yahoo.com
ORCID: 0000-0002-2130-6821

GİRİŞ

Konjenital diyafragma hernisi (KDH), diyaframın gelişimsel kusuruna bağlı olarak batın içi organların toraks boşluğuna yerleşmesine neden olan, mortalitesi yüksek doğumsal bir defekttir. Sıklığı 3000-5000 canlı doğumda 1'dir.[1,2] Vakaların yaklaşık %80'inde defekt sol tarafta görülürken, sağ taraflı herni oranı daha düşüktür.[3] KDH izole şekilde görülebileceği gibi, diğer sistem anomalilerinin eşlik ettiği kompleks formlarda da ortaya çıkabilir.[4] KDH'nde mortalite oranı literatürde %30 ile %60 arasında bildirilmektedir.[5-7] KDH'nin etiyolojisinde maternal özellikler ve genetik faktörlerin rol oynadığı düşünülmektedir; ancak mekanizma halen tam olarak aydınlatılamamıştır. Maternal sigara ve alkol kullanımı, genç anne yaşı ve maternal diyabetin KDH gelişiminde etkili olabileceğini gösteren çalışmalar bulunmaktadır.[8-10]

KDH, perinatal tıptaki gelişmeler sayesinde günümüzde antenatal dönemde tanı alabilen ve çeşitli tedavi yöntemlerinin denendiği bir hastalıktır.[11] Sağkalımın en önemli belirleyicileri pulmoner hipoplazi ve ilişkili pulmoner hipertansiyon (PHT)'dur.[12] KDH'li olgularda mortalite üzerine etkili olabilecek demografik ve klinik değişkenlerin incelendiği çok sayıda çalışma mevcuttur. Bu çalışmalarda sağ yerleşimli herni, erkek cinsiyet, ilk 24 saatte pCO_2 değerinin >60 mmHg olması, maternal düşük (<16) ya da ileri (>40) yaş, karaciğer ve mide herniasyonu mortalite ile yakından ilişkili bulunmuştur.[3,13-18]

Prematüre bebeklerde KDH'nin prognozunu etkileyen faktörlerin araştırıldığı bir çalışmada, 37 haftadan küçük doğan KDH'li bebeklerde sağkalım oranı anlamlı şekilde düşük bulunmuştur.[19] Bu olumsuz etki yalnızca düşük doğum ağırlığından değil, yetersiz akciğer gelişimi ve ilişkili komplikasyonlardan da kaynaklanmaktadır. Mekanik ventilasyon yönteminin sağkalım üzerine etkili olmadığı bildirilmiştir.[2,15,20,21] KDH'nde mortalite ile ilişkili faktörlerin belirlenmesi, operasyon zamanı ve tedavi modaliteleri üzerine yüksek kanıt düzeyine sahip öneriler geliştirerek rehber oluşturulması açısından önem taşımaktadır.

Bu çalışmanın amacı, kliniğimizde izlenen KDH tanılı hastaların demografik ve klinik verilerini inceleyerek mortaliteyi etkileyen faktörleri belirlemek ve literatüre katkı sunmaktır.

YÖNTEM

Hastanemiz üçüncü basamak yenidoğan yoğun bakım ünitesinde 01.12.2021-01.03.2025 tarihleri arasında KDH tanısı ile izlenen hastaların verileri retrospektif olarak incelendi. Hastaların demografik verileri ve antenatal özellikleri; cinsiyet, gestasyon yaşı, doğum şekli, doğumda canlandırma gereksinimi, anne yaşı, anne-baba akrabalığı, gebelik sayısı, diyafragma hernili kardeş öyküsü ve antenatal tanı varlığı hasta dosyalarından kaydedildi.

Antenatal veya postnatal dönemde KDH tanısı alan bebekler çalışmaya dâhil edildi. Verilerine ulaşılamayan ya da aile onamı olmayan olgular çalışma dışı bırakıldı. Tüm hastalara yenidoğan yoğun bakım yatışında umbilikal venöz kateterizasyon yapıldı ve kan gazı örneği alındı. Yaşamın ilk bir saatinde yapılan kan gazı analizlerinde pH, baz açığı (BE, mEq/L), laktat düzeyi (mmol/L), bikarbonat (mEq/L) ve pCO_2 (mmHg) değerleri ölçüldü. Ekokardiyografik incelemede konjenital kalp hastalığı (KKH) ve PHT varlığı değerlendirildi. PHT; triküspit kapakta yetmezlik, sağ atriyum basıncının 35 mmHg'nin üzerinde olması ve patent foramen ovale veya patent duktus arteriozus düzeyinde sağ-sol ya da bidireksiyonel şant varlığı olarak kabul edildi. Yenidoğan döneminde PHT etiyolojik olarak parankimal akciğer hastalıklarına bağlı vazokonstriksiyon, hipoplastik pulmoner vasküler yatak veya yapısal vasküler malformasyonlar şeklinde sınıflandırılır; KDH'nde genellikle hipoplazik pulmoner vasküler yatağa bağlı PHT görülür.[22]

Torakal ve abdominal boşluk ultrasonografik olarak değerlendirilerek herni içeriği (karaciğer, dalak, böbrek, mide, barsaklar) belirlendi. Klinik izlem sırasında periton diyaliz gereksinimi, hava kaçağı gelişimi, mekanik ventilatör ihtiyacı ve tipi ile inotrop gereksinimi hasta kayıtlarından elde edildi.

Mortalite gelişen ve gelişmeyen olgular iki gruba ayrıldı. Gruplar arasında antenatal ve postnatal özellikler karşılaştırıldı.

Çalışma için Sağlık Bilimleri Üniversitesi, Diyarbakır Gazi Yaşargil Eğitim ve Araştırma Hastanesi Etik Kurulu onayı alındı (karar numarası: 328, tarih: 07.02.2025).

İstatistiksel Analiz

İstatistiksel analizler SPSS v.22.0 yazılımı (SPSS Inc., Chicago, IL, ABD) kullanılarak yapıldı. Tanımlayıcı istatistikler sayı, ortalama ± standart sapma ve yüzde (%) değerleri olarak sunuldu. Değişkenlerin normal dağılıma uygunluğu Kolmogorov-Smirnov testi ile değerlendirildi. İki grup arasındaki kategorik değişkenlerin karşılaştırılmasında ki-kare testi, normal dağılım gösteren sürekli değişkenlerde Student t-testi, normal dağılmayanlarda ise Mann-Whitney U testi kullanıldı. $p<0,01$ anlamlı kabul edildi. Mortaliteyi etkileyen faktörlerin optimal kesim değerlerine göre duyarlılık ve özgüllüklerini belirlemek için alıcı işletim karakteristiği (AİK) analizi yapıldı. Eğri altındaki alan (EAA) değerleri hesaplandı. Parametrelerin EAA değerlerinin karşılaştırılmasında DeLong testi kullanıldı (duyarlılık +1 – özgüllük). EAA değeri 0,5-0,6 arasında ise kötü, 0,6-0,7 orta, 0,7-0,8 kabul edilebilir, 0,8-0,9 mükemmel ve >0,9 olağanüstü olarak değerlendirildi.

BULGULAR

Hastanemiz yenidoğan yoğun bakım ünitesinde toplam 30 hasta KDH tanısı ile takip edildi. Hastalardan yalnızca biri dış merkezden KDH ön tanısı ile sevk edilmişti. Olguların %63,3'ü sezaryen ile doğmuş olup, %93,3'ü antenatal dönemde tanı almıştı. Hastaların %96,6'sında sol diyafragma hernisi mevcuttu. Yenidoğanların ortalama doğum haftası 36,5 (minimum-maksimum: 29-40) hafta, ortalama doğum ağırlığı ise 2600 g (minimum-maks: 1000-3400) olarak bulundu (Tablo 1).

Hastaların %70'i (n=21) exitus olurken, sağkalım oranı %30 (n=9) idi. Erkek cinsiyet, doğum şekli, antenatal tanı varlığı ve anne yaşı ile mortalite arasında anlamlı ilişki saptanmadı. Herniye olan organlar arasında karaciğer (p=0,61) ve mide herniasyonu (p=0,68) açısından mortalite ile sağ kalan grup arasında anlamlı fark görülmedi. Mortalite gelişen hastaların ortalama doğum ağırlığı 2371±551 g, sağ kalan bebeklerin ise 2935±258 g idi; fark istatistiksel olarak anlamlı bulundu (p<0,01). Doğum haftası mortalite grubunda ortalama 35±2,5 hafta, sağ kalanlarda 37,5±1,7 hafta olup fark istatistiksel olarak anlamlı değildi (p=0,073) (Tablo 2).

Tablo 1. Hastaların demografik özellikleri

Değişken	Sayı (n=30)	(%) yüzde
Cinsiyet-erkek	19	63,3
Doğum şekli-sezaryen	11	36,6
Antenatal tanı	28	93,3
Sol herni	29	96,6
Anne yaşı (ortanca-aralık)	29 (19-40)	
5. dakika Apgar (ortanca-aralık)	4 (1-8)	
Doğum haftası (hafta, ortanca-aralık)	36,5 (29-40)	
Doğum ağırlığı (g, ortanca-aralık)	2600 (1000-3400)	

Sürekli değişkenler ortanca (aralık) veya ortalama (± standard sapma) olarak belirtilmiştir

Toplam 30 hastanın 21'i mortal seyretmiş olup, bu hastaların 9'una yaşamın ilk 24 saati içinde kaybedildikleri için ekokardiyografi yapılamamıştır. Ekokardiyografi yapılabilen 21 hastanın 12'sinde (%57,1) PHT saptanmıştır. PHT saptanan bu 12 hastanın 8'i mortal, 4'ü sağ kalan gruptadır (Tablo 2). Olguların 2'sinde (%6,6) siyanotik kalp hastalığı, 2'sinde (%6,6) ventriküler septal defekt veya atriyoventriküler septal defekt saptandı. Eşlik eden diğer anomaliler arasında santral sinir sistemi anomalisi, ekstremite kısalığı ve yarık damak-dudak yer aldı. Diğer sistemlere ait konjenital anomaliler mortalite ile ilişkili değildi (p>0,01).

Beşinci dakika APGAR skorları hayatta kalan bebeklerde daha yüksek (medyan: 7, minimum-maks: 5-9) olmasına rağmen fark istatistiksel olarak anlamlı bulunmadı (p=0,087). Hava kaçağı gelişimi mortalite grubunda daha sık görülmekle birlikte (n=18 vs. n=4), fark anlamlı değildi (p=0,16) (Tablo 2).

Kan gazı analizlerinde mortalite grubundaki bebeklerde pCO_2 düzeyleri (99-40 vs. 40-69 mmHg) ve laktat düzeyleri (2,9-5,9 vs. 2,1-6,9 mmol/L) anlamlı derecede yüksek bulundu (p<0,01) (Tablo 2). Mortalite ile ilişkili parametreler arasında laktat ve pCO_2 için yapılan AİK analizinde EAA

Tablo 2. Konjenital diyafragma hernisinde mortalite ile ilişkili faktörler

Değişken	Mortalite var (n=21,%70)	Mortalite yok (n=9, %30)	p değeri
Cinsiyet (erkek), n (%)	14 (66,7)	5 (55,6)	1,00
Doğum şekli (sezaryen), n (%)	7 (33,3)	4 (44,4)	0,41
Antenatal tanı, n (%)	20 (95,2)	8 (88,9)	1,00
Anne yaşı (yıl)	29 (19-40)	28,5 (14-37)	0,37
5. dakika Apgar	6 (3-9)	7 (5-9)	0,087
Doğum haftası (hafta)	35±2.5	37,5±1,7	0,073
Doğum ağırlığı (g)	2371±551	2935±258	<0,01
Karaciğer herniasyonu, n (%)	13(61,9)	3 (33,3)	0,61
Mide herniasyonu, n (%)	10(47,6)	5 (55,6)	0,68
pCO_2 (mmHg)	99-40	40-69	<0,01
Laktat (mmol/L)	2,9-5,9	2,1-6,9	<0,01
Hava kaçağı, n (%)	18 (85,7)	4 (44,4)	0,16
Pulmoner hipertansiyon, n (%)	8 (38,1)	4 (44,4)	0,31
Operasyon günü (ortanca)	0 (0-3)	1,5 (1-3)	0,04

Kalın yazılmış p değerleri istatistiksel anlamlılığı göstermektedir (p<0,05)

sırasıyla 0,79 (0,7-0,8) ve 0,98 (>0,9) olarak bulundu. pCO_2 değeri yüksek düzeyde anlamlı kabul edildi. Laktat ve pCO_2'nin kesim değerlerine göre yapılan analizde laktat için 2,3 mmol/L düzeyinde sensitivite %72,5 ve spesifisite %62,5; pCO_2 için >60 mmHg düzeyinde sensitivite %81,8 ve spesifisite %100 oranında mortalite ile ilişkiliydi.

AİK analizinde hastaların 5. dakika APGAR skoru, gestasyon yaşı ve doğum ağırlığının mortalite üzerine anlamlı etkisi saptanmadı (Tablo 3 ve 4, Şekil 1).

TARTIŞMA

KDH, perinatal tıptaki gelişmelere ve neonatal bakım olanaklarının iyileşmesine rağmen hâlâ önemli bir mortalite nedenidir. Bizim çalışmamızda mortalite oranı %70,9 olarak saptandı. Çok merkezli bir çalışmada gelişmiş ülkelerde mortalite oranı %20-30 iken, az gelişmiş ve gelişmekte olan ülkelerde bu oran %90'lara ulaşmaktadır.[23,24] Ülkemizde yapılan çalışmalarda mortalite oranı %59,5 olarak, bir diğerinde ise sağkalım oranı %32,8 olarak bildirilmiştir.[2,25] Bizim mortalite oranımızın literatüre göre yüksek olması, olgulardaki pulmoner hipoplazinin ciddiyetiyle ilişkili olabilir. Çünkü hastalarımızda hedef preduktal ve postduktal satürasyonlar yenidoğan yoğun bakım yatışından itibaren sağlanamamış, olguların %60'ı (n=20) preoperatif dönemde yüksek frekanslı mekanik ventilasyon (HFOV) ile izlenmiştir. Bu durum mevcut pulmoner rezervin yetersizliğini göstermektedir. Gelişmiş ülkelerde KDH sağkalım oranlarının inhalasyonla verilen nitrik oksit ve ekstrakorporeal membran oksijenizasyonu (ECMO) uygulamalarının yaygınlaşmasıyla son 10 yılda %20-30 düzeylerine düştüğü bildirilmiştir.[10] Merkezimizde solunum yetmezliği nedeniyle ekstrakorporeal tedavilerin uygulanamaması mortaliteyi etkileyen bir faktör olarak değerlendirilmektedir.

Tablo 3. Serum laktat, Apgar skoru, gestasyonel hafta ve parsiyel karbondioksit basıncı ile mortalite ilişkisi

Değişken	EAA (%95 GA)	p değeri
Serum laktat (mmol/L)	0,79	0,01
5. dakika Apgar skoru	0,29	0,041
Gestasyonel hafta	0,28	0,042
Doğum ağırlığı (g)	0,14	0,07
pCO_2 (mHg)	0,98	<0,01

EAA: Eğri altındaki alan, GA: Güven aralığı

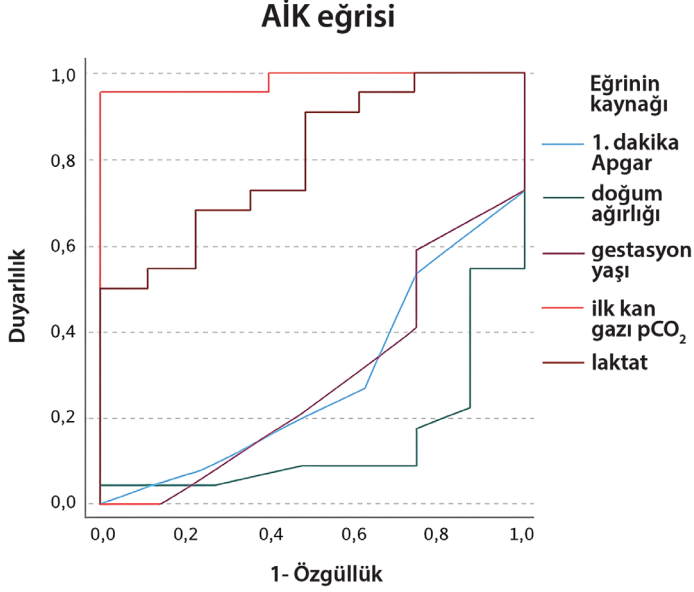
Tablo 4. Mortalite gelişen hasta grubunda serum laktat ve pCO_2 değerleri

Değişken	Eşik değeri	(%) Duyarlılık	(%) Özgüllük
Serum laktat (mmol/L)	2,3	72,5	62,5
pCO_2 (mmHg)	60	81,8	100

KDH'de erkek cinsiyetin mortaliteyi artırabileceği bildirilmiştir.[2,7,24,26] Çalışmamızda erkek cinsiyet oranı %61 olup, mortaliteyle ilişkili bulunmadı. Literatürde karaciğer ve mide herniasyonunun postnatal mortaliteyi artırdığı bildirilmiştir.[15,16] Bizim çalışmamızda da bu organların toraksta yer almasının mortaliteye etkisi istatistiksel olarak anlamlı değildi.

Doğum odasından itibaren KDH'li olgularda akciğer koruyucu ventilasyon stratejileri uygulanmalıdır. Hipoplazik akciğerlerin barotravmadan korunması kısa ve uzun dönem sonuçları doğrudan etkiler.[3,7,17,23] Başlangıç ventilasyon modu seçimlerinin sağkalım üzerindeki etkilerini inceleyen çalışmalarda HFOV ventilasyon ve perfüzyonu iyileştirdiği, ancak mortaliteyi etkilemediği bildirilmiştir.[18] Bizim olgularımızın %64'ünde başlangıç modu olarak HFOV uygulanmış, ancak sağkalım üzerine anlamlı etkisi saptanmamıştır. Hastalarımızın %56'sında orta veya ağır PHT saptanmış, medikal tedavi verilmiştir. Ekokardiyografik olarak değerlendirilen bu vakalarda PHT varlığı mortaliteyi etkilememiştir. Bunun hasta sayısının kısıtlı olmasından kaynaklandığını düşünmekteyiz.

Doğum ağırlığı ile mortalite arasındaki ilişki birçok çalışmada incelenmiştir.[27-30] Bu çalışmalardan yalnızca birinde doğum ağırlığının sağkalımı etkileyen anlamlı bir faktör olduğu belirtilmiştir.[27] Bizim çalışmamızda da exitus olan hastalarda doğum ağırlığı istatistiksel olarak anlamlı biçimde düşük bulundu. Prematürite, pulmoner hipoplaziyi şiddetlendirerek mortaliteyi artırabilir.



Şekil 1. Doğum ağırlığı, gestasyon yaşı, pCO_2, laktat ve 5. dakika Apgar skorunun mortalite ile ilişkisi

AİK analizinden elde edilen eğri altındaki alan değerleri Tablo 3'te sunulmuştur. AİK: Alıcı işletim karakteristiği

Retrospektif bir çalışmada 34. gebelik haftası ve altındaki olgularda mortalite oranları daha yüksek bildirilmiş, ancak fark istatistiksel olarak anlamlı bulunmamıştır.[31] Çalışmamızda gestasyon haftası mortaliteyi etkilemedi; ancak 32 haftanın altında doğan iki olgu exitus oldu. Hasta sayısının az olması bu ilişkinin istatistiksel olarak gösterilememesine neden olabilir.

Eşlik eden anomaliler, özellikle KKH, KDH'de sağkalımı önemli ölçüde etkileyen faktörlerdir. Kompleks KDH olgularında mortalite oranlarının daha yüksek olduğu bildirilmiştir.[2,4,7,10,11,31] Bizim serimizde 2 hastada siyanotik KKH ve ek olarak sinir sistemi anomalisi, ekstremite kısalığı ve yarık damak-dudak mevcuttu. Ancak ek anomali varlığı mortaliteyle ilişkili bulunmadı.

Postnatal dönemde mortaliteyi öngörmek ve ventilasyon stratejilerini, operasyon zamanını ve ECMO gereksinimini belirlemek için daha fazla veriye ihtiyaç vardır. İlk 24 saatteki kan gazı parametrelerinin bu amaçla kullanılabileceği bildirilmiştir. Bir çalışmada ilk 24 saatteki en yüksek pCO_2 değerinin mortalite ve ECMO gereksinimini öngördüğü rapor edilmiştir.[17] Bizim çalışmamızda pCO_2 >60 mmHg değeri mortalite ile anlamlı biçimde ilişkili bulundu.

Operasyon zamanına ilişkin literatürde net bir görüş birliği yoktur. Avrupa ve Kanada KDH rehberlerinde hastanın kardiyovasküler ve solunumsal stabilizasyonu sağlandıktan sonra elektif cerrahi yapılması önerilmektedir.[12,30] Bizim olgularımızda 9 hasta postnatal ilk 24 saat içinde ağır solunum yetmezliği nedeniyle kaybedildi. Klinik olarak stabil hale getirilen 5 olgu postnatal ilk 24 saatte, 6 olgu ise 24-48 saat arasında opere edildi; medyan operasyon günü 0,8 idi. Operasyon zamanının mortaliteyle ilişkisi bulunmadı.

Çalışma Kısıtlılıkları

Çalışmamızın retrospektif tasarımı ve sınırlı örneklem sayısı, sonuçların genellenebilirliğini kısıtlamaktadır. Postnatal ilk gün kaybedilen hastalarda klinik ve laboratuvar değerlendirmelerin tamamlanamaması sonuçların güvenilirliğini azaltmaktadır.

SONUÇ

KDH, antenatal dönemden itibaren multidisipliner yaklaşım gerektiren ciddi bir neonatal solunum problemidir. Sağkalım oranlarının artırılması için merkezlerin tedavi protokollerini bireyselleştirerek doğum odası stabilizasyonu, mekanik ventilasyon stratejileri, kan gazı hedefleri ve operasyon zamanını dikkatle belirlemesi gerekmektedir. Ayrıca seçilmiş hastalarda ekstrakorporeal destek tedavilerinin uygulanabilirliğinin değerlendirilmesi önemlidir.

Etik

Etik Kurul Onayı ve Hasta Onayı: Çalışma için Sağlık Bilimleri Üniversitesi, Diyarbakır Gazi Yaşargil Eğitim ve Araştırma Hastanesi Etik Kurulu onayı alındı (karar numarası: 328, tarih: 07.02.2025). Retrospektif bir çalışmalıdır.

Dipnotlar

Yazarlık Katkıları

Cerrahi ve Medikal Uygulama: S.Ç., Konsept: D.T., L.Ş., S.Ç., N.O., Dizayn: D.T., L.Ş., S.Ç., N.O., Veri Toplama veya İşleme: D.T., L.Ş., N.O., Analiz veya Yorumlama: D.T., N.O., Literatür Arama: D.T., L.Ş., N.O., Yazan: D.T., N.O.

Çıkar Çatışması: Yazarlar bu makale ile ilgili olarak herhangi bir çıkar çatışması bildirmemiştir.

Finansal Destek: Çalışmamız için hiçbir kurum ya da kişiden finansal destek alınmamıştır.

KAYNAKLAR

1. Stege G, Fenton A, Jaffray B. Nihilism in the 1990s: the true mortality of congenital diaphragmatic hernia. Pediatrics. 2003;112:532-5.
2. Kandemir İ, Alp Ünkar Z, Kersin SG, Köle MT, Yaman A. Early clinical outcomes of congenital diaphragmatic hernia and prognosis: a retrospective multicenter study. Bagcilar Med Bull. 2023;8:155-60.
3. Abramov A, Fan W, Hernan R, et al. Comparative outcomes of right versus left congenital diaphragmatic hernia: a multicenter analysis. J Pediatr Surg. 2020;55:33-8.
4. Paoletti M, Raffler G, Gaffi MS, Antounians L, Lauriti G, Zani A. Prevalence and risk factors for congenital diaphragmatic hernia: a global view. J Pediatr Surg. 2020;55:2297-307.
5. Deveci MF, Alagöz M, Gökce İK, Özdemir R. Konjenital diyafragma hernisi tanılı hastalar; 10 yıllık tek merkez deneyimi. KSÜ Tıp Fak Der. 2023;18:35-8.
6. Vasudev RB, Kumar N, Gadgade BD, Radhakrishna V, Basavaraju M, Anand A. Factors contributing to mortality in neonates with congenital diaphragmatic hernia and eventration. Afr J Paediatr Surg. 2023;20:85-8.
7. Gupta VS, Harting MT, Lally PA, et al. Mortality in congenital diaphragmatic hernia: a multicenter registry study of over 5000 patients over 25 years. Ann Surg. 2023;277:520-7.
8. Long AM, Bunch KJ, Knight M, Kurinczuk JJ, Losty PD; BAPS-CASS. One-year outcomes of infants born with congenital diaphragmatic hernia: a national population cohort study. Arch Dis Child Fetal Neonatal Ed. 2019;104:F643-7.
9. Mesas Burgos C, Ehrén H, Conner P, Frenckner B. Maternal risk factors and perinatal characteristics in congenital diaphragmatic hernia: a nationwide population-based study. Fetal Diagn Ther. 2019;46:385-91.
10. Donahoe PK, Longoni M, High FA. Polygenic causes of congenital diaphragmatic hernia produce common lung pathologies. Am J Pathol. 2016;186:2532-43.

11. Peppa M, De Stavola BL, Loukogeorgakis S, Zylbersztejn A, Gilbert R, De Coppi P. Congenital diaphragmatic hernia subtypes: comparing birth prevalence, occurrence by maternal age, and mortality in a national birth cohort. Paediatr Perinat Epidemiol. 2023;37:143-53.
12. Cordier AG, Russo FM, Deprest J, Benachi A. Prenatal diagnosis, imaging, and prognosis in congenital diaphragmatic hernia. Semin Perinatol. 2020;44:51163.
13. Snoek KG, Reiss IK, Greenough A, et al. Standardized postnatal management of infants with congenital diaphragmatic hernia in europe: the CDH EURO consortium consensus - 2015 update. Neonatology. 2016;110:66-74.
14. Oh C, Youn JK, Han JW, et al. Predicting survival of congenital diaphragmatic hernia on the first day of life. World J Surg. 2019;43:282-90.
15. Snoek KG, Capolupo I, Morini F, et al. Score for neonatal acute physiology-ii predicts outcome in congenital diaphragmatic hernia patients. Pediatr Crit Care Med. 2016;17:540-6.
16. Salas GL, Otaño JC, Cannizzaro CM, Mazzucchelli MT, Goldsmit GS. Congenital diaphragmatic hernia: postnatal predictors of mortality. Arch Argent Pediatr. 2020;118:173-9.
17. Sekhon MK, Fenton SJ, Yoder BA. Comparison of early postnatal prediction models for survival in congenital diaphragmatic hernia. J Perinatol. 2019;39:654-60.
18. Patel MJ, Bell CS, Lally KP, Lally PA, Katakam LI; Congenital Diaphragmatic Hernia Study Group. Lowest PaCO2 on the first day of life predicts mortality and morbidity among infants with congenital diaphragmatic hernia. J Perinatol. 2019;39:229-36.
19. Wild KT, Mathew L, Ades AM, et al. Association between initial ventilation mode and hospital outcomes for severe congenital diaphragmatic hernia. J Perinatol. 2024;44:1353-8.
20. Morini F, Capolupo I, van Weteringen W, Reiss I. Ventilation modalities in infants with congenital diaphragmatic hernia. Semin Pediatr Surg. 2017;26:159-65.
21. Semama C, Vu S, Kyheng M, et al. High-frequency oscillatory ventilation versus conventional ventilation in the respiratory management of term neonates with a congenital diaphragmatic hernia: a retrospective cohort study. Eur J Pediatr. 2022;181:3899-906.
22. Global PaedSurg Research Collaboration. Mortality from gastrointestinal congenital anomalies at 264 hospitals in 74 low-income, middle-income, and high-income countries: a multicentre, international, prospective cohort study. Lancet. 2021;398:325-39.
23. Rivas JFG, Clugston RD. The etiology of congenital diaphragmatic hernia: the retinoid hypothesis 20 years later. J Perinatol. 2023.
24. Aydin E, Torlak N, Haberman B, Lim FY, Peiro JL. The survivorship bias in congenital diaphragmatic hernia. Children (Basel). 2022;9:218.
25. Yamoto M, Tanaka Y, Fukumoto K, et al. Cardiac fetal ultrasonographic parameters for predicting outcomes of isolated left-sided congenital diaphragmatic hernia. J Pediatr Surg. 2015;50:2019-24.
26. Takahashi T, Koga H, Tanaka T, et al. Pulmonary artery size has prognostic value in low birth weight infants with congenital diaphragmatic hernia. Pediatr Surg Int. 2011;27:847-50.
27. Lazar DA, Ruano R, Cass DL, et al. Defining "liver-up": does the volume of liver herniation predict outcome for fetuses with isolated left-sided congenital diaphragmatic hernia? J Pediatr Surg. 2012;47:1058-62.
28. Hidaka N, Ishii K, Furutake Y, Yamamoto R, Sasahara J, Mitsuda N. Magnetic resonance fetal right lung volumetry and its efficacy in predicting postnatal short-term outcomes of congenital left-sided diaphragmatic hernia. J Obstet Gynaecol Res. 2014;40:429-38.
29. Partridge EA, Peranteau WH, Herkert L, et al. Rate of increase of lung-to-head ratio over the course of gestation is predictive of survival in left-sided congenital diaphragmatic hernia. J Pediatr Surg. 2016;51:703-5.
30. Paoletti M, Raffler G, Gaffi MS, Antounians L, Lauriti G, Zani A. Prevalence and risk factors for congenital diaphragmatic hernia: a global view. J Pediatr Surg. 2020;55:2297-307.
31. Canadian Congenital Diaphragmatic Hernia Collaborative; Puligandla PS, Skarsgard ED, et al. Diagnosis and management of congenital diaphragmatic hernia: a clinical practice guideline. CMAJ. 2018;190:E103-12.

DOI: 10.4274/forbes.galenos.2025.77045
Forbes J Med 2025;6(3):299-302

Case Report / Olgu Sunumu

Horizontal Gaze Palsy with Progressive Scoliosis (HGPPS) in a Consanguineous Family: A Case Report

Akraba Evliliği Olan Bir Ailede Horizontal Bakış Paralizisi ve Progresif Skolyoz: Bir Olgu Sunumu

Dilara Ece TOPRAK[1], Gözde Merve TEKEL[2], Yusuf Can DOĞAN[3], Esra IŞIK[3], Atilla Hikmet ÇİLENGİR[2]

[1]İzmir Democracy University Buca Seyfi Demirsoy Training and Research Hospital, Department of Pediatric Neurology, İzmir, Türkiye
[2]İzmir Democracy University Buca Seyfi Demirsoy Training and Research Hospital, Department of Radiology, İzmir, Türkiye
[3]Ege University Faculty of Medicine, Department of Pediatric Genetics, İzmir, Türkiye

Cite as: Toprak DE, Tekel GM, Doğan YC, Işık E, Horizontal gaze palsy with progressive scoliosis (HGPPS) in a consanguineous family: a case report. 2025;6(3):299-302

ABSTRACT

Horizontal gaze palsy with progressive scoliosis is a rare autosomal recessive disorder caused by biallelic variants in *ROBO3*, a gene essential for commissural axon crossing in the hindbrain and spinal cord. It is characterized by the congenital absence of horizontal eye movements and progressive scoliosis, typically presenting in childhood or adolescence.

Keywords: Horizontal gaze palsy with progressive scoliosis, HGPPS, *ROBO3*, brainstem, malformation

ÖZ

Horizontal bakış paralizisi ve ilerleyici skolyoz, beyin sapı ve omurilikte komissural akson geçişi için gerekli olan *ROBO3* genindeki bialelik varyantlardan kaynaklanan nadir bir otozomal resesif bozukluktur. Doğumsal horizontal göz hareketlerinin yokluğu ve genellikle çocukluk veya ergenlik döneminde ortaya çıkan ilerleyici skolyoz ile karakterizedir.

Anahtar Kelimeler: Horizontal bakış paralizisi ve ilerleyici skolyoz, HGPPS, *ROBO3*, beyin sapı, malformasyon

Received/Geliş: 14.09.2025
Accepted/Kabul: 28.11.2025
Publication Date/ Yayınlanma Tarihi: 05.12.2025

Corresponding Author/ Sorumlu Yazar:
Dilara Ece TOPRAK, MD,
İzmir Democracy University Buca Seyfi Demirsoy Training and Research Hospital, Department of Pediatric Neurology, İzmir, Türkiye
dilaraecetoprak@yahoo.com
ORCID: 0000-0002-3833-1230

INTRODUCTION

Horizontal gaze palsy with progressive scoliosis (HGPPS) is a rare autosomal recessive disorder first described in 1974.[1] It is characterized by a congenital absence of horizontal eye movements and progressive scoliosis, both of which usually become apparent during childhood or adolescence. The condition is caused by biallelic pathogenic variants in the Roundabout homolog 3 *(ROBO3)* gene, which encodes a receptor essential for commissural axon crossing within the hindbrain and spinal cord.[2] Impaired *ROBO3*-mediated signaling results in defective horizontal gaze control and abnormal spinal cord circuitry underlying scoliosis progression. To date, almost 100 patients with HGPPS have been reported and 55 *ROBO3* mutations have been identified, most of them occurring in consanguineous families.[3] We present a 17-year-old male with genetically confirmed HGPPS, highlighting the clinical, radiological, and genetic findings and intrafamilial recurrence.

CASE REPORT

A 17-year-old male was referred for evaluation of impaired horizontal eye movements, scoliosis, and decreased vision. His parents were consanguineous (first cousins). The family history revealed an older brother with a similar clinical phenotype, including limitation of horizontal gaze and scoliosis requiring surgical correction, whereas another brother was healthy. Both parents were unaffected. Segregation analysis could not be performed in the parents or the affected brother due to technical limitations; however, the same variant was presumed to be present in the affected brother.

The patient's medical history included congenital torticollis identified during infancy, along with horizontal gaze restriction observed in early childhood. Cervicothoracic and thoracolumbar scoliosis progressed with advancing age. Despite these findings, his cognitive and motor developmental milestones were appropriate for his chronological age.

Neurological examination revealed complete restriction of horizontal eye movements, with preserved upward and downward gaze. The remainder of the cranial nerve examination was normal, with no facial asymmetry. Examinations of the pyramidal and extrapyramidal systems were unremarkable. Ophthalmological evaluation showed bilateral optic atrophy, and the patient required corrective lenses. Spinal evaluation identified an S-shaped scoliosis, characterized by a left-convex curve at the cervical and upper thoracic levels and a right-convex curve at the lower thoracic and upper lumbar regions.

Cranial magnetic resonance imaging (MRI) demonstrated mild hypoplasia of the medulla oblongata and pons within the posterior fossa, along with a midline cleft extending to the floor of the fourth ventricle and involving the midsagittal planes of both structures (Figure 1). Spinal MRI demonstrated scoliosis, with a Cobb angle of 12°, showing left convexity at the cervical and upper thoracic levels and right convexity at the lower thoracic and upper lumbar levels (Figure 2). A long-segment syrinx measuring up to 3 mm was observed within the thoracic spinal cord at the level of the T8-9 disc. Additionally, minor central protrusions were noted at the C4-5 and C6-7 disc levels. The transverse processes of the L5 vertebra were sacralized. At the L4-5 disc level, degenerative signal loss and mild annular bulging were present. Genetic testing using a

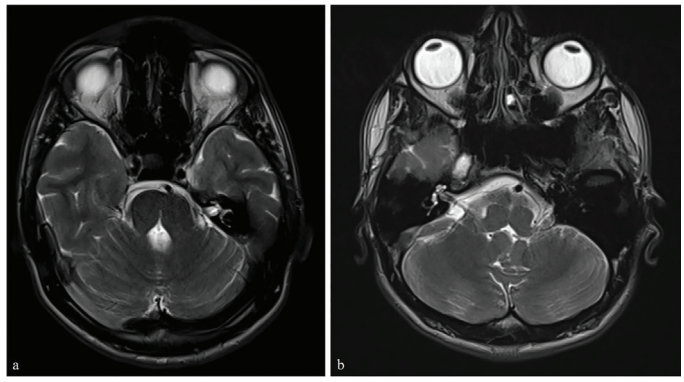


Figure 1. Axial T2-weighted cranial magnetic resonance imaging demonstrates a fissure extending from the posterior midsagittal aspects of the pons (a) and the medulla oblongata (b) to the floor of the fourth ventricle

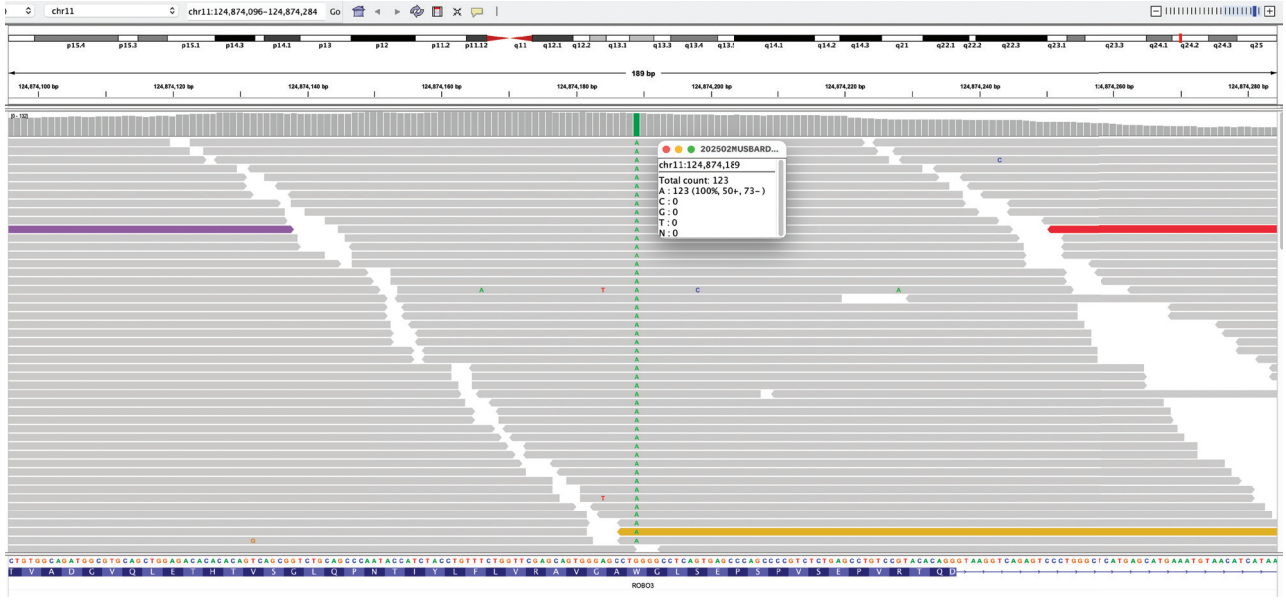


Figure 2. Integrative genomics viewer screenshot showing the homozygous c.1904G>A variant in *ROBO3*

skeletal dysplasia gene panel identified a homozygous, likely pathogenic variant in *ROBO3* (c.1904G>A, p.Trp635*), a gene previously associated with HGPPS (Figure 2). Based on these clinical and radiological findings, a diagnosis of HGPPS was considered. Notably, the homozygous c.1904G>A (p.Trp635*) variant identified in our patient has not been previously reported in the literature.

DISCUSSION

HGPPS is a distinct clinical entity that should be suspected in patients presenting with congenital horizontal gaze palsy and progressive scoliosis, particularly in the setting of consanguinity.[4] Our patient exhibited the classical clinical features, including absence of horizontal eye movements, torticollis in infancy, and progressive scoliosis in adolescence. The presence of an affected sibling in the family further supported the autosomal recessive inheritance pattern.

At initial evaluation, the differential diagnosis encompassed a broad spectrum of genetic and mitochondrial disorders, including Duane retraction syndrome, Moebius syndrome, congenital fibrosis of the extraocular muscles, congenital oculomotor apraxia, and polymerase gamma-related mitochondrial disease, because of the combination of ophthalmological and neurological manifestations.[5] However, the characteristic brain MRI findings-hypoplasia of the pons and medulla with a midline cleft extending into the floor of the fourth ventricle-together with progressive scoliosis were pivotal in narrowing the differential diagnosis. These features strongly suggested HGPPS, enabling us to establish a targeted genetic hypothesis and reach a definitive diagnosis more rapidly.

Neuroimaging abnormalities are highly suggestive of HGPPS and reflect impaired axonal guidance in the hindbrain resulting from ROBO3 dysfunction.[6] The identification of a homozygous likely pathogenic *ROBO3* variant (c.1904G>A) in our patient provided molecular confirmation of the diagnosis. Although segregation testing could not be performed, the affected brother is highly likely to carry the same variant, further supporting its pathogenic role. This variant has not been previously reported in any published case, representing a novel addition to the *ROBO3* mutational spectrum. This truncating variant introduces a premature stop codon predicted to result in loss of function, which aligns with the established disease mechanism. In the limited number of cases reported to date, loss-of-function variants have consistently been implicated in HGPPS, further reinforcing the pathogenicity of the truncating variant observed in our patient. Our case also highlights an unusual feature-bilateral optic atrophy-which has not been consistently reported in HGPPS and may represent either a coincidental finding or an expansion of the phenotype. Longitudinal follow-up is necessary to clarify its clinical relevance.

Management of HGPPS remains supportive, primarily consisting of ophthalmological monitoring and orthopedic interventions for scoliosis.[7] Early recognition is important not only to anticipate complications but also to guide genetic counseling and to avoid unnecessary investigations for other neurogenetic or mitochondrial disorders.

CONCLUSION

We report a consanguineous family with two affected siblings who present with the classical features of HGPPS and who carry a homozygous, likely pathogenic ROBO3 variant. This case underlines the importance of considering HGPPS in patients with congenital horizontal gaze palsy and progressive scoliosis, and emphasizes the diagnostic utility of brain MRI and genetic testing. In our patient, the characteristic MRI findings, together with scoliosis, were the key diagnostic clues that directed us toward HGPPS and enabled earlier genetic confirmation, thereby avoiding extensive investigations for other mitochondrial or neurogenetic disorders. Recognition of this rare disorder is essential for timely counseling, surveillance, and supportive management.

Ethics

Informed Consent: The authors confirm that written informed consent was obtained from the patient (and/or the patient's legal guardian) for publication of the clinical details and accompanying images in this report.

Footnotes

Authorship Contributions

Concept: G.M.T., Data Collection or Processing: D.E.T., G.M.T., Y.C.D., E.I., A.H.Ç., Analysis or Interpretation: D.E.T., G.M.T., Y.C.D., E.I., A.H.Ç., Literature Search: D.E.T., G.M.T., Writing: D.E.T., G.M.T.

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

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

REFERENCES

1. Dretaki Dretakis EK, Kondoyannis PN. Congenital scoliosis associated with encephalopathy in five children of two families. J Bone Joint Surg Am. 1974;56:1747-50.
2. Deniz A, Çomu S, Güngör M, Anık Y, Kara B. Compound heterozygous ROBO3 mutation in two siblings presenting with horizontal gaze palsy without scoliosis: case-based review. J Pediatr Genet. 2021;13:116-22.
3. Yi S, Qin Z, Zhou X, et al. Early onset horizontal gaze palsy and progressive scoliosis due to a noncanonical splicing-site variant

and a missense variant in the ROBO3 gene. Mol Genet Genomic Med. 2023;11:e2215.

4. Volk AE, Carter O, Fricke J, et al. Horizontal gaze palsy with progressive scoliosis: three novel ROBO3 mutations and descriptions of the phenotypes of four patients. Mol Vis. 2011;17:1978-86.
5. Whitman MC, Engle EC. Ocular congenital cranial dysinnervation disorders (CCDDs): insights into axon growth and guidance. Hum Mol Genet. 2017;26:R37-R44.
6. Jen JC, Chan WM, Bosley TM, et al. Mutations in a human ROBO gene disrupt hindbrain axon pathway crossing and morphogenesis. Science. 2004;304:1509-13.
7. Dolar Bilge A. Horizontal gaze palsy with progressive scoliosis: a case report and literature review. Neuroophthalmology. 2019;43:334-36.

2025 REFEREE INDEX - 2025 HAKEM DİZİNİ

Ali Turhan
Alper Şener
Atilla Hikmet Çilengir
Ayça Töre Başer Tüfekçi
Aydın İşisağ
Aylin Kılınç Uğurlu
Aysel Başer
Ayşe Anık
Ayşe Berna Anıl
Ayşe Serra Artan
Ayşe Şimşek
Banu Karaca
Betül Siyah Bilgin
Burak Sayar
Burçin Tuna
Caner Baysan
Caner Turan
Cevdet Duran
Çağlar Keskin
Deniz Anuk İnce
Didem Aytimur
Dilek Altun
Emine Ayça Cimbek
Emre Emekli
Eren Soyaltın
Erhan Ateş
Erkan Güler
Erol Erden Ünlüer
Ezgi Yangın Ergon
Fatma Ceren Sarıoğlu
Ferruh Acet
Funda Yavanoğlu Atay
Gamze Gökalp
Giray Kolcu
Gökşin Nilüfer Demirci Derici
Hande Gazeteci Tekin
Hüseyin Altunhan
İlhan Yaylım
İlker Gül
İnanç Karakoyun
Kemal Gökçek
Marie Bou Nassif
Mehmet Coşkun
Mehmet Eyüboğlu
Mehmet Güney
Mehmet Kenan Kanburoğlu
Mehmet Üstün
Mehtap Yüksel Eğrilmez
Merve Keskinkılıç
Metin Güçlü
Murat Örmen
Mustafa Kemal Alimoğlu
Mustafa Özgür Pirgon
Orkun Sarıoğlu
Osman Akdeniz
Özlem Gürsoy Doruk
Özüm Erkin
Özgür Tosun
Pelin Kıyat
Rahime Bedir Fındık
Safa Gürsoy
Safiye Aktaş
Selçuk Sinan Çelik
Sema Rala
Serdar Babacan
Sermet Sezigen
Sinem Sarı Gökay
Süheyla Serin Senger
Şadiye Dur
Şule Demir
Tarık Salman
Turgay Turan
Tülay Demircan
Türkan Atasever Rezanko
Ufuk Atlıhan
Ulaş Akgün
Utku Oflazoğlu
Ümit Çıraklı