e-ISSN: 2149-3189

European Research Journal

Volume 11 Issue 6 November 2025

Available at https://dergipark.org.tr/en/pub/eurj

Published by Prusa Medical Publishing

The European Research Journal

Aim and Scope

The European Research Journal (EuRJ) is an international, independent, double-blind peer reviewed, Open Access and online publishing journal, which aims to publish papers on all the related areas of basic and clinical medicine.

Editorial Board of the European Research Journal complies with the criteria of the International Council of Medical Journal Editors (ICMJE), the World Association of Medical Editors (WAME), and Committee on Publication Ethics (COPE).

The journal publishes a variety of manuscripts including original research, case reports, invited review articles, technical reports, how-to-do it, interesting images and letters to the editor. The European Research Journal has signed the declaration of the Budapest Open Access Initiative. All articles are detected for similarity or plagiarism. Publication language is English.

EuRJ recommends that all of our authors obtain their own ORCID identifier which will be included on their article.

The journal is published bimonthly (January, March, May, July, September, and November).

Abstracting and Indexing

The journal is abstracted and indexed with the following: ULAKBIM TR Index (ULAKBİM TR DİZİN), NLM Catalog (NLM ID: 101685727), Google Scholar (h-index: 15), EMBASE, ProQuest Central, EBSCO Academic Search Ultimate, J-Gate, EZB, TURK MEDLINE, Turkish Citation Index, ResearhGate, SOBIAD, ScienceGate, Publons, (Clarivate Web of Science)

Publisher

The European Research Journal (EuRJ)
Prusa Medical Publishing
Konak Mh. Kudret Sk. Şenyurt İş Mrk. Blok No:6 İç kapı no: 3
Nilüfer/Bursa-Türkiye
info@prusamp.com

https://dergipark.org.tr/en/pub/eurj
https://www.prusamp.com

e-ISSN: 2149-3189

EDITORIAL BOARD

EDITOR-IN-CHIEF

EDITORS

OWNER ON BEHALF OF THE PRUSA MEDİCAL PUBLISHING

ASSISTANT EDITOR

SECTION EDITORS

Meliha KASAPOGLU AKSOY, MD.,
Associate Professor,
University of Health Sciences, Bursa Yuksek Ihtisas Training & Research Hospital,
Department of Physical Therapy and Rehabilitation,
Bursa, Türkiye

Arda ISIK, MD.,
Associate Professor,
Medeniyet University School of Medicine,
Department of General Surgery,
Istanbul, Türkiye

Kadir Kaan OZSIN, MD.,
Associate Professor,
University of Health Sciences, Bursa Yuksek Ihtisas Training & Research Hospital,
Department of Cardiovascular Surgery,
Bursa, Türkiye

Cihan AYDIN, MD.,
Associate Professor,
Tekirdağ Namık Kemal University, Faculty of Medicine,
Department of Cardiology,
Tekirdağ, Türkiye

Sayad KOCAHAN, PhD.,
Professor,
University of Health Sciences, Gülhane Medical Faculty,
Department of Physiology,
Ankara, Türkiye

Gokhan OCAKOGLU, PhD.,
Associate Professor,
Uludag University School of Medicine,
Department of Biostatistics,
Bursa, Türkiye

Nurullah DOGAN, MD.,
Associate Professor,
Doruk Nilüfer Hospital,
Department of Radiology,
Bursa, Türkiye

Ömer Faruk KARATAS, PhD.,
Professor,
Erzurum Technical University,
Department of Molecular Biology and Genetics,
Erzurum, Türkiye

Serhat YALÇINKAYA, PhD.,
Associate Professor,
Priavate Bursa NEV Health Group,
Department of Thoracic Surgery,
Bursa, Türkiye

LANGUAGE EDITOR

ETHICAL EDITOR

SCIENTIFIC ADVISORY BOARD

Nurcan ÖZYAZICIOĞLU,
Professor,
Department Nursing and Health Sciences
Bursa Uludağ Unibersity
Bursa, Türkiye

Burcu DİNÇGEZ, MD.,
Professor,
University of Health Sciences, Bursa Yuksek Ihtisas Training & Research Hospital,
Department of Gynecology and Obstetrics,
Bursa, Türkiye

Yenal DUNDAR, MD.,
Consultant Psychiatrist
University of Liverpool,
Liverpool, UK

Başar CANDER, MD.,
Professor,
Bezmialem Vakif University,
Department of Emergency Medicine
İstanbul, Türkiye

Aylin COLPAN, MD.,
Associate Professor,
Jefferson University-Lehigh Valley Hospital
Department of Infectious Diseases
Allentown, ABD

Table of Contents

Original Articles

Review

Case Reports

The European Research Journal 2025;11(6):1029-1042 Original Article

DOI: https://doi.org/10.18621/eurj.1672422 *Artificial Intelligence and Rehabilitation*

DeepTherapy: A mobile platform for osteoarthritis rehabilitation utilizing chain-of-thought reasoning and deep learning

Turgay Tugay Bilgin[1], Muhammed Ferit Avcı[2], Selim Mahmut Günay[3], Büşra Şahin[4], Cetin Sayaca[4], Lale Altan[5], Özden Özkal[4], Tuğberk Coşkun[6], Hakan Özkaynak[2]

[1]Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Bursa Technical University, Bursa Türkiye; [2]Bursa Nilüfer Science and Art Center, Bursa, Türkiye; [3]Department of Physical Therapy and Rehabilitation, Faculty of Health Sciences, Hatay Mustafa Kemal University, Hatay, Türkiye; [4]Department of Physical Therapy and Rehabilitation, Faculty of Health Sciences, Bursa Uludağ University, Bursa, Türkiye; [5]Department of Physical Therapy and Rehabilitation, Faculty of Medicine, Bursa Uludağ University, Bursa, Türkiye; [6]Arnavutköy District Health Directorate, Istanbul, Türkiye

ABSTRACT

Objectives: To develop and evaluate an AI-driven mobile platform that integrates deep learning-based exercise analysis with large language model (LLM) feedback for enhancing osteoarthritis (OA) rehabilitation accessibility and effectiveness.

Methods: A deep learning framework was developed using Long Short-Term Memory (LSTM) architecture to classify exercise phases from video data of 10 rehabilitation exercises. The dataset consisted of approximately 800,000 frames collected from 20 healthy volunteers. A feedback system utilizing chain-of-thought reasoning in LLMs (GPT-4o and Claude 3.5 Sonnet) was implemented to generate targeted corrective feedback. Evaluation was conducted with OA patients (n=2) and physiotherapists (n=7) using the Intraclass Correlation Coefficient (ICC) and Likert scales.

Results: The developed LSTM models achieved 97.8% accuracy in exercise phase classification. Strong agreement between system-generated scores and expert evaluations was demonstrated (ICC=0.85). Physiotherapists slightly preferred Claude's outputs (52.4% vs 47.6%) but rated GPT-4o higher on clinical relevance (4.57/5 vs 4.13/5), clarity (4.71/5 vs 4.38/5), and helpfulness (4.50/5 vs 4.29/5).

Conclusions: DeepTherapy effectively addresses critical limitations in rehabilitation monitoring by providing qualitative movement assessment, identifying incorrect movements, and offering detailed guidance on technique improvement, potentially increasing rehabilitation accessibility while maintaining quality of care.

Keywords: Osteoarthritis, rehabilitation, deep learning, large language models, chain-of-thought reasoning, mobile health

Osteoarthritis (OA) is a degenerative joint disease that affects over 528 million people globally, placing substantial pressure on healthcare systems and significantly impairing the daily lives of those affected [1]. Physiotherapy represents an essential component in OA treatment, helping to reduce

Received: April 9, 2025 **Accepted:** June 26, 2025 **Available Online:** June 30, 2025 **Published:** November 4, 2025

How to cite this article: Bilgin TT, Avcı MF, Günay SM, et al. DeepTherapy: A mobile platform for osteoarthritis rehabilitation utilizing chain-of-thought reasoning and deep learning. Eur Res J. 2025;11(6):1029-1042. doi: 10.18621/eurj.1672422

Corresponding author: Turgay Tugay Bilgin, PhD., Prof., Phone: +90 224 300 35 54, E-mail: turgay.bilgin@btu.edu.tr

Available Online at https://dergipark.org.tr/en/pub/eurj

pain, improve mobility, and maintain independence. However, traditional physiotherapy often requires frequent in-person visits, making it costly (up to 2,000 USD per year per patient) and inaccessible for many patients, with over 50% of those in need not receiving adequate therapy [2, 3].

Mobile health (mHealth) platforms have emerged as promising solutions to address these accessibility challenges, demonstrating significant potential in improving healthcare delivery and patient engagement for chronic conditions [4, 5]. Recent studies have shown that smartphone-based rehabilitation interventions can achieve clinically meaningful outcomes, with mobile apps showing convincing effects for chronic disease management when properly implemented with evidence-based care pathways [6, 7]. Co-design studies specifically for knee osteoarthritis have established functional requirements for mobile platforms, including symptom tracking, visual progress monitoring, goal setting, and personalized exercise planning [8].

Current technological approaches to exercise assessment focus on binary movement tracking without assessing execution quality or providing detailed performance feedback, typically registering whether a movement occurred but unable to identify insufficient or incorrect movements [9, 10]. Research in osteoarthritis rehabilitation shows exercise therapy significantly reduces pain (effect size 0.56) and improves function, with benefits peaking at 2-month intervals [11, 12]. Meta-analyses demonstrate that patients with higher baseline pain and poorer physical function benefit more from structured therapeutic exercise, emphasizing personalized treatment importance [13]. Contemporary studies show 84% of OA research focuses on knee osteoarthritis with exercise therapy as primary intervention, with trends toward remote delivery and technology-supported care [14].

Recent deep learning models show significant potential in human activity recognition and rehabilitation monitoring. LSTM networks have emerged as powerful tools for analyzing temporal patterns in sequential data, particularly for human motion and therapeutic exercises. Studies successfully applied LSTMs to classify and segment rehabilitation exercises using wearable sensors and video-based motion capture, achieving high accuracy in recognizing movement phases and identifying repetitions [15–17]. LSTMs have also been used for predictive modeling of human motion, including joint trajectory estimation and movement intention anticipation, demonstrating utility in rehabilitation and assistive technologies [18, 19].

Hybrid CNN-LSTM architectures have gained popularity in human activity recognition by combining CNNs' spatial feature extraction with LSTMs' temporal modeling. These models have been successfully applied to recognize lower-limb rehabilitation exercises and traditional therapies like Baduanjin, outperforming conventional geometric feature-based methods [9, 16]. CNN-LSTM models have also generated reference trajectories for lower-limb exoskeletons using healthy motion data, advancing rehabilitation robotics and personalized exercise monitoring [17, 20].

Recent AI advances show significant potential for enhancing rehabilitation through precision interventions. Machine learning approaches achieve over 85% accuracy in estimating clinical scores from wearable sensors, enabling precise motor recovery tracking [21]. Network meta-analyses demonstrate that AI-assisted interventions, particularly with virtual reality and intelligent feedback, produce superior functional improvements compared to conventional therapy [22, 23]. Reviews document AI rehabilitation evolution from early concepts to current personalized and telerehabilitation implementations [24].

Parallel advances in natural language processing introduce new opportunities for personalized rehabilitation guidance. Large Language Models with chain-of-thought (CoT) prompting demonstrate remarkable capabilities in healthcare reasoning and decision-making. CoT variants include Iter-CoT for enhanced reasoning through error correction [25], Auto-CoT for improved generalization via diverse question sampling [26], and Coarse-to-Fine CoT for structured decision-making through gradual reasoning refinement [27].

While CoT prompting has been applied in question answering, education, and medical reasoning, its potential for patient-specific rehabilitation feedback remains largely unexplored. Recent studies show LLMs can generate accurate, contextually relevant feedback in physiotherapy education, suggesting applicability in direct patient guidance [28, 29]. However, challenges remain in ensuring LLM-generated feedback is clinically valid, interpretable, and individually tailored. This study integrates CoT prompting into a compre-

hensive AI-driven rehabilitation system to provide automated, physiotherapist-like feedback based on objective kinematic performance values [30-32].

Wearable sensor technology has become increasingly sophisticated in rehabilitation monitoring, with inertial measurement units (IMUs), electromyography (EMG), and flexible sensors enabling comprehensive movement analysis and real-time feedback [33]. Deep learning frameworks specifically designed for movement classification using wearable sensors have demonstrated the feasibility of personalized rehabilitation progress monitoring through objective kinematic assessment [34]. User evaluation studies of wearable sensor-based exercise biofeedback systems report high system usability scores (90.8/100) and adherence rates approaching 80%, indicating strong clinical acceptability [35].

To address these limitations, this paper presents DeepTherapy, an AI-driven rehabilitation system that leverages deep learning-based kinematic analysis and LLM-assisted exercise recommendations to provide personalized rehabilitation that is more affordable and widely accessible. The system utilizes sequence-based deep learning models to automatically segment exercise repetitions and compute biomechanical quality scores, capturing precise movement patterns and highlighting deviations from therapeutic standards. Additionally, it employs chain-of-thought reasoning in LLMs to interpret exercise performance data, reason like a physiotherapist, and provide targeted corrective feedback.

The primary contributions of this work include the development and optimization of a deep learning-based framework that automatically segments exercise repetitions and computes biomechanical quality scores; the design of a feedback system that utilizes chain-of-thought prompting in LLMs to interpret exercise performance data and provide targeted corrective feedback; and the implementation of a mobile application that integrates exercise analysis and AI-driven feedback to assist patients and therapists in monitoring rehabilitation progress.

METHODS

The overall architecture of the proposed DeepTherapy is illustrated in Fig. 1. The system consists of four main components: data collection and preparation, exercise phase classification, quality scoring, and personalized feedback generation. Initially, exercise videos captured via smartphone undergo pose estimation to extract frame-level skeletal keypoints, which are then preprocessed and formatted into structured format containing normalized coordinate vectors. These features serve as input to a deep learning model that

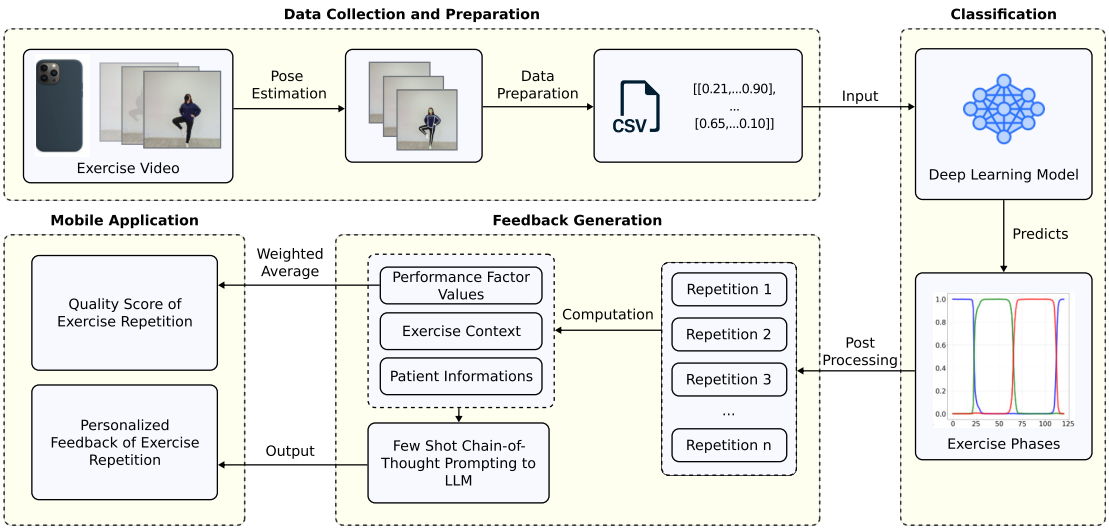

Fig. 1. System architecture of the DeepTherapy platform. Overview diagram illustrating the major components of the system, including data collection and preparation, exercise phase classification, quality scoring, and personalized feedback generation.

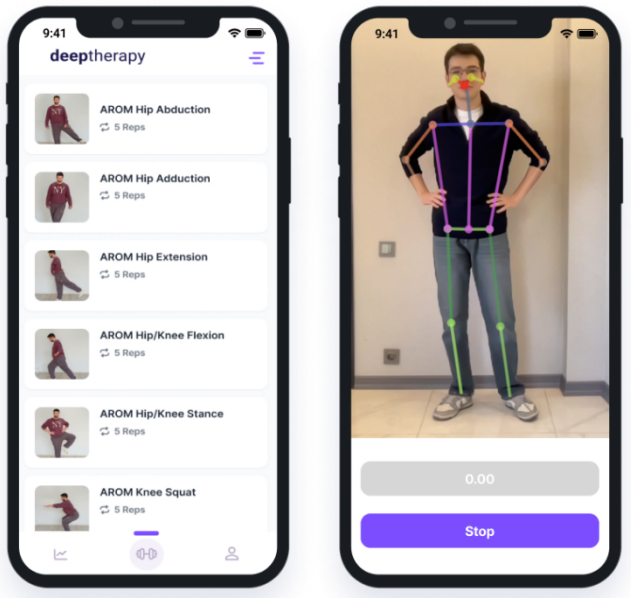

Fig. 2. DeepTherapy mobile platform interface. Screenshots of the mobile application showing the exercise selection screen and the real-time pose tracking interface during exercise performance.

analyzes temporal patterns to classify exercise phases and segment individual repetitions. For each segmented repetition, a continuous quality score is computed based on multiple performance factors derived from biomechanical analysis. Finally, a feedback generation module utilizes LLMs, guided by CoT prompting strategies, to translate the analysis into personalized corrective feedback for patients. Fig. 2 shows the mobile interface of the DeepTherapy platform.

The implementation of digital health solutions for chronic disease management has shown strong evidence for effectiveness when incorporating proper evaluation frameworks and technology adoption strategies [36]. Recent research demonstrates that digital health interventions achieve significant cost-effectiveness benefits for behavioral changes among patients with chronic conditions, particularly when combined with evidence-based intervention protocols [37]. Usability evaluation studies of digital platforms for orthopedic rehabilitation show high patient engagement rates and satisfaction scores, validating the clinical acceptability of mobile-based rehabilitation systems [38].

Data Collection

As there was no publicly available dataset specifically designed for hip and knee OA rehabilitation exercises, a new dataset was created to support the development and evaluation of deep learning models. A team of specialized physiotherapists designed an exercise protocol consisting of 10 therapeutic exercises targeting individuals with hip and knee osteoarthritis (Table 1). Several of these exercises were performed separately on each side of the body, leading to a total of 17 distinct activities.

In the data collection phase, a total of 20 healthy adult volunteers (10 males and 10 females) aged 21-26 years (mean age: 23 years) were participated. Participants were selected based on their competency in exercise performance, ensuring proper execution and minimizing the risk of injury. All participants provided informed consent before the data collection phase was initiated.

Each participant performed all 17 activities under supervision, completing five consecutive repetitions of each exercise at a controlled, slow speed to reflect the physical capabilities of OA patients. Recordings were conducted in a controlled indoor environment specifically arranged to optimize motion-capture quality (Fig. 3). Walls were uniformly colored and plain to minimize background interference, and lighting was carefully managed to be bright yet diffused, effectively eliminating shadows that could negatively affect data quality. A multi-camera setup was used to comprehen-

Table 1. Exercises included in protocol

#	Exercise name	Side
01	AROM Hip Abduction	Per Side
02	AROM Hip Adduction	Per Side
03	AROM Hip Extension	Per Side
04	AROM Hip/Knee Flexion Forward Lunge with Kegel	Per Side
05	AROM Knee Squat	Bilateral
06	Dynamic Hip Flexion (Knee to Chest)	Per Side
07	Hamstring Stretch (Standing, 90 Degrees)	Bilateral
08	Dynamic Knee Lunge Backward Diagonal	Per Side
09	Hip Adductor Stretch (Standing)	Bilateral
10	AROM Hip/Knee Stance	Per Side

AROM=Active Range of Motion

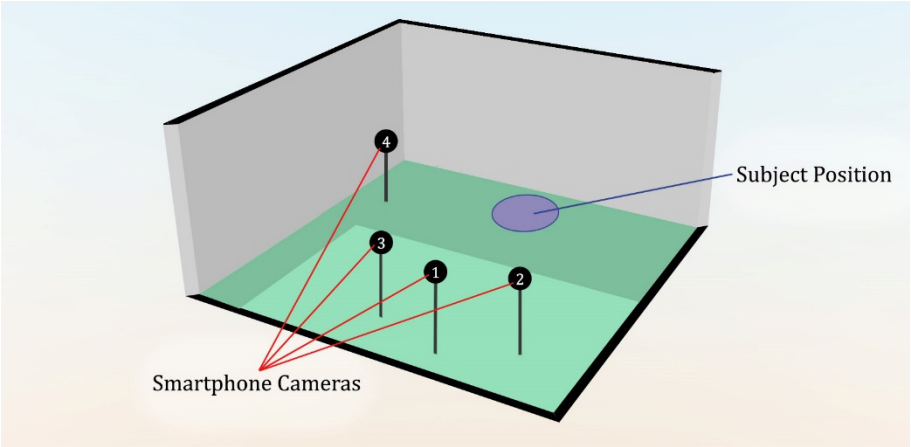

Fig. 3. Illustrated recording environment. Visualization of the controlled indoor environment used during data collection, featuring a multi-camera smartphone configuration.

sively capture the participants' movements from different perspectives. Four identical smartphones (1080p resolution at 30 frames per second) mounted on stable tripods captured videos simultaneously from multiple angles: frontal (90°), slightly left (105°), slightly right (75°), and lateral (0°).

In total, the dataset includes video recordings from 20 participants, each performing 17 activities with five repetitions, recorded from four camera views, amounting to 1,360 individual video samples (20 participants × 17 activities × 4 views).

Data Preparation

Following the data collection, videos underwent an initial processing phase involving frame extraction and resizing the original recording frame rate of 30 frames per second. This resulting in approximately 800,000 frames. Domain expert physiotherapists manually all frames and assigned numeric labels corresponding to specific phases of each exercise. Exercise phases were defined as clinically significant segments of movement, such as resting, execution, and recovery phases, essential for segmenting individual repetitions within an exercise. Labeling these phases provided ground-truth data necessary for training and evaluating the performance of the deep learning model.

Two principal approaches exist for extracting motion-related features from video data: deriving latent spatiotemporal embeddings directly from raw frames using CNNs, and extracting interpretable kinematic features from pose estimation outputs. While CNN-based feature extraction can capture spatial and motion patterns from video, such representations are often high-dimensional, opaque, and computationally intensive—posing challenges for real-time, on-device deployment in physiotherapy settings. In contrast, pose estimation-based pipelines yield structured, interpretable features such as joint coordinates, angles, and distances that are directly relevant to biomechanical analysis. This approach not only supports physiotherapy-specific domain reasoning but also enables downstream integration with LLMs. Accordingly, a pose-based feature extraction strategy was adopted in this study to ensure both clinical interpretability and compatibility with mobile implementation constraints.

Following frame labeling, pose estimation was applied to each frame using MediaPipe, a widely adopted framework featuring a lightweight architecture optimized for real-time processing on mobile devices. MediaPipe utilizes a deep neural network-based pipeline to detect 33 landmarks on the human body in three dimensions (3D) along with visibility scores. In this study, the (x, y) pixel coordinates and visibility scores of selected landmarks were extracted from each frame. As the dataset was collected in a controlled environment with stable lighting and camera positioning, no temporal smoothing or denoising filters were necessary and were therefore not applied to the pose landmarks. In addition to raw landmark coordinates, further kinematic features were computed to capture the dynamics of exercise performance. Specifically, angle features were calculated to describe the relative

orientation between connected body segments. For three $A(x_a, y_a)$, $B(x_b, y_b)$, $C(x_c, y_c)$ points and the angle θ at B was derived using the dot product and arccos function. The cosine of the angle is calculated as shown in Equation (1):

$$cos(\theta) = \frac{(x_a - x_b)(x_c - x_b) + (y_a - y_b)(y_c - y_b)}{\sqrt{(x_a - x_b)^2 + (y_a - y_b)^2} \times \sqrt{(x_c - x_b)^2 + (y_c - y_b)^2}} \quad (1)$$

The angle θ was then obtained by applying the arccos function and converting from radians to degrees, as shown in Equation (2):

$$\theta = \arccos(\cos(\theta)) \times \frac{180}{\pi} \quad (2)$$

Additionally, the spatial relationship between two landmarks $P_1(x_1, y_1)$ and $P_2(x_2, y2_a)$ was quantified using the Euclidean distance formula, as presented in Equation (3):

$$\text{Distance} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3)$$

To standardize the extracted kinematic features and ensure invariance to participant location and body size, spatial normalization was implemented through pose centering and scale normalization. First, all landmark coordinates were translated relative to the mid-hip landmark, which served as the origin, resulting in a body-centric coordinate system invariant to the participant's position within the video frame. Subsequently, coordinates were scaled using the shoulder width as a reference metric, defined as the Euclidean distance between the left and right shoulder landmarks. This normalization procedure ensured consistency and comparability of spatial features across participants, regardless of individual body dimensions and camera distances.

Model Development

A sequence-based deep learning model was developed to analyze temporal patterns within each of the 10 therapeutic exercises, individually trained using exercise-specific data. The selected architecture was based on a LSTM network, chosen for its proven effectiveness in capturing sequential dependencies inherent in human motion data. Each model accepted input data structured as sequences composed of consecutive frame-level feature vectors, with each vector containing normalized landmark coordinates, angle features, distances, and visibility scores. Thus, the input shape of each sequence was represented as $(n_{window} \times n_{features})$, where n_window denotes the number of consecutive frames per sequence, and n_features represents the total number of extracted kinematic features per frame.

The LSTM-based architecture included two stacked LSTM layers with 128 and 64 hidden units, respectively. Each LSTM layer was followed by batch normalization, to enhance training stability, and dropout layers with a dropout rate of 0.3, to prevent overfitting. Following these layers, a dense layer with 32 units and Rectified Linear Unit (ReLU) activation was included, introducing nonlinearity into the model. The final output layer comprised three neurons with softmax activation, providing probabilistic predictions of the exercise phase label for each frame. Fig. 4 illustrates the detailed structure of the developed LSTM model.

For model training, frame-level feature vectors and their corresponding phase labels were used to construct input sequences using a sliding window approach with a fixed sequence length of 16 frames and a stride of 1, resulting in 15-frame overlap between

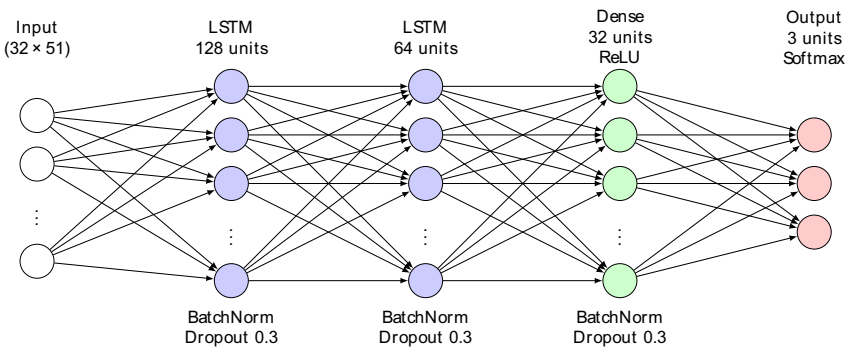

Fig. 4. Deep learning model architecture. Diagram of the LSTM-based neural network, including the sequence input layer, two LSTM layers with batch normalization and dropout, a dense layer, and a softmax output layer.

consecutive sequences. Each sequence was assigned the phase label of its final frame. To improve the model's ability to recognize and generalize across phase transitions, the overlapping windows often included segments from two or more consecutive phases (e.g., rest to eccentric, or eccentric to concentric). This design reflects the natural continuity of human motion and enables the LSTM to learn transitional dynamics rather than isolated steady-state segments. While each sequence was labeled by its final frame, the model leveraged the full temporal context to infer phase boundaries and intra-phase variations.

The dataset was partitioned into training and validation sets using participant-based splitting to ensure that data from the same participants did not appear in both sets, thus preventing data leakage and ensuring unbiased evaluation. The models were trained using the Adam optimizer with a learning rate of 0.001 and sparse categorical cross-entropy loss function, suitable for multiclass classification tasks. Training was conducted over a maximum of 50 epochs with a batch size of 32. To prevent overfitting and improve generalization, early stopping was applied by monitoring the validation loss with a patience of 10 epochs, and model checkpoints were saved based on the best validation accuracy. Dropout layers and batch normalization were incorporated within the model architecture to further enhance training stability and performance.

Scoring and Feedback System Development

Predicted phase sequences obtained from the deep learning model were post-processed to improve temporal consistency and reduce misclassification noise. A rule-based noise filtration method was applied, incorporating prediction confidence scores and temporal smoothing. Specifically, short-duration phase predictions with low confidence were corrected based on the surrounding phase context, and brief, isolated state transitions were suppressed to eliminate spurious predictions. Following the filtering process, exercise repetitions were segmented by detecting specific phase transition patterns corresponding to the expected sequence of resting, eccentric, and concentric phases. Additional criteria, such as a minimum duration threshold and valid phase order, were enforced to ensure that only complete and clinically meaningful repetitions were identified.

Quantitative performance factors were computed for each segmented exercise repetition to provide an objective assessment of exercise quality. The quality score of each repetition was calculated based on multiple performance factors relevant to the execution of the exercise. These factors included the correctness of the detected phase sequence, the temporal consistency of phase durations relative to expected patterns, the average confidence scores of model predictions within the repetition, and exercise-specific kinematic indicators such as range of motion, movement smoothness, and symmetry between sides. Each factor was normalized and assigned a predefined weight reflecting its clinical importance. The final quality score for each repetition was computed by aggregating these weighted factors, as shown in Equation (4), yielding a continuous score that captured overall movement quality of the performed exercise.

$$S = \sum_{i=1}^{N} w_i \cdot f_i \qquad (4)$$

Where S represents the final quality score, f_i denotes the normalized value of the i-th performance fac-

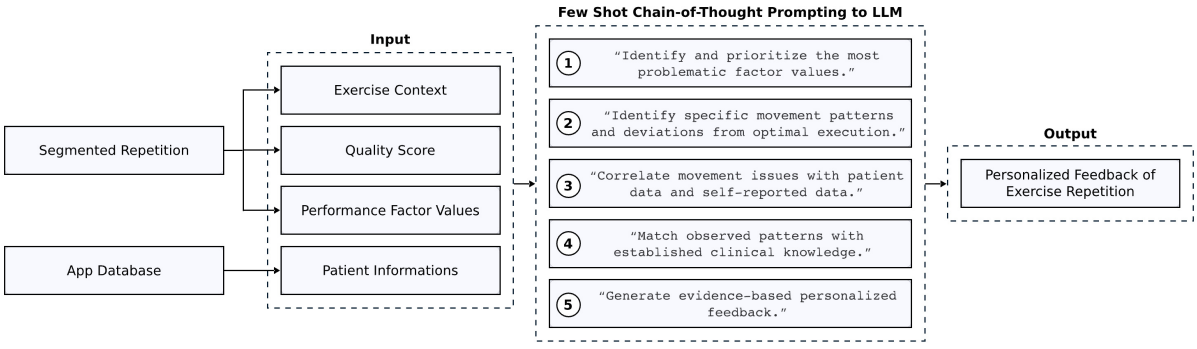

Fig. 5. Feedback generation process. Schematic showing how segmented exercise data and quality metrics are processed using LLMs with chain-of-thought prompting to generate personalized feedback.

tor, w_i is the predefined weight assigned to each factor, and N is the total number of factors considered in the scoring system.

A feedback generation system was developed to translate the computed quality score and performance factors into clinically meaningful feedback. This system utilized LLMs to automatically generate feedback statements based on the quantitative performance evaluation of each segmented repetition. To guide the feedback generation process, a structured prompt template was designed using example feedback provided by experienced physiotherapists. These examples were incorporated into a CoT prompting strategy, which explicitly structured the reasoning process of the LLM to improve the relevance and clinical accuracy of the generated feedback.

The input to the feedback system consisted of segmented repetition data, including the computed quality score, performance factor values, exercise context, and additional participant information such as self-reported feedback (perceived pain and discomfort) and demographic details. This information was organized into a structured prompt following a few-shot CoT template. The template was designed to encourage step-by-step reasoning by instructing the LLM. Example prompts and corresponding expert-written feedback were included to further guide the LLM's response. Two LLMs were evaluated within this system: OpenAI's GPT-4o and Anthropic's Claude 3.5 Sonnet. Both models processed the same CoT-based prompt template to ensure consistency in feedback generation. Fig. 5 illustrates the structure of the feedback generation system and the designed CoT prompt.

Mobile Application Development

A mobile application was developed to facilitate exercise assessment and feedback delivery. The application was implemented using the Flutter framework to ensure cross-platform compatibility on both iOS and Android devices. The system integrated the core components of repetition segmentation, quality scoring, and feedback generation within an interactive user interface. The application included two user modes: a patient interface for recording rehabilitation exercises and receiving feedback, and a therapist portal for monitoring patient performance and adjusting rehabilitation programs. Exercise data were processed after session completion, and feedback was provided to users based on the analysis of their recorded performance.

To enhance reliability during real-time inference, the mobile application ensures that feedback and scoring are performed only when all required landmarks in each frame exceed a predefined visibility threshold and a complete sequence of 16 valid frames is available. This constraint helps mitigate issues related to occlusions, poor lighting, or partial views during exercise performance without introducing additional filtering or smoothing.

Evaluation

The proposed system was evaluated through a multi-stage assessment process to examine the reliability of the quality scoring mechanism, the clinical relevance of the generated feedback, and the classification performance of the deep learning models. For the evaluation of the scoring system, two volunteer patients diagnosed with hip or knee OA performed the prescribed rehabilitation exercises. Each exercise repetition was assessed by the proposed system and independently scored by seven expert physiotherapists, each with over ten years of clinical experience. Physiotherapists assigned quality scores on a scale of 1 to 10 based on their expert assessment of exercise execution. The agreement between the system-generated scores and the physiotherapist ratings was evaluated using the Intraclass Correlation Coefficient (ICC) [39] to measure scoring reliability and consistency.

The feedback generation component was evaluated through expert review. For each exercise repetition, physiotherapists were presented with two feedback statements generated by the two LLMs (OpenAI's GPT-4o and Anthropic's Claude 3.5 Sonnet). Physiotherapists selected their preferred feedback and subsequently rated the selected output based on three criteria: clinical relevance, clarity, and helpfulness, using a five-point Likert scale.

Furthermore, the classification performance of the deep learning models for exercise phase recognition was evaluated on the validation set. Model performance was assessed using standard classification metrics, including accuracy, precision, recall, and F1-score. Fig. 6 illustrates the overall evaluation process, including the assessment of system-generated quality scores and LLM-generated feedback.

The study was reviewed and approved by the Bursa Uludağ University Health Research Ethics

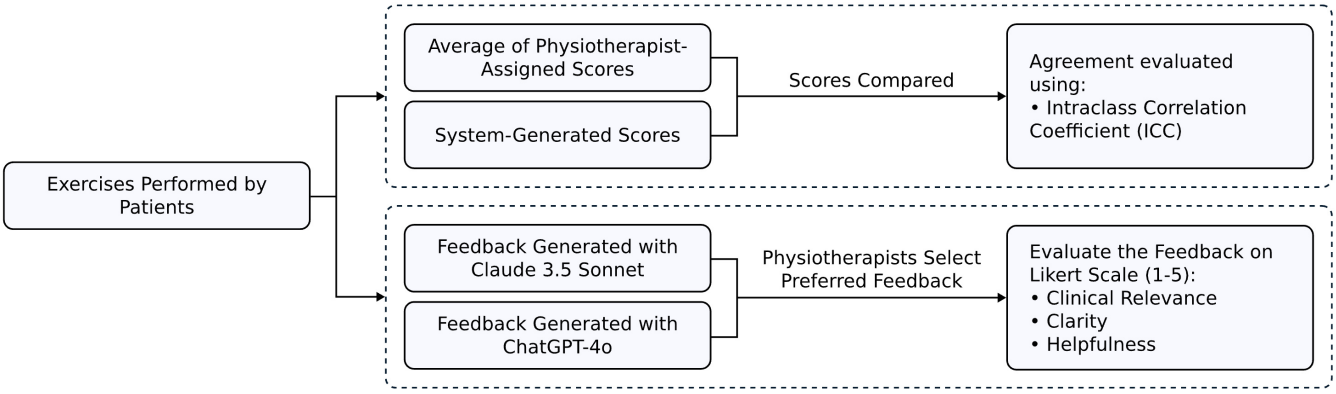

Fig. 6. Evaluation methodology. Diagram summarizing the evaluation workflow, including expert scoring comparison and LLM feedback review.

Committee. All participants provided written informed consent prior to participation, including consent for the use of anonymized video recordings and evaluation data for research and publication purposes.

RESULTS

Deep Learning Model Performance

The deep learning models demonstrated varying levels of performance, with the top model reaching 97.8% accuracy while the baseline achieved 96.1%. Figures 7 and 8 show the training and validation accuracy curves over epochs for the baseline and top models respectively, illustrating their convergence behavior and learning dynamics.

Detailed performance metrics for the baseline and top models are presented in Table 2, which includes precision, recall, and F1 scores for each phase classification. Confusion matrices and performance visualizations are shown in Figures 9 and 10 for the baseline and top models respectively. The baseline model achieved F1 scores of 0.96, 0.94, and 0.92 for rest, eccentric, and concentric phases respectively. The top model demonstrated improved performance with F1 scores of 0.99, 0.98, and 0.98 for the three phases.

Agreement Between Physiotherapist and System Scores

The reliability of the proposed scoring system was evaluated by comparing the system-generated quality scores with the scores assigned by expert physiotherapists. The Intraclass Correlation Coefficient (ICC) [39] between the two sets of scores was calculated as

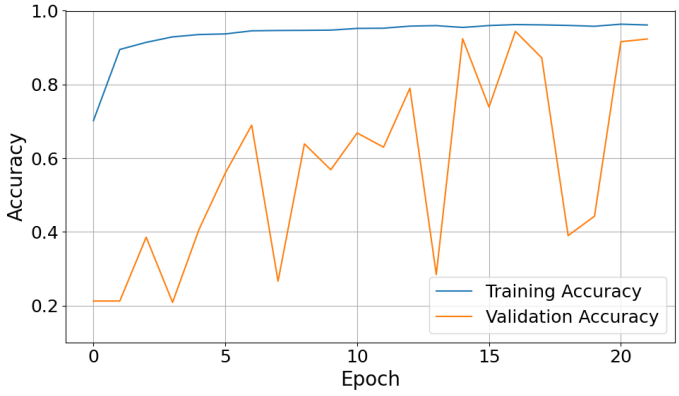

Fig. 7. Baseline model accuracy curves. Training and validation accuracy over epochs for the baseline model, demonstrating learning progression during training.

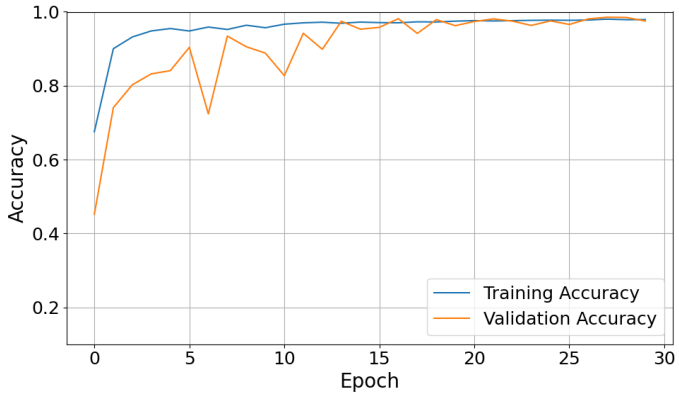

Fig. 8. Top model accuracy curves. Training and validation accuracy curves for the top-performing model, demonstrating learning progression during training.

Table 2. Performance metrics for baseline and top models

Model	Phase	Precision	Recall	F1 Score
Baseline	0	0.98	0.95	0.96
	1	0.92	0.96	0.94
	2	0.89	0.95	0.92
Top	0	0.99	0.99	0.99
	1	0.98	0.98	0.98
	2	0.98	0.97	0.98

0.85, indicating excellent agreement and inter-rater reliability. Figure 11 illustrates the correlation between system-assigned scores and the average physiotherapist scores. The data points closely follow the line of perfect agreement (1:1 line), demonstrating the system's ability to replicate expert evaluation of exercise quality.

Evaluation of LLM-Generated Feedback

The feedback generated by the two LLMs was evaluated based on physiotherapist preferences and three specific evaluation metrics: clinical relevance, clarity, and helpfulness. Physiotherapists slightly favored the outputs of Claude 3.5 Sonnet, selecting them in 52.4% of cases, compared to 47.6% for ChatGPT-4o. The distribution of model preferences is presented in Table 3.

However, as summarized in Table 4, ChatGPT-4o outperformed Claude 3.5 Sonnet across all evaluation criteria. ChatGPT-4o achieved higher average scores in clinical relevance (4.57/5), clarity (4.71/5), and helpfulness (4.52/5), compared to Claude 3.5 Sonnet, which received average scores of 4.13/5, 4.38/5, and 4.19/5, respectively. These results suggest that although Claude 3.5 Sonnet was marginally preferred in overall selection, the feedback produced by ChatGPT-4o was rated higher in clinical quality and clarity by expert physiotherapists.

DISCUSSION

The high ICC (0.85) between system-generated scores and professional physiotherapist evaluations demonstrates the system's capability to assess exercise performance with expert-level accuracy. This strong agreement validates the approach of integrating deep

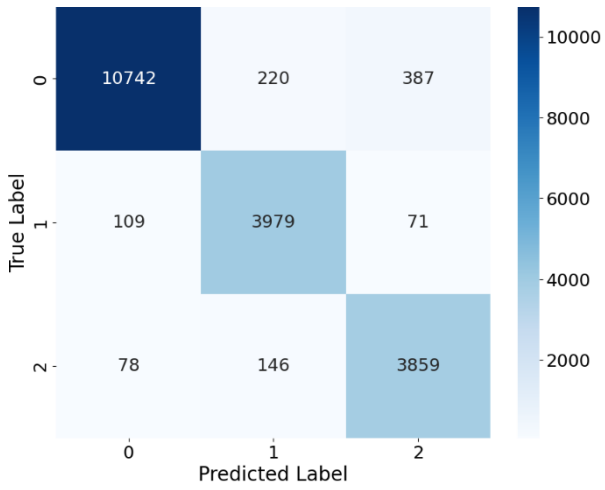

Fig. 9. Baseline model confusion matrix. Confusion matrix displaying classification performance of the baseline model across rest, eccentric, and concentric phases.

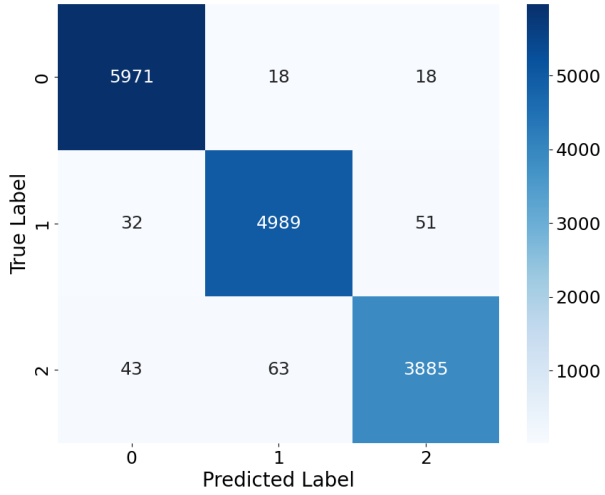

Fig. 10. Top model confusion matrix. Confusion matrix displaying classification performance of the top model across rest, eccentric, and concentric phases.

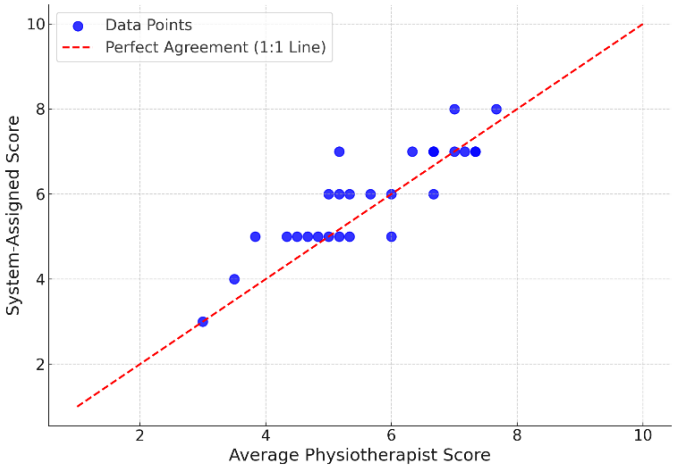

Fig. 11. Score correlation between system and physiotherapists. Scatter plot showing agreement between system-assigned and physiotherapist-assigned quality scores, including a 1:1 reference line.

learning-based movement analysis with biomechanical quality scoring for rehabilitation assessment. Such correlation suggests that automated systems can effectively supplement human expertise in rehabilitation settings, potentially extending the reach of professional guidance to patients who might otherwise have limited access.

The comparative analysis of LLM feedback reveals interesting insights about the nature of effective rehabilitation communication. While physiotherapists slightly preferred Claude 3.5 Sonnet's feedback qualitatively, GPT-4o consistently achieved superior ratings in clarity, relevance, and helpfulness. This suggests that GPT-4o may provide more technically precise and instructionally clear feedback, while Claude's outputs might have qualities that are more subjectively appealing to practitioners, potentially in tone or presentation style. These findings highlight the importance of considering both technical accuracy and communication style when designing AI-driven feedback systems for healthcare applications.

The DeepTherapy system represents several significant advancements over existing technological approaches to rehabilitation. The system quantifies exercise quality through AI-generated quality scores, significantly advancing current binary tracking methods that only detect whether a movement was performed without assessing its quality. This capability allows for more nuanced tracking of patient progress and more targeted interventions, addressing a critical limitation in existing rehabilitation technologies.

Table 3. Physiotherapist preferences for LLM-generated feedback

LLM model	Preference rate (%)
Claude 3.5 Sonnet	52.4%
ChatGPT-4o	47.6%

LLM=large language model

By utilizing chain-of-thought prompting in LLMs to analyze performance data and deliver corrective feedback, the system provides guidance similar to that of professional physiotherapists, addressing the critical gap in personalized feedback provision. This approach moves beyond simple performance detection to incorporate clinical reasoning and personalized recommendations, bringing AI-assisted rehabilitation closer to the experience of working with a human therapist.

The mobile platform enhances the accessibility, affordability, and convenience of osteoarthritis rehabilitation compared to traditional methods, potentially increasing therapy adherence and effectiveness. By making quality rehabilitation guidance available through a widely accessible device, DeepTherapy could significantly expand the reach of effective rehabilitation services, particularly for patients in underserved areas or with mobility limitations that make regular clinic visits challenging.

Despite the promising results, several limitations should be addressed in future work. The current evaluation was conducted with a limited number of OA

Table 4. Evaluation metrics of LLM-generated feedback by physiotherapists

LLM model	Clinical relevance	Clarity	Helpfulness
Claude 3.5 Sonnet	4.13	4.38	4.29
ChatGPT-4o	4.57	4.71	4.50

LLM=large language model

patients (n=2) and physiotherapists (n=7), which may not fully represent the diversity of clinical presentations and professional perspectives. A larger-scale validation with a more diverse patient population and clinical expertise would strengthen the generalizability of the findings. Additionally, the system was developed and tested on a specific set of exercises for hip and knee OA, and may require adaptation for other conditions or exercise protocols.

Future research directions include expanding the system's capabilities by incorporating a broader range of exercises and increasing dataset diversity to improve model robustness across different patient populations and movement patterns. This expansion would enhance the system's utility across various rehabilitation contexts and patient needs. Implementing Retrieval-Augmented Generation (RAG) with open-source LLMs could enhance domain-specific feedback quality and enable greater scalability in various clinical settings, potentially reducing dependence on proprietary LLM services while maintaining or improving feedback quality.

Integrating EMG (Electromyography) data processing represents another promising direction to enhance the system's capabilities, leveraging multi-modal LLMs for more precise movement analysis and adaptive rehabilitation guidance based on muscle activation patterns. This additional data stream could provide deeper insights into movement quality and muscular engagement, further refining the assessment and feedback capabilities of the system.

Limitations

The current evaluation was conducted with a limited number of OA patients and physiotherapists, which may not fully represent the diversity of clinical presentations and professional perspectives. A larger-scale validation with a more diverse patient population and clinical expertise would strengthen the generalizability of the findings. Additionally, the system was developed and tested on a specific set of exercises for hip and knee OA, and may require adaptation for other conditions or exercise protocols. The dataset consisted of healthy volunteers (n=20) aged 21-26 years performing exercises in controlled laboratory conditions, which may not accurately reflect the movement patterns and capabilities of actual OA patients who typically present with pain, limited mobility, and varying degrees of functional impairment.

The current study lacks long-term efficacy assessment and real-world deployment validation beyond the controlled laboratory environment. The pose estimation approach, while suitable for controlled conditions, may face accuracy challenges in diverse lighting conditions, camera angles, and backgrounds typical of home-based rehabilitation settings. Furthermore, the LLM-generated feedback, despite showing clinical relevance, requires continuous validation for factual accuracy and potential bias, particularly when applied to diverse patient populations with varying cultural backgrounds and health literacy levels. The mobile application's dependence on stable internet connectivity and smartphone compatibility may limit accessibility for elderly patients or those with limited digital literacy, potentially exacerbating healthcare disparities in rehabilitation access.

CONCLUSION

DeepTherapy effectively addresses the key limitations identified in current rehabilitation technologies by registering not only whether a movement was performed but also its quality; identifying insufficient or incorrect movements through detailed biomechanical analysis; and providing detailed guidance on how to improve technique through LLM-generated personalized feedback.

The system demonstrates strong agreement with expert physiotherapist assessments and provides clinically relevant feedback that could potentially enhance the accessibility and effectiveness of rehabilitation for individuals with osteoarthritis. By combining deep learning-based movement analysis with LLM-generated feedback in a mobile platform, DeepTherapy represents a promising approach to addressing the global burden of osteoarthritis through technology-assisted rehabilitation, ultimately working toward more accessible, effective, and personalized care for millions of affected individuals worldwide.

Ethics Approval and Consent to Participate

The study was approved by the Bursa Uludag University Health Research Ethics Committee (Decision no: 2025-2/19; date: 22.01.2025). All participants provided written informed consent prior to participation, including consent for the use of anonymized video

recordings and evaluation data for research and publication purposes.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: TTB, MFA, TÇ, HÖ; Study Design: TTB, MFA, TÇ, HÖ; Supervision: TTB, MFA, TÇ, HÖ; Funding: TTB, MFA, TÇ, HÖ; Materials: MFA, SMG, BŞ, ÇS, LA, ÖÖ, TÇ; Data Collection and/or Processing: MFA, SMG, BŞ, ÇS, LA, ÖÖ, TÇ; Statistical Analysis and/or Data Interpretation: TTB, MFA, TÇ, HÖ; Literature Review: TTB, MFA, TÇ, HÖ; Manuscript Preparation: TTB, MFA; and Critical Review: TTB, MFA.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgment

The authors would like to thank the volunteers who participated in the data collection phase, as well as the expert physiotherapists who evaluated the system's performance. We also acknowledge the technical support provided by the Faculty of Health Sciences at Bursa Uludağ University.

Generative Artificial Intelligence Statement

The author(s) declare that artificial intelligence tools were used in accordance with academic ethical standards during the preparation of this manuscript. Large Language Models GPT-4o (OpenAI) and Claude 3.5 Sonnet (Anthropic) were utilized within the experimental framework of the study to generate physiotherapy feedback. Additionally, Trinka.ai was employed to assist with grammar correction and paraphrasing during the manuscript editing process. All AI tools were used under the full control and supervision of the authors. Their purpose was to support system evaluation, not to replace human expertise.

Editor's note

All statements made in this article are solely those of the author(s) and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. World Health Organization. Osteoarthritis. 2023 Jul 14. Available at: https://www.who.int/news-room/fact-sheets/detail/osteoarthritis.
2. Losina E, Paltiel AD, Weinstein AM, et al. Lifetime medical costs of knee osteoarthritis management in the United States: impact of extending indications for total knee arthroplasty. Arthritis Care Res (Hoboken). 2015;67(2):203-215. doi: 10.1002/acr.22412.
3. World Health Organization. Rehabilitation Fact sheet. WHO Newsroom. 2024 Apr 22. Available at: https://www.who.int/news-room/fact-sheets/detail/rehabilitation.
4. Mrklas KJ, Barber T, Campbell-Scherer D, et al. Co-Design in the Development of a Mobile Health App for the Management of Knee Osteoarthritis by Patients and Physicians: Qualitative Study. JMIR Mhealth Uhealth. 2020;8(7):e17893. doi: 10.2196/17893.
5. Dieter V, Janssen P, Krauss I. Efficacy of the mHealth-Based Exercise Intervention re.flex for Patients With Knee Osteoarthritis: Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth. 2024;12:e54356. doi: 10.2196/54356.
6. Beresford L, Norwood T. The Effect of Mobile Care Delivery on Clinically Meaningful Outcomes, Satisfaction, and Engagement Among Physical Therapy Patients: Observational Retrospective Study. JMIR Rehabil Assist Technol. 2022;9(1):e31349. doi: 10.2196/31349.
7. Sun S, Simonsson O, McGarvey S, Torous J, Goldberg SB. Mobile phone interventions to improve health outcomes among patients with chronic diseases: an umbrella review and evidence synthesis from 34 meta-analyses. Lancet Digit Health. 2024;6(11):e857-e870. doi: 10.1016/S2589-7500(24)00119-5.
8. Goh SL, Persson MSM, Stocks J, et al. Efficacy and potential determinants of exercise therapy in knee and hip osteoarthritis: a systematic review and meta-analysis. Ann Phys Rehabil Med. 2019;62(5):356-365. doi: 10.1016/j.rehab.2019.04.006.
9. Bevilacqua A, Ciampi G, Argent R, Caulfield B, Kechadi T. Combining real-time segmentation and classification of rehabilitation exercises with LSTM networks and pointwise boosting. Proceedings of the AAAI Conference on Artificial Intelligence. 2020;34(8):13229-13234. doi: 10.1609/aaai.v34i08.7028.
10. Krutz P, Rehm M, Lang Z, Dix M, Patalas-Maliszewska J. Classification of Sports Exercises and Repetition Counting based on Inertial Measurement Data. Sensors and Electronic Instrumentation Advances: Proceedings of the 9th International Conference on Sensors and Electronic Instrumentation Advances, 20-22 September 2023 Funchal (Madeira Island), Portugal, 2023, pp. 35-39. https://sensorsportal.com/SEIA_2023/SEIA_2023_Proceedings.pdf.

11. Whittaker JL, Truong LK, Dhiman K, Beck C. Osteoarthritis year in review 2020: rehabilitation and outcomes. Osteoarthritis Cartilage. 2021;29(1):21-30. doi: 10.1016/j.joca.2020.10.005.
12. Runhaar J, Holden MA, Hattle M, et al; STEER OA Patient Advisory Group; OA Trial Bank Exercise Collaborative. Mechanisms of action of therapeutic exercise for knee and hip OA remain a black box phenomenon: an individual patient data mediation study with the OA Trial Bank. RMD Open. 2023;9(3):e003220. doi: 10.1136/rmdopen-2023-003220.
13. Adans-Dester C, Hankov N, O'Brien A, et al. Enabling precision rehabilitation interventions using wearable sensors and machine learning to track motor recovery. NPJ Digit Med. 2020;3(1):121. doi: 10.1038/s41746-020-00328-w.
14. Zhu Y, Wang C, Li J, Zeng L, Zhang P. Effect of different modalities of artificial intelligence rehabilitation techniques on patients with upper limb dysfunction after stroke-A network meta-analysis of randomized controlled trials. Front Neurol. 2023;14:1125172. doi: 10.3389/fneur.2023.1125172.
15. Chen J, Wang J, Yuan Q, Yang Z. CNN-LSTM Model for Recognizing Video-Recorded Actions Performed in a Traditional Chinese Exercise. IEEE J Transl Eng Health Med. 2023;11:351-359. doi: 10.1109/JTEHM.2023.3282245.
16. Rahman ZU, Ullah SI, Salam A, Rahman T, Khan I, Niazi B. Automated Detection of Rehabilitation Exercise by Stroke Patients Using 3-Layer CNN-LSTM Model. J Healthc Eng. 2022;2022:1563707. doi: 10.1155/2022/1563707.
17. Ren B, Zhang Z, Chi Z, Chen S. Motion Trajectories Prediction of Lower Limb Exoskeleton Based on Long Short-Term Memory (LSTM) Networks. Actuators 2022;11(3):73. doi: 10.3390/act11030073.
18. Görgülü YE, Tasdelen K. Human Activity Recognition and Temporal Action Localization Based on Depth Sensor Skeletal Data. 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), Istanbul, Turkey, 2020, pp. 1-5, doi: 10.1109/ASYU50717.2020.9259886.
19. He J, Guo Z, Shao Z, Zhao J, Dan G. An LSTM-Based Prediction Method for Lower Limb Intention Perception by Integrative Analysis of Kinect Visual Signal. J Healthc Eng. 2020;2020:8024789. doi: 10.1155/2020/8024789.
20. Liang F-Y, Zhong CH, Zhao X, et al. Online Adaptive and LSTM-Based Trajectory Generation of Lower Limb Exoskeletons for Stroke Rehabilitation. 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO), Kuala Lumpur, Malaysia, 2018, pp. 27-32. doi: 10.1109/ROBIO.2018.8664778.
21. J. Sun, Y. Gong, C. Lin, Y. Shen, J. Guo, and N. Duan, "Enhancing Chain-of-Thoughts Prompting with Iterative Bootstrapping in Large Language Models," arXiv.org, Apr. 2023, doi: 10.48550/arXiv.2304.11657.
22. Zhang Z, Zhang A, Li M, Smola AJ. Automatic Chain of Thought Prompting in Large Language Models. arXiv: 2210.03493, Oct 2022. doi: 10.48550/arXiv.2210.03493.
23. Nguyen HH, Liu Y, Zhang C, Zhang T, Yu PS. CoF-CoT: Enhancing Large Language Models with Coarse-to-Fine Chain-of-Thought Prompting for Multi-domain NLU Tasks. arXiv: 2310.14623, Oct 2023, doi: 10.48550/arxiv.2310.14623.
24. Villagrán I, Hernández R, Schuit G, et al. Implementing Artificial Intelligence in Physiotherapy Education: A Case Study on the Use of Large Language Models (LLM) to Enhance Feedback. IEEE Trans Learn Technol. 2024;17:2025-2036. doi: 10.1109/TLT.2024.3450210.
25. Nachane SS, Gramopadhye O, Chanda P, et al. Few shot chain-of-thought driven reasoning to prompt LLMs for open ended medical question answering. arXiv:2403.04890, Mar 2024. doi: 10.48550/arxiv.2403.04890.
26. Masikisiki B, Marivate V, Hlope Y. Investigating the Efficacy of Large Language Models in Reflective Assessment Methods through Chain of Thoughts Prompting. arXiv:2310.00272, Sep 2023, doi: 10.48550/arXiv.2310.00272.
27. Zou A, Zhang Z, Zhao H, Tang X. Meta-CoT: Generalizable Chain-of-Thought Prompting in Mixed-task Scenarios with Large Language Models. arXiv: 2310.06692, Oct 2023. doi: 10.48550/arxiv.2310.06692.
28. Cheng X, Li J, Zhao WX, Wen J-R. ChainLM: Empowering Large Language Models with Improved Chain-of-Thought Prompting. arXiv:2403.14312, Mar 2024. doi: 10.48550/arXiv.2403.14312.
29. Mahmoud H, Aljaldi F, El-Fiky A, et al. Artificial Intelligence machine learning and conventional physical therapy for upper limb outcome in patients with stroke: a systematic review and meta-analysis. Eur Rev Med Pharmacol Sci. 2023;27(11):4812-4827. doi: 10.26355/eurrev_202306_32598.
30. Senadheera I, Hettiarachchi P, Haslam B, et al. AI Applications in Adult Stroke Recovery and Rehabilitation: A Scoping Review Using AI. Sensors (Basel). 2024;24(20):6585. doi: 10.3390/s24206585.
31. Maceira-Elvira P, Popa T, Schmid AC, Hummel FC. Wearable technology in stroke rehabilitation: towards improved diagnosis and treatment of upper-limb motor impairment. J Neuroeng Rehabil. 2019;16(1):142. doi: 10.1186/s12984-019-0612-y.
32. Panwar M, Biswas D, Bajaj H, et al. Rehab-Net: Deep Learning Framework for Arm Movement Classification Using Wearable Sensors for Stroke Rehabilitation. IEEE Trans Biomed Eng. 2019;66(11):3026-3037. doi: 10.1109/TBME.2019.2899927.
33. Argent R, Slevin P, Bevilacqua A, Neligan M, Daly A, Caulfield B. Wearable Sensor-Based Exercise Biofeedback for Orthopaedic Rehabilitation: A Mixed Methods User Evaluation of a Prototype System. Sensors (Basel). 2019;19(2):432. doi: 10.3390/s19020432.
34. De Fazio R, Mastronardi VM, De Vittorio M, Visconti P. Wearable Sensors and Smart Devices to Monitor Rehabilitation Parameters and Sports Performance: An Overview. Sensors (Basel). 2023;23(4):1856. doi: 10.3390/s23041856.
35. Kristoffersson A, Linden M. A Systematic Review of Wearable Sensors for Monitoring Physical Activity. Sensors (Basel). 2022;22(2):573. doi: 10.3390/s22020573.
36. Pong C, Tseng RMWW, Tham YC, Lum E. Current Implementation of Digital Health in Chronic Disease Management: Scoping Review. J Med Internet Res. 2024;26:e53576. doi: 10.2196/53576.
37. Kyaw TL, Ng N, Theocharaki M, Wennberg P, Sahlen KG. Cost-effectiveness of Digital Tools for Behavior Change Interventions Among People With Chronic Diseases: Systematic Review. Interact J Med Res. 2023;12:e42396. doi: 10.2196/42396.
38. Papadopoulos P, Soflano M, Connolly T. A Digital Health Intervention Platform (Active and Independent Management System) to Enhance the Rehabilitation Experience for Orthopedic Joint Replacement Patients: Usability Evaluation Study. JMIR Hum Factors. 2024;11:e50430. doi: 10.2196/50430.
39. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63. doi: 10.1016/j.jcm.2016.02.012.

The European Research Journal 2025;11(6):1043-1056 Original Article

DOI: https://doi.org/10.18621/eurj.1701524 *Medical Physiology*

Tranilast protects from sepsis-induced acute kidney injury in rat via the STAT-3 signaling pathway

Mümin Alper Erdoğan[1], Arife Erdoğan[2], Oytun Erbaş[3]

[1]Department of Physiology, İzmir Katip Çelebi University, Faculty of Medicine, İzmir, Türkiye; [2]Department of Emergency Medicine, University of Health Sciences, İzmir Tepecik Training and Research Hospital, Izmir, Türkiye; [3]Biruni University, Faculty of Medicine, BAMER, İstanbul, Türkiye

ABSTRACT

Objectives: Sepsis-induced acute kidney injury (SI-AKI) is a major contributor to morbidity and mortality among critically ill patients. This experimental study evaluated the renoprotective effects of Tranilast, an anti-inflammatory and antifibrotic agent, in a rat model of polymicrobial sepsis, with a focus on modulation of the signal transducer and activator of transcription 3 (STAT-3) signaling pathway.

Methods: Thirty-six female Wistar albino rats were randomly assigned to three groups: Sham control, cecal ligation and puncture (CLP)+saline, and CLP+Tranilast (300 mg/kg/day). Sepsis was induced by CLP. Survival was monitored for five days. Biochemical parameters including plasma tumor necrosis factor alpha (TNF-α), neutrophil gelatinase-associated lipocalin (NGAL), heat shock protein 27 (HSP-27), malondialdehyde (MDA), blood urea nitrogen (BUN), and renal STAT-3 expression were assessed via ELISA and spectrophotometric assays. Histopathological evaluation of renal tissues was performed to assess tubular injury, inflammation, and hemorrhage. Results are expressed as mean±standard error of the mean (SEM).

Results: Tranilast significantly improved survival in septic rats (75% vs. 50% in CLP+saline), reduced plasma MDA and TNF-α levels, lowered BUN and NGAL concentrations, and suppressed renal STAT-3 expression ($P<0.05$). It also enhanced HSP-27 levels, suggesting activation of cytoprotective responses. Histological analysis demonstrated reduced tubular necrosis, luminal debris, inflammation, and hemorrhage in Tranilast-treated rats.

Conclusions: Tranilast provides significant renoprotection in SI-AKI by reducing oxidative stress, inflammation, and STAT-3 activity while enhancing cytoprotective mechanisms. These findings support its potential as an adjunctive therapeutic agent for managing sepsis-related organ injury.

Keywords: Sepsis, acute kidney injury, tranilast, STAT-3 pathway, oxidative stress, inflammation

Sepsis, a potentially fatal malfunction of organs caused by an uncontrolled immune response to infection, is still one of the leading causes of death and disability among severely sick people across the globe. Up to half of all septic patients may get acute kidney damage (AKI), which is a serious consequence of the infection that greatly raises death rates and healthcare expenditures [1, 2].

The pathophysiology of sepsis-induced AKI (SI-AKI) is multifactorial, involving systemic inflammation, hemodynamic instability, endothelial dysfunction, and oxidative stress. Unlike traditional

Received: May 18, 2025 **Accepted:** August 9, 2025 **Available Online:** August 17, 2025 **Published:** November 4, 2025

How to cite this article: Erdoğan MA, Erdoğan A, Erbaş O. Tranilast protects from sepsis-induced acute kidney injury in rat via the STAT-3 signaling pathway. Eur Res J. 2025;11(6):1043-1056. doi: 10.18621/eurj.1701524

Corresponding author: Mümin Alper Erdoğan, PhD., DVM, Assoc. Prof., Phone: +90 232 329 35 35, E-mail: alpero86@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

ischemic AKI, SI-AKI often presents with preserved or even increased renal blood flow, highlighting the role of inflammation and cellular injury rather than perfusion deficits [3]. To model this complex condition, we utilized the cecal ligation and puncture (CLP) technique, which is widely regarded as a clinically relevant method to induce polymicrobial sepsis in rodents and closely mimics the human pathophysiological course [4].

Key players in the progression of renal tubular injury in sepsis include proinflammatory cytokines like the tumor necrosis factor alpha (TNF-α) and reactive oxidative mediators like malondialdehyde (MDA) [5]. Heat shock protein 27 (HSP-27) and neutrophil gelatinase-associated lipocalin (NGAL) have both been shown to be early indicators of cellular stress responses and kidney damage, respectively, in a number of different experimental and clinical contexts [6, 7]. New data also points to signal transducer and activator of transcription 3 (STAT-3) as an important player in SI-AKI's inflammatory and apoptotic pathways. The phosphorylation of STAT-3 has been shown to contribute to renal cell injury by promoting proinflammatory gene expression and enhancing oxidative damage [8].

Tranilast (N-[3,4-dimethoxycinnamoyl]-anthranilic acid), originally developed as an antiallergic agent, has recently garnered attention for its anti-inflammatory, antifibrotic, and antioxidant properties. It is well-known that tranilast prevents the secretion of inflammatory cytokines and regulates the activity of immune cells, namely by suppressing the signaling pathways of nuclear factor-kappa B (NF-κB) and STAT-3 [9, 10]. Studies have shown that Tranilast ameliorates organ damage in various inflammatory models, including pulmonary fibrosis, myocardial ischemia-reperfusion injury, and diabetic nephropathy [11-13]. However, its protective role against sepsis-induced renal injury and the underlying mechanisms, particularly the involvement of STAT-3 modulation, remain largely unexplored.

Recognizing the critical need for effective interventions in SI-AKI, we aimed to investigate the therapeutic potential of Tranilast in this clinically relevant sepsis model.

We hypothesized that Tranilast would attenuate renal inflammation, oxidative stress, and tubular damage in septic rats, potentially via downregulation of the STAT-3 signaling pathway.

To test this hypothesis, we assessed histopathological and biochemical markers of kidney injury, including TNF-α, NGAL, HSP-27, blood urea nitrogen (BUN), MDA, and STAT-3 levels, in CLP-induced septic rats treated with Tranilast. Our results contribute to a better understanding of the mechanisms underlying SI-AKI and offer novel insights into the anti-inflammatory and cytoprotective effects of Tranilast in this context.

METHODS

Animals

A total of 36 female Wistar albino rats, aged 10-12 weeks and weighing between 200 and 250 grams, were used in this study. This study's research received approval from the Demiroğlu Science University Animals Ethical Board (Ethical number: 0825015107) and was done in compliance with the National Institutes of Health (U.S.A.). Guideline for the Care and Utilization of Experimental Animals. The animals used in the experiment were obtained from the Experimental Animal Laboratory at Demiroğlu Science University. The rats were housed in steel cages with 12-hour light/dark cycles, fed freely, and maintained in a temperature-controlled environment at 22±2 °C. Only female rats were used in this study to reduce variability due to hormonal cycles and behavioral differences commonly observed in males. While we acknowledge that sex-specific immune responses may exist, using one sex allowed for a more consistent evaluation of the intervention's effects.

Experimental procedures

To create a sepsis model, 36 female Wistar albino rats (10-12 weeks old, 200-250 g) were randomly assigned into three experimental groups (n=12 per group) as follows:

(1) Group 1: Sham-operated Control (Sham Control, n=12): Rats underwent laparotomy and bowel manipulation under sterile conditions, but without cecal ligation or puncture. They received oral gavage with vehicle solution only.

(2) Group 2: Sepsis+Saline (CLP+Saline, n=12): Rats underwent CLP to induce polymicrobial sepsis, followed by intraperitoneal administration of 0.9% NaCl at a dose of 30 mL/kg/day for 5 consecutive days.

(3) Group 3: Sepsis+Tranilast (CLP+Tranilast, n=12): Rats underwent CLP and received Tranilast at a dose of 300 mg/kg/day, administered orally once daily for 5 consecutive days. The dose of Tranilast (300 mg/kg/day, orally) was selected based on previous reports demonstrating efficacy in reducing inflammation and organ damage in rodent models [14]. Tranilast (purity >98%, Sigma-Aldrich, Cat# T5316) was dissolved in 0.9% saline and administered by oral gavage in a volume of 1 mL.

The sham group served as a true surgical control to distinguish the effects of surgical stress from those of sepsis. Tranilast or saline was administered once daily by oral gavage for 5 consecutive days post-CLP. All animals were monitored for survival over 5 days, after which biological sampling and tissue analyses were performed. The sample size in each group remained consistent throughout the experimental period, and the final number of animals included in the analyses is reported in the figure legends and table footnotes.

All volume replacement treatments were administered in similar amounts to Groups 2 and 3: 30 mL/kg % 0.9 NaCl once throughout the initial hour of the surgical process then 10 ml/kg/day % 0.9 NaCl i.p.) every 12 hours of the day. Intraperitoneal saline (30 mL/kg) was administered for fluid resuscitation after CLP, based on previously established models. Although this dose is close to the recommended upper limit, it was well tolerated by all animals without respiratory distress or mortality immediately following injection [15].

After five days, the trial was completed. Nine rats perished - three from the CLP and Tranilast group and six from the CLP and saline group. The animals were euthanized after the trial was complete so that blood samples could be taken for biochemical analysis and kidneys could be removed for histological analysis. Blood samples were collected into EDTA tubes and centrifuged at 3000 rpm for 15 minutes at 4°C. Plasma was aliquoted and stored at -80°C until biochemical analysis. The left kidney was fixed in 10% formalin for histopathological analysis, while the right kidney was snap-frozen in liquid nitrogen and used for ELISA. A total of 36 rats (n = 12 per group) were included at baseline. Survival and mortality were monitored daily, and final sample sizes for each group at each endpoint are specified in the Table 1 and Fig. 1. Due to CLP-induced mortality, final sample sizes were: Sham Control (n=12), CLP+Saline (n=6), CLP+Tranilast (n=9) for biochemical and histological analysis.

Procedure for Cecal Ligation and Puncture (CLP)

To expose the cecum and adjacent intestine, a 3 cm midline laparotomy was carried out under aseptic circumstances. Following its secure tying at its base beneath the iliocecal valve using a 3.0 silken suture, a 22-gauge needle was used to puncture the cecum. The cecum was then lightly squeezed to expel a little amount of waste from the puncture area. After returning the cecum to the peritoneal cavity, 4-0 polyglactin sutures were used to seal the laparotomy incision. The animals were allowed some downtime after surgery before being returned to their homes. The sham group of rats had just laparotomy in an aseptic environment: they did not undergo cecum piercing or ligation. Rats were considered septic in this model five hours after CLP [16].

Histopathological Analysis of Renal Tissue

Three days of preservation in a solution of 0.1 M phosphate-buffer sodium chloride (PBS) with 10% formaldehyde was used for the kidney after its extraction. Hematoxylin and eosin (H&E) staining was applied to 5 µm formalin-fixed kidney slices. All areas were pho-

Table 1. Survival of rats following CLP-induced sepsis with or without tranilast treatment

Group	Number of animals at baseline	Surviving animals-day 1	Surviving animals-day 2	Surviving animals-day 3	Surviving animals-day 4	Surviving animals-day 5
CLP + Saline	**12**	**12**	**10**	**10**	**8**	**6**
CLP + Tranilast	**12**	**12**	**11**	**11**	**10**	**9**

CLP=cecal ligation and puncture. Survival data recorded daily for 5 days post-CLP surgery in placebo and Tranilast-treated groups. Tranilast improved overall survival compared to saline-treated controls.

tographed using Olympus C-5050 camera models that were attached to Olympus BX51 microscope.

An unbiased observer conducted a morphological evaluation on ten microscopic fields per slice at a magnification of ×20 utilizing a computer-based image processing system (Image-Pro Express 1.4.5, Media Cybernetics, Inc. USA) without knowledge of the study group.

The degree of tubular dilatation, luminal necrotic debris, tubular epithelial necrosis, interstitial inflammation, and hemorrhage in kidney sections from each rat in each group were assessed semi-quantitatively using the following rating system: (a) score 0: 0 to 5 percent; (b) score 1: 6 to 20 percent; (c) score 2: 21 to 40 percent; (d) score 3: 41 to 60 percent; (e) score 4: 61 to 80 percent; and (f) score 5: 81 to 100 percent [15].

Plasma and Kidney Biochemical Measurements

Organs were quickly removed after decapitation and kept at 20°C until biochemical assays were completed. Before being centrifuged at 5000 g for 15 minutes, kidneys were homogenized using a glass homogenizer in five volumes of phosphate-buffered saline (PBS), which was five times the volume of the recovered tissue (pH 7.4). After collecting the supernatant, Bradford's technique was used to quantify the total protein content in the innerve homogenates using bovine serum albumin as a reference. Renal STAT-3 protein levels were measured using a rat-specific ELISA kit: Rat STAT-3 ELISA Kit, Cat. MBS703634, MyBioSource (San Diego, CA, USA), following the manufacturer's protocol.

Biochemistry Evaluation of BUN

An automated analysis method was used to determine the BUN concentrations spectrophotometrically. Concentrations of BUN were reported in mg/dL.

Assessment of Plasma Levels of TNF-a, NGAL, and HSP-27

Plasma levels of TNF-α, NGAL, and HSP-27 were measured using commercially available ELISA kits: Rat TNF α ELISA Kit, Cat. ab100784, Abcam (Cambridge, UK), Rat NGAL ELISA Kit, Cat. ab119602, Abcam (Cambridge, UK), Rat HSP 27 ELISA Kit, Cat. A2296, Antibodies.com (Cambridge, UK). All assays were performed according to the manufacturer's instructions.

Evaluation of Lipid Peroxidation

For quantification of MDA, a TBARS Assay Kit was used: TBARS (Thiobarbituric Acid Reactive Substances) Assay Kit, Cat. 10009055, Cayman Chemical (Ann Arbor, MI, USA). To put it briefly, the plasma samples were combined, incubated for 60 minutes at 100 °C, and then treated with TBARS reagent and trichloroacetic acid. After chilling briefly on ice, the

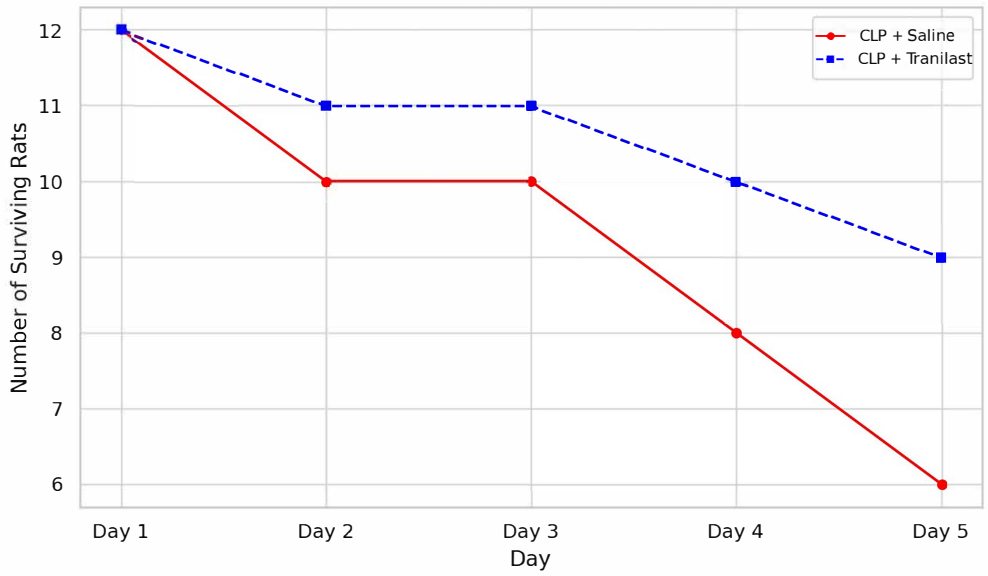

Fig. 1. Survival of rats over 5 days.

samples were centrifuged for 20 minutes at 3000 rpm, and the supernatant's absorbance value was measured at 535 nanometer. Calibration was done using tetraethoxypropane, and MDA levels were given in nM.

Statistical Analysis

Statistical analysis was performed using GraphPad Prism version 9.0 (GraphPad Software, San Diego, CA, USA). Mann-Whitney U was used as a non-parametric test for all quantitative data. The Student's-t test was used to assess the variations among the groups. Mean values together with the standard error of the median (SEM) are often used for data visualization. We considered it statistically significant if the value of p was lower than 0.05. The following notation was used to indicate statistical significance in Table 2 and corresponding figures: *P<0.01, **P<0.001 vs. Sham control group; #P<0.05, ##P<0.001 vs. CLP+Saline group. The following notation was used to indicate statistical significance in Table 3 and corresponding figures: *P<0.001 indicates a significant difference compared to the Sham control group; #P<0.01 and ##P<0.001 indicate significant differences compared to the CLP+Saline group.

RESULTS

Survival Analysis

Tranilast therapy considerably increased the five-day survival rate of rats who had sepsis caused by CLP (Table 1, Fig. 1). Between Day 1 and Day 5, the CLP+Saline (placebo) group's survival rate gradually decreased from 100% to 50% (6/12 rats). In contrast, rats receiving Tranilast showed improved survival outcomes, with 9 of 12 animals (75%) alive at the end of the observation period. These results suggest that Tranilast confers a survival advantage under septic conditions, potentially by mitigating systemic organ injury.

Survival data recorded daily for 5 days post-CLP surgery in placebo and Tranilast-treated groups. Tranilast improved overall survival compared to saline-treated controls.

Plasma and Renal Biochemical Parameters

The biochemical analysis of plasma and renal tissue samples revealed significant alterations in key markers of oxidative stress, inflammation, and renal function across experimental groups, as summarized in Table 2 and Fig. 2.

Comparison of plasma and kidney biochemical parameters among sham control, CLP+Saline (placebo), and CLP+Tranilast groups. Tranilast treatment significantly reduced levels of MDA, TNF-α, BUN, NGAL, and renal STAT-3 while increasing HSP-27 expression, suggesting attenuation of oxidative stress, inflammation, and kidney injury (P<0.001 vs. control; P<0.01 vs. Control; P<0.001 vs. CLP+Saline;P<0.05 vs. CLP+Saline).

Table 2. Biochemical markers of renal injury and inflammation in experimental groups

	Sham control	CLP and saline group	CLP and Tranilast group
Plasma MDA (nM)	45.3±2.2	124.9±3.8**	88.5±5.5##
Plasma TNF alfa (pg/ml)	15.2±4.5	101.6±2.7*	71.1±1.8#
Plasma BUN (mg/dl)	17.9±0.2	36.5±1.6*	29.7±0.9#
Plasma NGAL (ng/ml)	65.8±1.4	215.1±6.8**	145.7±5.5##
Plasma HSP-27 (ng/ml)	1.06±0.2	1.95±0.1*	3.07±0.4#
Kidney STAT-3 (pg/mg protein)	0.88±0.2	2.09±0.1**	1.12±0.3#

Data are shown as mean±standard error of the median. CLP=cecal ligation and puncture, MDA=malondialdehyde, TNF-α=tumor necrosis factor alpha, BUN=blood urea nitrogen, NGAL=neutrophil gelatinase-associated lipocalin, HSP-27=heat shock protein 27, STAT-3= signal transducer and activator of transcription 3

Comparison of plasma and kidney biochemical parameters among Sham control, CLP+Saline (placebo), and CLP+Tranilast groups. Tranilast treatment significantly reduced levels of MDA, TNF-α, BUN, NGAL, and renal STAT-3 while increasing HSP-27 expression, suggesting attenuation of oxidative stress, inflammation, and kidney injury.

*P<0.01 vs. control, **P< 0.001 vs. control, #P<0.05 vs. CLP+Saline ##P<0.001 vs. CLP+Saline.

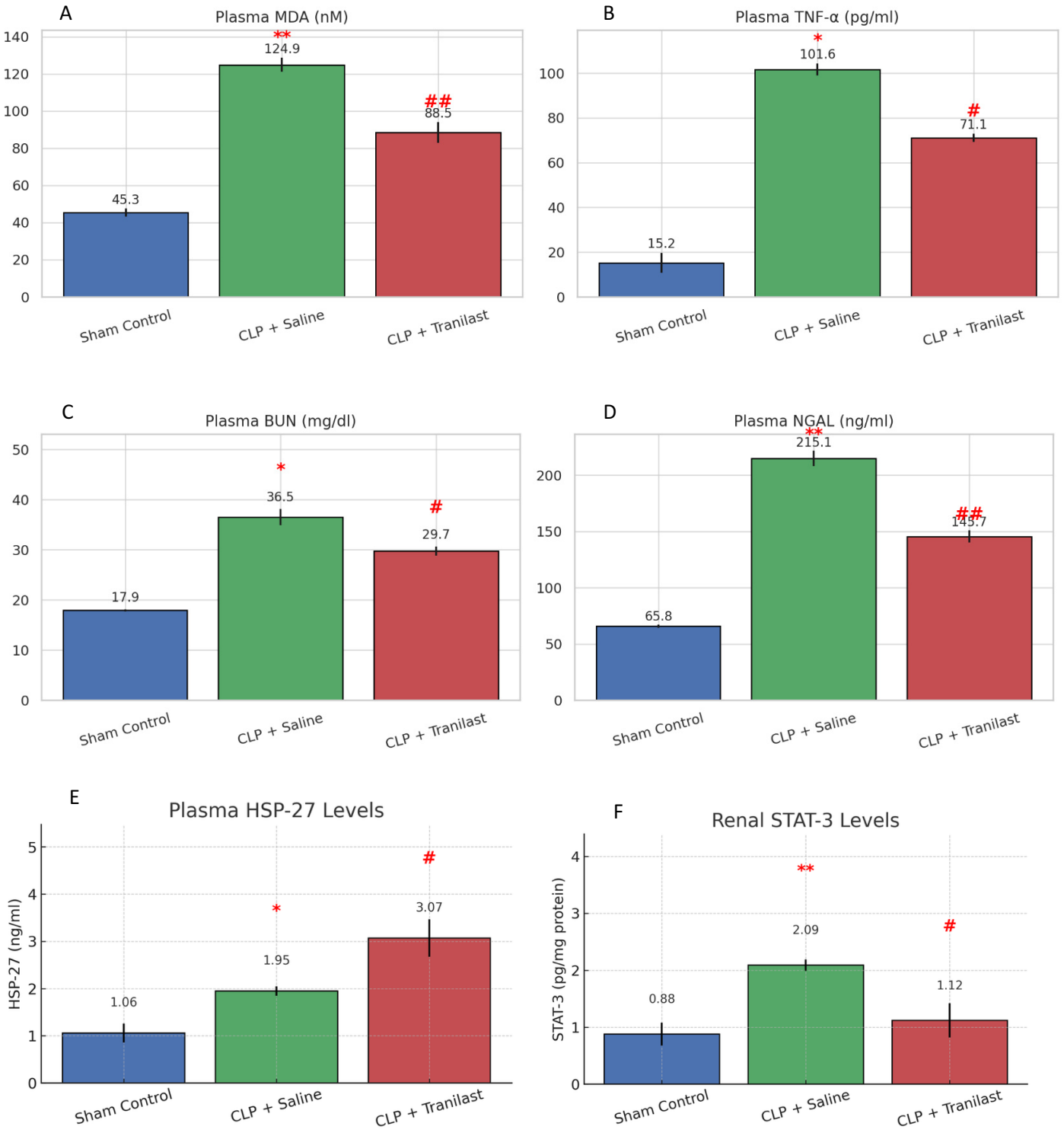

Fig. 2. Effects of tranilast on renal injury and inflammatory markers in CLP-induced sepsis model. Bar graphs illustrate the impact of Tranilast treatment on various plasma and kidney biochemical markers in septic rats. Compared to the CLP+Saline group, Tranilast significantly reduced levels of (A) plasma malondialdehyde (MDA), (B) tumor necrosis factor-alpha (TNF-α), (C) blood urea nitrogen (BUN), (D) neutrophil gelatinase-associated lipocalin (NGAL), and (F) renal STAT-3, while increasing (E) plasma heat shock protein 27 (HSP-27). These findings suggest that Tranilast alleviates oxidative stress, inflammation, and renal injury associated with sepsis. Data are expressed as mean±SEM. *P<0.01, **P<0.001 vs. Sham Control; #P<0.05, ##P<0.001 vs. CLP+Saline.

Table 3. Histopathological evaluation scores of renal tissue damage in experimental groups

	Sham control	CLP and saline group	CLP and tranilast group
Tubular epithelial necrosis	0.1±0.1	3.9±0.1*	1.2±0.1##
Luminal necrotic debris	0.1±0.1	3.1±0.2*	1.9±0.3#
Tubular dilatation	0.1±0.2	1.6±0.4*	1.1±0.2#
Hemorrhage	0.1±0.1	0.8±0.2*	0.4±0.1#
Interstitial inflammation	0.1±0.1	1.1±0.2*	0.5±0.1#

Data are shown as mean±standard error of the median. CLP=cecal ligation and puncture,
*P<0.001 indicates a significant difference compared to the Sham control group,
#P<0.01 and ##P<0.001 indicate significant differences compared to the CLP+Saline group.

1. Lipid Peroxidation Levels

Plasma MDA levels, a marker of oxidative stress, were significantly elevated in the CLP+Saline group (124.9±3.8 nM) compared to the Sham Control group (45.3±2.2 nM; P<0.001). Treatment with Tranilast significantly reduced MDA levels to 88.5±5.5 nM (P<0.001 vs. CLP+Saline), suggesting an antioxidant effect of Tranilast.

2. Tumor Necrosis Factor-Alpha (TNF-α)

A significant increase in plasma TNF-α levels was observed in the CLP + Saline group (101.6±2.7 pg/ml) compared to the Sham Control (15.2±4.5 pg/mL; P<0.01). Tranilast administration significantly attenuated the TNF-α elevation (71.1±1.8 pg/mL; P<0.05 vs. CLP+Saline), indicating an anti-inflammatory effect.

3. Blood Urea Nitrogen (BUN)

The CLP+Saline group had significantly higher BUN levels (36.5±1.6 mg/dL), a sign of renal failure, than the Sham Control group (17.9±0.2 mg/dL; P<0.01). Tranilast treatment significantly reduced BUN levels to 29.7±0.9 mg/dL (P<0.05 vs. CLP+Saline), reflecting improved renal function.

4. Neutrophil Gelatinase-Associated Lipocalin (NGAL)

The CLP+Saline group had considerably higher plasma NGAL levels (215.1±6.8 ng/mL) than the Sham Control group (65.8±1.4 ng/mL; P<0.001). Tranilast administration significantly decreased NGAL levels to 145.7±5.5 ng/mL (P<0.001 vs. CLP+Saline), suggesting attenuation of acute tubular injury.

5. Heat Shock Protein 27 (HSP-27)

HSP-27 levels, which reflect cellular stress and protective responses, were significantly increased in both experimental groups. The CLP+Saline group exhibited elevated HSP-27 (1.95±0.1 ng/ml) compared to the Sham Control (1.06±0.2 ng/ml; P<0.01), while CLP+Tranilast rats showed a further elevation (3.07±0.4 ng/ml; P<0.05 vs. CLP+Saline), indicating a potential cytoprotective upregulation associated with Tranilast.

6. Kidney STAT-3 Expression

In comparison to controls (0.88±0.2 pg/mg; P<0.001), the CLP+Saline group had considerably higher levels of renal tissue STAT-3 protein (2.09±0.1 pg/mg protein). Tranilast significantly suppressed this increase (1.12±0.3 pg/mg; P<0.05 vs. CLP+Saline), suggesting inhibition of the pro-inflammatory STAT-3 signaling pathway.

Histopathological Evaluation of the Kidney

Histopathological scoring revealed extensive renal damage in the CLP+Saline group, as evidenced by significantly increased tubular epithelial necrosis (3.9±0.1), luminal necrotic debris (3.1±0.2), tubular dilatation (1.6±0.4), interstitial inflammation (1.1±0.2), and hemorrhage (0.8±0.2) in comparison to the regular control group (P<0.001 for all parameters). In contrast, rats treated with Tranilast demonstrated a substantial reduction in renal injury scores, with tubular epithelial necrosis decreasing to 1.2±0.1 (P<0.001 vs. CLP+Saline), luminal necrotic debris to 1.9±0.3 (P<0.01), tubular dilatation to 1.1±0.2 (P<0.05), hemorrhage to 0.4±0.1 (P<0.05), and interstitial inflammation to 0.5±0.1 (P<0.01) (Table 3, Fig. 3).

Histological examination using H&E staining supported the quantitative findings. As shown in Fig. 4

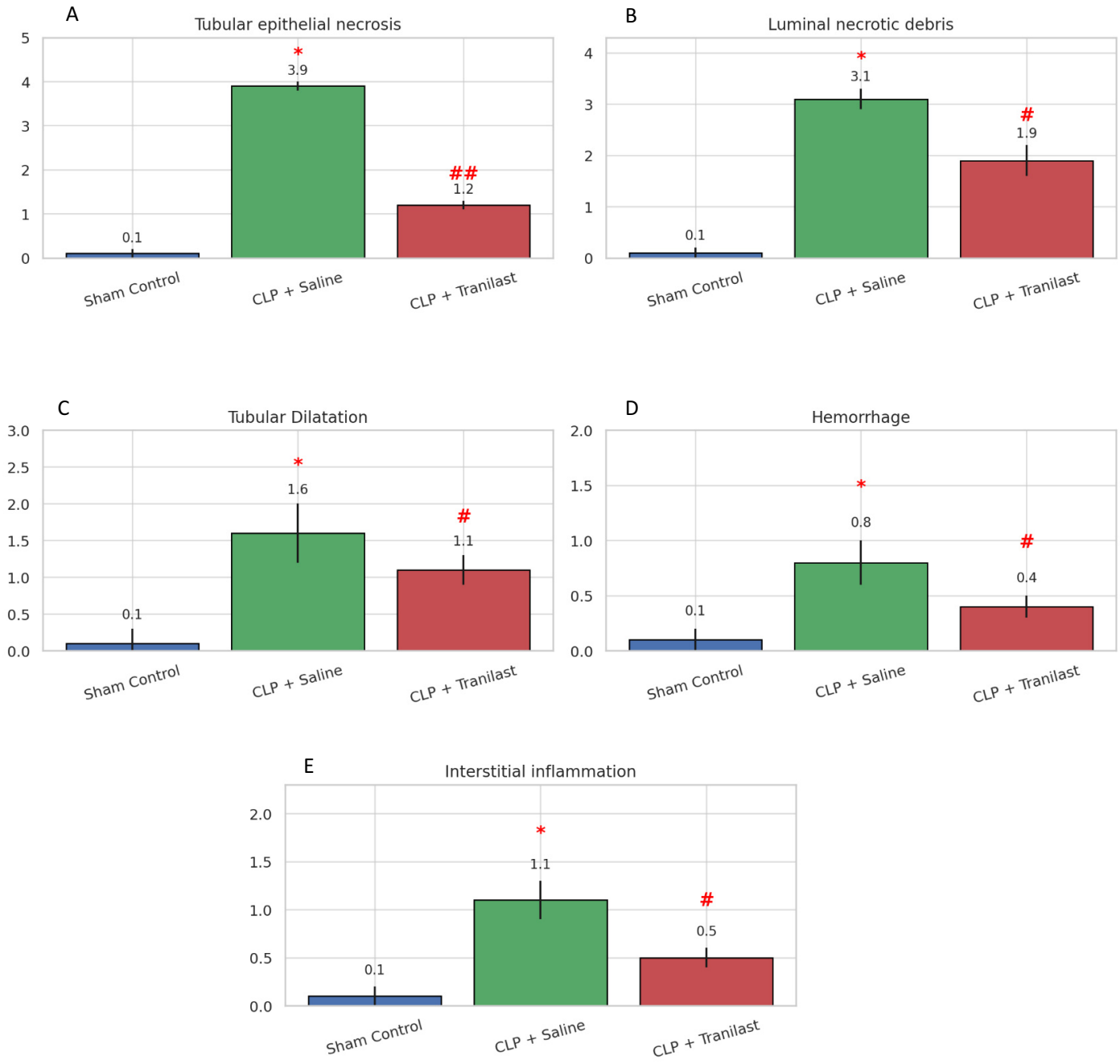

Fig. 3. Histopathological assessment of renal tissue damage in sepsis and the protective effects of tranilast. Histopathological scoring of kidney tissue damage parameters among experimental groups (Sham Control, CLP+Saline, and CLP+Tranilast). The graphs display the following features: (A) Tubular epithelial necrosis, (B) Luminal necrotic debris, (C) Tubular dilatation, (D) Hemorrhage, (E) Interstitial inflammation. Sepsis induced by CLP significantly increased renal tissue damage scores. Tranilast treatment resulted in a marked reduction in necrosis, inflammation, and other pathological findings. Data are expressed as mean ± SEM. *P<0.001 vs. Sham Control; #P<0.01, ##P<0.001 vs. CLP+Saline.

panels A-B, kidneys from the control group exhibited normal glomerular and tubular architecture without signs of injury. In the CLP+Saline group (Fig. 4 panels C-D), widespread tubular epithelial necrosis, tubular dilatation, and dense inflammatory infiltration were apparent. Conversely, sections from the CLP+Tranilast group (Fig 4 panels E-F) showed markedly preserved tubular structures, minimal necrosis, and reduced interstitial cell infiltration, indicating partial protection of renal tissue architecture.

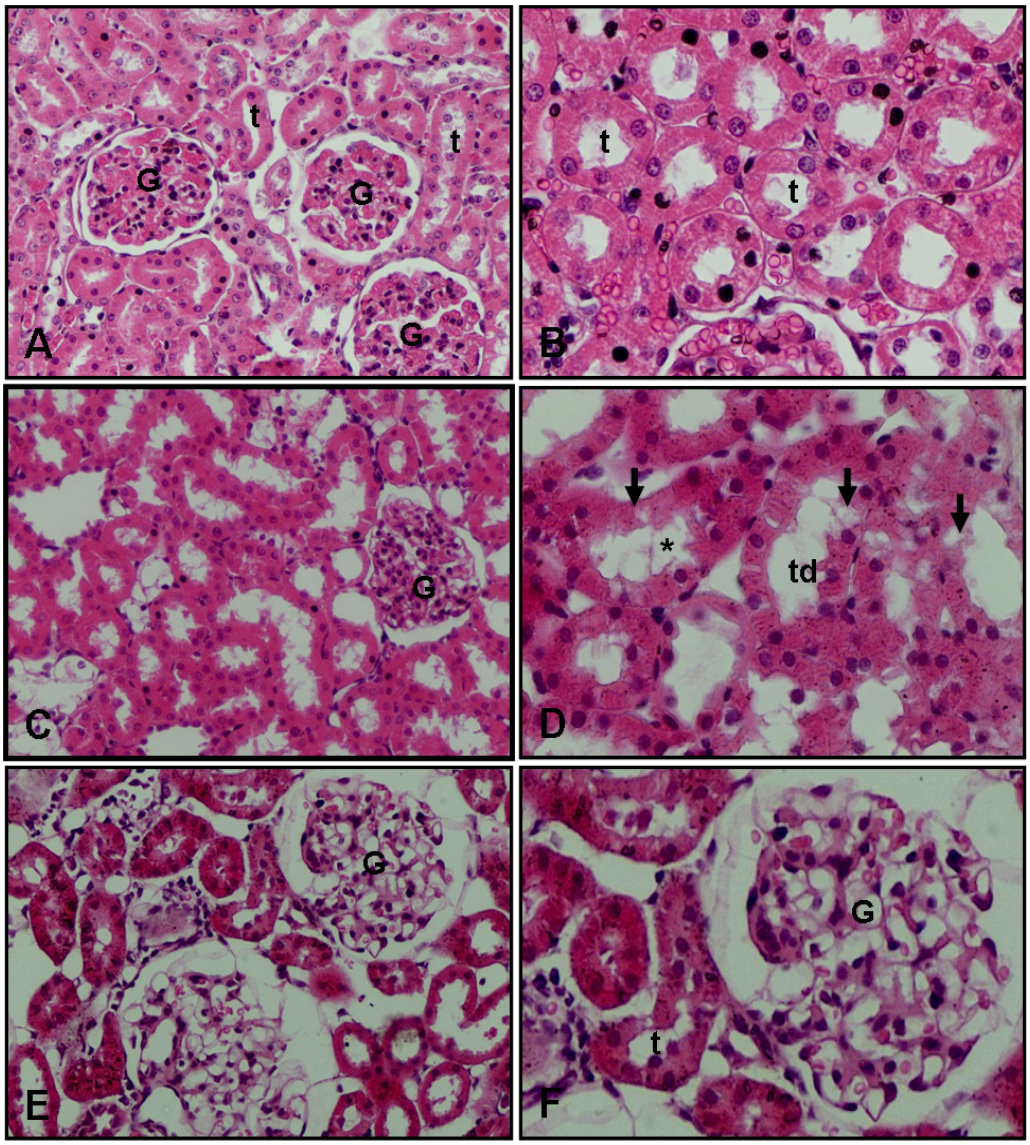

Fig. 3. Representative histological sections of renal tissue in experimental groups (H&E staining; ×20 and ×40 magnification). Histological evaluation of kidney tissue across control and experimental groups using hematoxylin and eosin (H&E) staining. Panels A-B (Control): Renal architecture appears normal, with well-preserved tubular structures (t) and intact glomeruli (G). Panels C-D (CLP+Saline): Significant pathological changes are observed, including marked tubular dilatation (td), epithelial necrosis (black arrows), and disrupted tubular integrity. Panels E-F (CLP+Tranilast): Administration of Tranilast results in notable preservation of renal structure, with reduced tubular necrosis and decreased dilatation, indicating protective effects against sepsis-induced kidney injury.

DISCUSSION

AKI caused by sepsis is still a major problem in intensive care units; it causes a lot of death and illness even though there are ways to help patients [1]. In a rodent sepsis model developed by CLP, our work shows that Tranilast, an anti-inflammatory and antifibrotic drug, greatly reduces kidney damage. The protective effects of Tranilast were mediated through suppression of oxidative stress, inflammation, and STAT-3 signaling, alongside upregulation of cytoprotective HSP-27. These findings align with emerging evidence on the role of STAT-3 in sepsis-associated organ damage and highlight Tranilast as a potential therapeutic candidate for SI-AKI [2].

The survival benefit observed in Tranilast-treated septic rats (75% vs. 50% in controls) underscores its systemic protective effects. Excessive generation of

reactive oxygen species (ROS) is a hallmark of sepsis, which damages cells and causes lipid peroxidation [3]. Our data show that Tranilast significantly reduced plasma MDA, a marker of oxidative stress, suggesting its antioxidant properties mitigate sepsis-induced tissue injury. Similar findings were reported in myocardial ischemia-reperfusion models, where Tranilast decreased ROS generation and preserved organ function [5]. The reduction in oxidative stress may also explain the diminished renal tubular necrosis observed histologically, as lipid peroxidation directly contributes to epithelial cell death in AKI [6].

TNF-α, a key proinflammatory cytokine, plays a central role in sepsis pathophysiology by promoting endothelial dysfunction and leukocyte activation [7]. Our study confirms that Tranilast suppresses TNF-α elevation, consistent with prior reports demonstrating its inhibitory effects on NF-κB and cytokine release [8]. This anti-inflammatory action likely contributes to the preserved renal histoarchitecture seen in Tranilast-treated rats, as TNF-α drives tubular apoptosis and interstitial inflammation in AKI [9]. Furthermore, Tranilast's ability to reduce NGAL, an early biomarker of tubular injury, reinforces its renoprotective role by limiting neutrophil-mediated damage [10].

A novel finding of our study is the significant suppression of renal STAT-3 by Tranilast. These results are consistent with Tan *et al*. [17], who demonstrated in a CLP induced septic mouse model that the flavonoid pectolinarigenin attenuates AKI by inhibiting Janus kinase 2 (JAK2)/STAT3 signaling and preserving mitochondrial integrity. STAT-3 is a critical mediator of inflammation and apoptosis in sepsis, and its hyperactivation exacerbates organ injury [10]. Lee *et al*. [12] demonstrated that STAT-3 inhibition alleviates LPS-induced acute kidney injury, a widely used experimental model of sepsis-associated AKI, by reducing renal tubular apoptosis and cytokine production. Our results corroborate these findings, as Tranilast-treated rats exhibited lower STAT-3 levels alongside improved renal function (reduced BUN) and histopathology. Given that STAT-3 drives transcription of proinflammatory genes (e.g., IL-6, MCP-1), its inhibition by Tranilast may disrupt the vicious cycle of inflammation and renal damage [13]. This observation is supported by studies demonstrating that targeted STAT 3 deletion in renal tubular epithelial cells significantly reduces renal inflammation and tubular damage in LPS induced sepsis models [18].

Interestingly, Tranilast increased plasma HSP-27 levels, a small heat shock protein that confers cellular resistance to stress [4]. HSP-27 stabilizes actin cytoskeletons, inhibits apoptosis, and enhances cell survival under oxidative and inflammatory insults [19]. In sepsis, HSP-27 upregulation has been associated with improved outcomes, possibly by mitigating mitochondrial dysfunction and endoplasmic reticulum stress [16]. Furthermore, in a UUO model, enhanced HSP 27 expression conferred protection to tubular cells, reinforcing the potential cytoprotective role of Tranilast via HSP-27 [20]. Additionally, automated models have shown that activation of the KEAP1/Nrf2/HO-1 axis can reduce renal damage in CLP-induced sepsis, suggesting a synergistic antioxidant mechanism [21]. Our findings suggest that Tranilast enhances endogenous protective mechanisms, further shielding renal tubules from sepsis-induced injury.

Current sepsis management relies on antibiotics, fluid resuscitation, and vasopressors, yet specific therapies for SI-AKI remain lacking [22]. Our data position Tranilast as a multifaceted agent targeting inflammation (TNF-α), oxidative stress (MDA), and STAT-3 signaling. Our observed decrease in MDA is in alignment with Qiongyue *et al*. [23], who reported that irisin attenuated ferroptosis and reduced MDA via activation of the SIRT1/Nrf2 pathway in CLP induced AKI. Unlike corticosteroids, which have mixed results in sepsis, Tranilast offers a more targeted approach with fewer immunosuppressive risks [24]. Clinical trials in fibrotic disorders have already established Tranilast's safety profile, supporting its potential repurposing for sepsis [25].

While our study provides compelling preclinical evidence, limitations include the absence of dose-response experiments and long-term renal function assessments. Future studies should explore Tranilast's effects on other STAT-3 downstream targets (e.g., SOCS3, Bcl-2) and validate findings in large-animal models [26]. Additionally, combination therapies with existing sepsis treatments (e.g., anticoagulants, immunomodulators) warrant investigation [27].

The high mortality rates observed in our CLP model reflect the clinical reality captured in global sepsis data, where acute kidney injury develops in 40-50% of septic patients and doubles mortality risk [28]. The Global Burden of Disease Study highlights sepsis

as responsible for 20% of all global deaths, with AKI representing a key determinant of poor outcomes [26, 29]. Our findings demonstrating Tranilast's survival benefit (75% vs 50%) take on added significance when considering that even a 10% reduction in sepsis-associated AKI mortality could save approximately 400,000 lives annually worldwide [29]. The mechanistic parallels between our rodent model and human disease are supported by multinational studies showing similar patterns of tubular injury and inflammation in septic AKI across diverse populations [30].

The biomarker patterns we observed (NGAL elevation with concurrent STAT-3 activation) mirror findings from large clinical cohorts where NGAL outperformed traditional markers like creatinine in early AKI detection [6, 31] Importantly, our demonstration that Tranilast simultaneously reduces NGAL while suppressing STAT-3, suggesting a unified therapeutic approach targeting both injury biomarkers and their upstream regulators. This aligns with current clinical trends toward biomarker-guided AKI management, though human studies must validate whether Tranilast can modify the well-documented "NGAL trajectory" associated with poor renal recovery [30, 31]. Our NGAL and oxidative stress results mirror those obtained with N-acetylcysteine, which reduces NGAL and KIM 1 levels in CLP rat kidneys, emphasizing the importance of antioxidant strategies [32].
While fluid resuscitation and vasopressors remain cornerstone sepsis therapies [24], our data suggest Tranilast could address the unmet need for direct renal protection. The 30% reduction in BUN with Tranilast compares favorably to effects seen with current supportive care in clinical registries [29]. Notably, the histopathological protection (reduced tubular necrosis scores from 3.9 to 1.2) exceeds benefits reported for renal replacement therapy in septic AKI, which often improves biochemical parameters without modifying underlying tubular injury [28]. This supports preclinical evaluation of Tranilast as an adjunct to conventional sepsis management.

Our study intentionally focused on early AKI mechanisms, but the long-term renal outcomes highlighted by Ronco *et al*. [28] warrant investigation. Approximately 30% of sepsis survivors develop chronic kidney disease, possibly through maladaptive repair processes that our observed HSP-27 induction might mitigate. Additionally, the multinational AKI epidemiology described by Uchino *et al*. [29] suggests Tranilast should be tested across different sepsis phenotypes. The drug's pleiotropic actions may prove particularly valuable in elderly patients who exhibit both exaggerated STAT-3 activation and poor renal recovery post-sepsis [30].

While promising, our findings must be contextualized within broader sepsis trial experiences. The survival benefit, though statistically significant in rodents, may not directly translate to humans given species differences in sepsis responses [24, 26]. The CLP model, while well-validated, doesn't fully replicate the heterogeneity of human septic AKI where comorbidities and treatment variability significantly impact outcomes [28, 29]. Furthermore, the optimal dosing strategy must be refined, as our single-dose protocol differs from the prolonged STAT-3 inhibition potentially needed in clinical sepsis [11, 31].

Tranilast’s ability to suppress proinflammatory signaling appears to play a central role in reducing renal injury markers in our model. Supporting this observation, previous studies have demonstrated Tranilast’s protective effects in various organ injury settings. For instance, Canovai *et al*. [33] reported that Tranilast significantly improved tissue integrity and survival in a rat model of intestinal ischemia-reperfusion injury, primarily through upregulation of heme oxygenase 1 (HO 1), a critical antioxidant defense enzyme. Additionally, Yin *et al*. [14] showed that Tranilast attenuated renal interstitial fibrosis by inhibiting mast cell infiltration and suppressing TGF β signaling in diabetic nephropathy, further underscoring its antifibrotic and anti-inflammatory properties. Of particular relevance to sepsis-induced AKI, Huang *et al*. [34] highlighted the capacity of Tranilast to inhibit NLRP3 inflammasome activation in chronic kidney disease models, a key pathway implicated in both sterile inflammation and septic renal injury. Collectively, these findings suggest that Tranilast exerts multifaceted nephroprotective effects - via modulation of oxidative stress, inhibition of proinflammatory cytokine cascades, and suppression of inflammasome activity - that are consistent with the protective outcomes observed in our sepsis model.

Limitations

This study has several limitations that should be acknowledged. First, although only female rats were

used to reduce variability, the estrous cycle was not monitored, which may have introduced hormonal influences on inflammatory and renal responses. Future studies should include cycle staging to better account for estrogen-related immune modulation. Second, although the 'Sham Control' group underwent laparotomy and bowel manipulation to account for surgical stress, the study did not include a completely non-operated control group. Inclusion of such a group in future studies would help further differentiate the effects of surgical intervention from those of sepsis-induced injury. Third, although the intraperitoneal saline volume (30 ml/kg) was based on prior sepsis models, it approached the upper tolerability limit, and may have influenced peritoneal and systemic responses. Fourth, Tranilast's pharmacokinetics and tissue distribution were not investigated, and although a well-established dose was selected based on prior literature, dose-response analysis was not performed. Finally, estrous cycle variability, potential volume overload from fluid resuscitation, and the lack of post-treatment inflammatory time-course measurements limit the generalizability and mechanistic insight. These factors should be addressed in future studies to better elucidate the therapeutic potential of Tranilast in sepsis-induced AKI.

CONCLUSION

In a rat CLP model, this research shows that Tranilast's complex effects on oxidative stress, inflammation, and cellular signaling pathways provide substantial protection against sepsis- induced acute kidney damage. The marked reduction in renal dysfunction biomarkers (BUN, NGAL) and histopathological damage, coupled with improved survival rates, highlights Tranilast's therapeutic potential in sepsis-associated organ injury. The drug's ability to simultaneously suppress the pro-inflammatory STAT-3 pathway while enhancing cytoprotective HSP-27 expression suggests a unique mechanism that addresses both the inflammatory and cellular stress components of septic AKI. These findings are particularly promising given the lack of targeted therapies currently available for sepsis-induced renal damage. The preservation of renal architecture observed in treated animals, evidenced by reduced tubular necrosis and inflammatory infiltration, further supports Tranilast's organ- protective properties. While these preclinical results require validation in clinical settings, they provide a strong rationale for investigating Tranilast as a potential adjunctive therapy in sepsis management. The study not only advances our understanding of the pathophysiological mechanisms in septic AKI but also identifies a clinically available compound with immediate translational potential. Future research should focus on optimizing dosing regimens, exploring combination therapies with standard sepsis treatments, and investigating potential benefits in other sepsis-affected organs. These findings contribute meaningfully to the ongoing search for effective therapeutic interventions against the devastating consequences of sepsis-associated multi-organ dysfunction.

Ethics Approval and Consent to Participate

The study protocol was reviewed and approved by the Demiroğlu Science University Animal Experiments Local Ethics Committee (HADYEK) (Decision No: 0825015107; date: 15.02.2023). All experimental procedures involving animals were conducted in accordance with the ethical standards of the Guide for the Care and Use of Laboratory Animals published by the U.S. National Institutes of Health. All efforts were made to minimize animal suffering and to reduce the number of animals used.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: MAE, AE, OE; Study Design: MAE, AE, OE; Supervision: MAE, OE; Funding: N/A; Materials: MAE, OE; Data Collection and/or Processing: MAE, OE; Statistical Analysis and/or Data Interpretation: MAE, AE, OE; Literature Review: MAE, OE; Manuscript Preparation: MAE, AE, OE and Critical Review: MAE, AE, OE.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Peerapornratana S, Manrique-Caballero CL, Gómez H, Kellum JA. Acute kidney injury from sepsis: current concepts, epidemiology, pathophysiology, prevention and treatment. Kidney Int. 2019;96(5):1083-1099. doi: 10.1016/j.kint.2019.05.026.
2. Zarbock A, Gomez H, Kellum JA. Sepsis-induced acute kidney injury revisited: pathophysiology, prevention and future therapies. Curr Opin Crit Care. 2014;20(6):588-595. doi: 10.1097/MCC.0000000000000153.
3. Gomez H, Ince C, De Backer D, et al. A unified theory of sepsis-induced acute kidney injury: inflammation, microcirculatory dysfunction, bioenergetics, and the tubular cell adaptation to injury. Shock. 2014;41(1):3-11. doi: 10.1097/SHK.0000000000000052.
4. Hubbard WJ, Choudhry M, Schwacha MG, et al. Cecal ligation and puncture. Shock. 2005;24 Suppl 1:52-57. doi: 10.1097/01.shk.0000191414.94461.7e.
5. Schrier RW, Wang W. Acute renal failure and sepsis. N Engl J Med. 2004;351(2):159-169. doi: 10.1056/NEJMra032401.
6. Mishra J, Ma Q, Prada A, et al. Identification of neutrophil gelatinase-associated lipocalin as a novel early urinary biomarker for ischemic renal injury. J Am Soc Nephrol. 2003;14(10):2534-2543. doi: 10.1097/01.asn.0000088027.54400.c6.
7. Golan-Lagziel T, Lewis YE, Shkedi O, Douvdevany G, Caspi LH, Kehat I. Analysis of rat cardiac myocytes and fibroblasts identifies combinatorial enhancer organization and transcription factor families. J Mol Cell Cardiol. 2018;116:91-105. doi: 10.1016/j.yjmcc.2018.02.003.
8. Tan Z, Liu Q, Chen H, et al. Pectolinarigenin alleviated septic acute kidney injury via inhibiting Jak2/Stat3 signaling and mitochondria dysfunction. Biomed Pharmacother. 2023;159:114286. doi: 10.1016/j.biopha.2023.114286.
9. Hiratochi M, Takamoto M, Tatemichi S, Sugane K. Inhibition of interleukin 5 production with no influence on interleukin 4 production by an anti-allergic drug, tranilast, in Toxocara canis-infected mice. Int J Immunopharmacol. 2000;22(6):463-471. doi: 10.1016/s0192-0561(00)00013-8.
10. Lee H, Jeong AJ, Ye SK. Highlighted STAT3 as a potential drug target for cancer therapy. BMB Rep. 2019;52(7):415-423. doi: 10.5483/BMBRep.2019.52.7.152.
11. Nakamura Y, Oscherwitz J, Cease KB, et al. Staphylococcus δ-toxin induces allergic skin disease by activating mast cells. Nature. 2013;503(7476):397-401. doi: 10.1038/nature12655.
12. Lee SH, Kim KH, Lee SM, et al. STAT3 blockade ameliorates LPS-induced kidney injury through macrophage-driven inflammation. Cell Commun Signal. 2024;22(1):476. doi: 10.1186/s12964-024-01841-1.
13. Jeong S, Kang C, Park S, et al. Eletrophilic Chemistry of Tranilast Is Involved in Its Anti-Colitic Activity via Nrf2-HO-1 Pathway Activation. Pharmaceuticals (Basel). 2021;14(11):1092. doi: 10.3390/ph14111092.
14. Yin DD, Luo JH, Zhao ZY, Liao YJ, Li Y. Tranilast prevents renal interstitial fibrosis by blocking mast cell infiltration in a rat model of diabetic kidney disease. Mol Med Rep. 2018;17(5):7356-7364. doi: 10.3892/mmr.2018.8776.
15. Remick DG. Pathophysiology of sepsis. Am J Pathol. 2007;170(5):1435-1444. doi: 10.2353/ajpath.2007.060872.
16. Erbaş O, Taşkıran D. Sepsis-induced changes in behavioral stereotypy in rats; involvement of tumor necrosis factor-alpha, oxidative stress, and dopamine turnover. J Surg Res. 2014;186(1):262-268. doi: 10.1016/j.jss.2013.08.001.
17. Tan Z, Liu Q, Chen H, et al. Pectolinarigenin alleviated septic acute kidney injury via inhibiting Jak2/Stat3 signaling and mitochondria dysfunction. Biomed Pharmacother. 2023;159:114286. doi: 10.1016/j.biopha.2023.114286.
18. Tan SM, Zhang Y, Cox AJ, Kelly DJ, Qi W. Tranilast attenuates the up-regulation of thioredoxin-interacting protein and oxidative stress in an experimental model of diabetic nephropathy. Nephrol Dial Transplant. 2011;26(1):100-110. doi: 10.1093/ndt/gfq355.
19. Başol N, Erbaş O, Çavuşoğlu T, Meral A, Ateş U. Beneficial effects of agomelatine in experimental model of sepsis-related acute kidney injury. Ulus Travma Acil Cerrahi Derg. 2016;22(2):121-126. doi: 10.5505/tjtes.2015.29499.
20. Vidyasagar A, Reese S, Acun Z, Hullett D, Djamali A. HSP27 is involved in the pathogenesis of kidney tubulointerstitial fibrosis. Am J Physiol Renal Physiol. 2008;295(3):F707-716. doi: 10.1152/ajprenal.90240.2008.
21. Li X, Zhou W, Chen J, et al. Circ_001653 alleviates sepsis associated-acute kidney injury by recruiting BUD13 to regulate KEAP1/NRF2/HO-1 signaling pathway. J Inflamm (Lond). 2024;21(1):37. doi: 10.1186/s12950-024-00409-7.
22. Vincent JL, Jones G, David S, Olariu E, Cadwell KK. Frequency and mortality of septic shock in Europe and North Amer-

ica: a systematic review and meta-analysis. Crit Care. 2019;23(1):196. doi: 10.1186/s13054-019-2478-6.
23. Qiongyue Z, Xin Y, Meng P, et al. Post-treatment With Irisin Attenuates Acute Kidney Injury in Sepsis Mice Through Anti-Ferroptosis via the SIRT1/Nrf2 Pathway. Front Pharmacol. 2022;13:857067. doi: 10.3389/fphar.2022.857067.
24. Rhodes A, Evans LE, Alhazzani W, et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. Intensive Care Med. 2017;43(3):304-377. doi: 10.1007/s00134-017-4683-6.
25. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016;315(8):801-810. doi: 10.1001/jama.2016.0287.
26. Rudd KE, Johnson SC, Agesa KM, et al. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study. Lancet. 2020;395(10219):200-211. doi: 10.1016/S0140-6736(19)32989-7.
27. Hotchkiss RS, Moldawer LL, Opal SM, Reinhart K, Turnbull IR, Vincent JL. Sepsis and septic shock. Nat Rev Dis Primers. 2016;2:16045. doi: 10.1038/nrdp.2016.45.
28. Ronco C, Bellomo R, Kellum JA. Acute kidney injury. Lancet. 2019;394(10212):1949-1964. doi: 10.1016/S0140-6736(19)32563-2.
29. Uchino S, Kellum JA, Bellomo R, et al; Beginning and Ending Supportive Therapy for the Kidney (BEST Kidney) Investigators. Acute renal failure in critically ill patients: a multinational, multicenter study. JAMA. 2005;294(7):813-818. doi: 10.1001/jama.294.7.813.
30. Bagshaw SM, Uchino S, Bellomo R, et al; Beginning and Ending Supportive Therapy for the Kidney (BEST Kidney) Investigators. Septic acute kidney injury in critically ill patients: clinical characteristics and outcomes. Clin J Am Soc Nephrol. 2007;2(3):431-439. doi: 10.2215/CJN.03681106.
31. Poston JT, Koyner JL. Sepsis associated acute kidney injury. BMJ. 2019;364:k4891. doi: 10.1136/bmj.k4891.
32. Wu W, Lan W, Jiao X, et al. Pyroptosis in sepsis-associated acute kidney injury: mechanisms and therapeutic perspectives. Crit Care. 2025;29(1):168. doi: 10.1186/s13054-025-05329-3.
33. Canovai E, Farré R, De Hertogh G, et al. Tranilast Reduces Intestinal Ischemia Reperfusion Injury in Rats Through the Up-regulation of Heme-Oxygenase (HO)-1. J Clin Med. 2025;14(9):3254. doi: 10.3390/jcm14093254.
34. Huang G, Zhang Y, Zhang Y, Ma Y. Chronic kidney disease and NLRP3 inflammasome: Pathogenesis, development and targeted therapeutic strategies. Biochem Biophys Rep. 2023;33:101417. doi: 10.1016/j.bbrep.2022.101417.

The European Research Journal 2025;11(6):1057-1066 Original Article

DOI: https://doi.org/10.18621/eurj.1668332 *Physical Medicine and Rehabilitation*

Subacromial injection failure in shoulder impingement: is somatic amplification the missing link?

Ömer Faruk Bucak[1], Emine Cengiz Çavuşoğlu[2], Adil Öncel[3], Eser Kalaoğlu[4]

[1]Department of Physical Medicine and Rehabilitation, University of Health Sciences, Başakşehir Çam and Sakura City Hospital, İstanbul, Türkiye; [2]Department of Psychiatry, University of Health Sciences, Başakşehir Çam and Sakura City Hospital, İstanbul, Türkiye; [3]Department of Physical Therapy and Rehabilitation, Nişantaşı University Faculty of Medicine, BHT Clinic İstanbul Tema Hospital, İstanbul, Türkiye; [4]Department of Physical Therapy and Rehabilitation, İstanbul Physical Therapy Rehabilitation Training and Research Hospital, İstanbul, Türkiye

ABSTRACT

Objectives: To investigate whether somatosensory amplification - a heightened sensitivity to normal bodily sensations - affects the clinical response to corticosteroid injection in shoulder impingement syndrome (SIS) patients who did not respond to conservative treatment.

Methods: This prospective observational study included 70 patients with SIS and persistent pain despite at least four weeks of physical therapy and non-steroidal anti-inflammatory drugs (NSAIDs). All patients received a standardized corticosteroid injection into the subacromial space under ultrasound guidance. Pain intensity was measured using the Visual Analog Scale (VAS) at rest (VAS-rest) and during movement (VAS-movement), while shoulder-related disability was assessed with the Shoulder Pain and Disability Index (SPADI) at baseline and one month after the injection. Somatosensory amplification - referring to increased sensitivity to normal bodily sensations—was evaluated using the Somatosensory Amplification Scale (SSAS), a brief self-report questionnaire administered at the one-month follow-up by a psychiatrist blinded to clinical outcomes. Treatment response was defined as a ≥30% reduction in VAS-movement, based on the minimal clinically important difference (MCID). Patients were categorized as responders or non-responders accordingly.

Results: Thirty (42.9%) patients were classified as responders and 40 (57.1%) as non-responders. At 1-month, non-responders had significantly higher SSAS scores than responders (30.7 [27.4-32.6] vs. 21.5 [19.4-23.4], P<0.001). Compared to non-responders, the responder group demonstrated more significant improvements in VAS-rest, VAS-movement, and SPADI scores (P<0.001 for all). Spearman correlation analysis revealed strong positive correlations between SAS and VAS-rest (r=0.732), VAS-movement (r=0.748), and SPADI (r=0.734); P<0.001 for all.

Conclusions: Higher levels of somatosensory amplification were linked to a lower likelihood of benefiting from subacromial injection in SIS. These findings support the importance of including psychosomatic assessment in routine care to improve treatment outcomes and guide more personalized management.

Keywords: Shoulder impingement, ultrasound, subacromial injection, somatosensory amplification, minimal clinically important difference (MCID), chronic pain

Received: April 2, 2025 **Accepted:** May 7, 2025 **Available Online:** June 13, 2025 **Published:** November 4, 2025

How to cite this article: Bucak ÖF, Cengiz Çavuşoğlu E, Öncel A, Kalaoğlu E. Subacromial injection failure in shoulder impingement: is somatic amplification the missing link? Eur Res J. 2025;11(6):1057-1066. doi: 10.18621/eurj.1668332

Corresponding author: Ömer Faruk Bucak, MD., Phone: +90 212 909 60 00, E-mail: omerbucak46@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

Shoulder pain is among the most prevalent musculoskeletal complaints, with a lifetime prevalence ranging from 7% to 26% in the general population [1]. Rotator cuff pathologies and Shoulder Impingement Syndrome (SIS) are leading causes of shoulder pain, with SIS accounting for 44% to 65% of outpatient shoulder consultations [2, 3].

SIS is characterized by repetitive mechanical conflict between the supraspinatus tendon and the overlying acromial arch, progressively leading to tendon degeneration, bursitis, and a reduction in subacromial space. Over time, these structural changes impair glenohumeral mobility. In advanced cases, this process may lead to adhesive capsulitis [4]. The focus of current treatments is primarily on relief of shoulder symptoms. Conservative treatments such as physical therapy, therapeutic exercise, and non-steroidal anti-inflammatory drugs (NSAIDs) are generally effective in relieving symptoms [5].

However, a significant proportion of patients fail to achieve significant clinical improvement with these approaches and are consequently referred to ultrasound-guided subacromial corticosteroid injections [6]. Despite the widespread acceptance of subacromial injections as a second-line intervention, variability in treatment response persists as a significant clinical challenge [7]. Even under such conservative therapeutic regimes, a surgical procedure is performed in approximately one-third of cases due to the failure to provide adequate pain relief. Unfortunately, the surgical procedure does not always successfully establish the desired pain relief [8]. Treatment response in SIS is influenced by multiple biomechanical, inflammatory and psychosocial factors and approximately 50% of affected patients continue to report persistent pain even one year after intervention [9].

While biomechanical and anatomical factors such as rotator cuff pathology or acromial morphology are often considered primary contributors to therapeutic resistance, emerging evidence suggests that psychosocial and somatic factors may play a more critical role than previously acknowledged [10, 11]. Among these factors, somatic amplification, which is defined as the tendency to perceive normal bodily sensations as intense, disturbing and serious symptoms of illness, has attracted attention in recent literature [12, 13]. Patients with elevated levels of somatic amplification often report increased pain intensity, decreased treatment response, and poor functional outcomes, regardless of the underlying structural pathology [14].

In clinical practice, it is not uncommon for patients with SIS who undergo subacromial injection to experience minimal or no relief, despite technically successful procedures. These cases often prompt further diagnostic investigations, while the potential role of psychological and somatic contributors is overlooked. This gap highlights the need for a more integrative clinical approach that considers both biological and psychosocial dimensions in understanding treatment failure [15]. When patients are evaluated with the Somatosensory Amplification Scale (SSAS), it may provide insight into the biopsychosocial mechanisms underlying persistent shoulder pain despite interventional treatment [13].

This study aims to investigate whether elevated somatosensory amplification contributes to poor response to subacromial corticosteroid injection in patients with SIS. By identifying this psychosomatic factor, we aim to improve clinical decision-making and support more individualized treatment strategies in routine musculoskeletal care.

METHODS

Study Design

This study was designed as a prospective, observational cohort investigation conducted in the outpatient clinic of the Department of Physical Medicine and Rehabilitation at a tertiary care university hospital. The data collection period spanned from 17 December 2024 to 17 March 2025. The study protocol was reviewed and approved by the local ethics committee in accordance with the Declaration of Helsinki (ethics approval no: 2024-12-11, protocol number: 2024/262).

Participants

A total of 70 patients aged 18 to 65 years with a clinical diagnosis of SIS were prospectively enrolled and had previously received ultrasound-guided subacromial corticosteroid injections as part of their routine clinical care.

Inclusion criteria are the following: (1) Adults aged 18 to 65 years; (2) Clinical diagnosis of SIS confirmed by physical examination and positive impingement signs; (3) Pain duration >3 months; (4)

Inadequate response to conservative treatment (physical therapy modalities, exercise regimens, and NSAIDs use) of ≥4 weeks; and (5) Ability and willingness to provide informed consent and participate in follow-up evaluations.

Exclusion criteria included Previous surgical intervention on the affected shoulder; Diagnosis of inflammatory rheumatic disease (e.g., rheumatoid arthritis); Presence of full-thickness rotator cuff tear or adhesive capsulitis confirmed via clinical and/or imaging evaluation; Major psychiatric illness (e.g., bipolar disorder, schizophrenia); Use of systemic corticosteroids or opioid analgesics in the preceding 4 weeks; Engagement in physical therapy, manual therapy, or injection therapy in the prior 3 months; Pregnancy or current breastfeeding; and Incomplete data or loss to follow-up during the study period.

Ultrasound-Guided Injection Protocol

All injections were performed by the same experienced physiatrist under real-time ultrasound guidance using a Hitachi Aloka Arietta 65 (Hitachi Medical Systems, Tokyo, Japan) ultrasound device equipped with a high frequency (6-18 MHz) linear transducer. With the patient seated and the affected shoulder exposed, a linear high-frequency transducer was used to visualize the subacromial-subdeltoid (SASD) bursa in a transverse plane. A lateral in-plane approach was utilized to guide a 22-gauge needle into the SASD bursa. A total volume of 3 mL consisting of 1 mL triamcinolone acetonide (40 mg/mL) and 2 mL 1% lidocaine hydrochloride was injected aseptically.

Outcome Measures

All participants were evaluated at two time points: prior to the injection (baseline) and at the 1-month follow-up visit. At baseline, the following variables were recorded for each participant: age, sex, duration of symptoms (in months), dominant extremity (right/left), and affected shoulder side.

Primary Outcome Measures

The primary outcome of the study was the SSAS score obtained at the 1-month follow-up. The SSAS is a self-reported psychometric instrument designed to assess the degree to which individuals perceive normal somatic sensations as intense, unpleasant, or symptomatic of serious illness. Originally developed by Barsky *et al*. [12], the scale consists of 10items, each rated on a 5-point Likert scale ranging from 1 ("not at all true") to 5 ("extremely true"). Total scores range from 10 to 50, with higher scores indicating a greater tendency toward somatosensory amplification. The Turkish version of the scale was translated, culturally adapted, and validated by Güleç and Sayar [13]. In the present study, the SSAS was administered at the 1-month follow-up by a trained psychiatrist who was blinded to each patient's treatment response status. This timing was selected to minimize bias from pre-treatment psychological state and to better reflect somatic perception following the intervention.

Secondary Outcome Measures

Pain severity was assessed using the visual analogue scale (VAS) at both baseline and 1-month follow-up. Patients were asked to rate their shoulder pain at rest and during movement on a 10 cm horizontal line, anchored by “no pain” (0) and “worst imaginable pain” (10). The VAS is widely used for musculoskeletal pain due to its simplicity, sensitivity, and validity in clinical trials [16]. Shoulder-specific disability was evaluated using the Shoulder Pain and Disability Index (SPADI), a validated self-report questionnaire initially developed by Roach *et al*. [17]. The SPADI consists of 13 items, distributed across two subscales: pain (5 items) and disability (8 items). Each item is rated from 0 to 10, and scores are normalized to a percentage scale, with higher values indicating greater impairment. The Turkish version was adapted and validated by Bumin *et al*. [18]. In this study, SPADI was administered at baseline and repeated at the 1-month follow-up to assess functional improvement post-injection.

Definition of Treatment Response

Treatment response was defined as a ≥30% reduction in VAS score during movement at the 1-month follow-up compared to baseline. This threshold was selected based on recommendations from previous pain literature, in which a 30% reduction in pain intensity is accepted as the Minimal Clinically Important Difference (MCID) for musculoskeletal pain syndromes [19]. Patients who achieved this degree of improvement were classified as responders, while those who did not were categorized as non-responders.

The detailed structure of the study population and assessment timeline is summarized in Fig. 1.

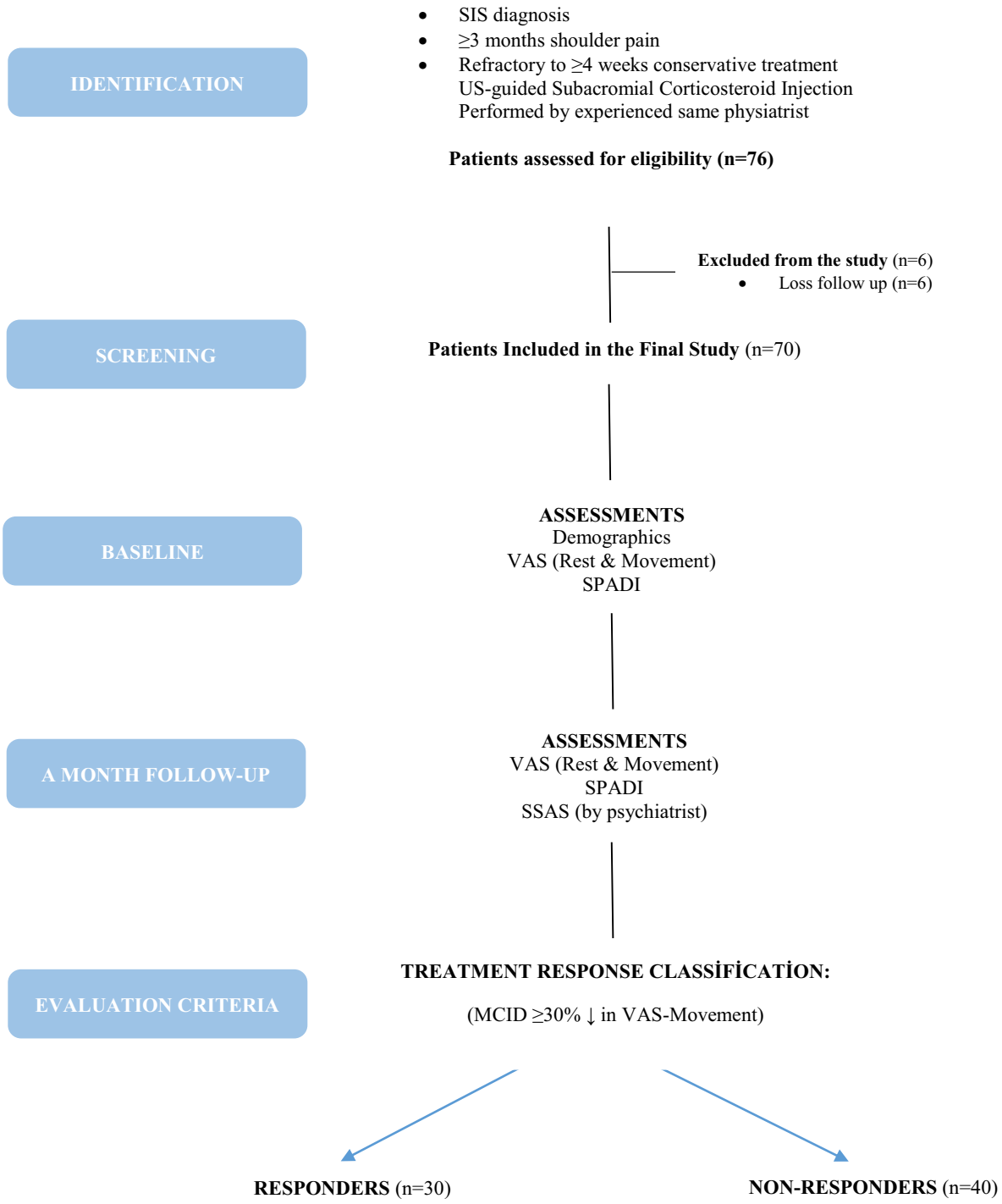

Fig. 1. Flow Chart of the study. SIS=Shoulder Impingement Syndrome, VAS=Visual Analog Scale, SPADI=Shoulder Pain and Disability Index, SSAS=Somatosensory Amplification Scale, MCID=Minimal Clinically Important Difference

Statistical Analysis

All statistical analyses were conducted using IBM SPSS Statistics version 22.0 (IBM Corp., Armonk, NY, USA). The normality of data distribution was assessed using the Kolmogorov–Smirnov test. Continuous variables with normal distribution were expressed as mean ± standard deviation, while non-normally distributed variables were presented as median with interquartile range (IQR). Categorical variables were summarized as frequency and percentage. Comparisons between independent groups (responders vs. non-responders) were performed using the Mann–Whitney U test for non-parametric data. Within-group comparisons between baseline and 1-month follow-up were analyzed using the Wilcoxon signed-rank test. The relationship between SSAS scores and clinical variables (VAS, SPADI) was evaluated using the Spearman correlation coefficient, as the variables did not meet the assumption of normality. A correlation was considered statistically meaningful when $P< 0.05$. Statistical significance was defined as $P<0.05$ for all analyses.

RESULTS

A total of 70 patients diagnosed with SIS were included in the final analysis. Based on their response to ultrasound-guided subacromial corticosteroid injection, 30 (42.9%) patients were classified as responders, while 40 (57.1%) patients were classified as non-responders, using a ≥30% reduction in VAS-Movement score as the threshold for treatment response.

As shown in Table 1, the mean age of patients in the responder group was 41.77±11.19 years, while it was 45.68±10.40 years in the non-responder group. The average duration of pain was slightly longer in the responder group (13.60±6.14 months) compared to the non-responders (11.83±5.31 months), but this difference was not statistically significant ($P>0.05$). In terms of gender distribution, both groups were balanced: 46.7% male and 53.3% female in the responder group, compared to 47.5% male and 52.5% female in the non-responder group. Similarly, the affected shoulder side and dominant side were evenly distributed across groups, with no significant between-group differences ($P>0.05$).

Table 2 presents the comparison of VAS and SPADI scores before and after treatment. At baseline, VAS scores (rest and movement) and SPADI scores were statistically similar between groups ($P>0.05$ for all). However, at the 1-month follow-up, the responder group demonstrated significantly greater improvements. VAS-Rest scores in responders showed a significant reduction, whereas non-responders exhibited a less pronounced improvement ($P<0.001$). Similarly, VAS-Movement scores significantly decreased in responders compared to non-responders ($P<0.001$). SPADI scores improved significantly more in responders than in non-responders ($P<0.001$). Within-group comparisons also confirmed statistically significant improvements in all parameters for both groups ($P<0.001$).

Furthermore, SSAS scores at 1-month were sig-

Table 1. Demographic and clinical characteristics of responder and non-responder groups

		Responder group (n=30)	Nonresponder group (n=40)
Age (years)		41.77±11.19	45.68±10.40
Duration of pain (months)		13.60±6.14	11.83±5.31
Gender	**Male**	14 (46.7%)	19 (47.5%)
	Female	16 (53.3%)	21 (52.5%)
Affected side	**Right**	15 (50.0%)	17 (42.5%)
	Left	15 (50.0%)	23 (57.5%)
Dominant side	**Right**	15 (50.0%)	26 (65.0%)
	Left	15 (50.0%)	14 (35.0%)

The data were shown as mean±standard deviation or n (%)

Table 2. Comparison of clinical outcome measures between responders and non-responders at baseline and 1-month follow-up

		Responder Group (n=30)	Nonresponder Group (n=40)	P value
VAS rest score	Pre-injection	6.0 (5.0-7.0)	6.0 (6.0-6.0)	0.664[U]
	1st month post-injection	2.0 (2.0-2.3)	4.0 (3.3-5.0)	**<0.001**[U]
	P* value	**<0.001**[W]	**<0.001**[W]	
VAS movement score	Pre-injection	8.0 (7.0-9.0)	8.0 (7.0-8.0)	0.904 [U]
	1st month post-injection	3.0 (3.0-4.0)	7.0 (6.0-7.0)	**<0.001**[U]
	P* value	**<0.001**[W]	**<0.001** [W]	
SPADI score	PreiInjection	64.9 (57.8-72.3)	68.1 (58.7-73.5)	0.469[U]
	1st month post-injection	25.9 (23.0-35.4)	53.3 (45.3-63.6)	**<0.001**[U]
	P* value	**<0.001**[W]	**<0.001**[W]	
SSAS score		21.5 (19.4-23.4)	30.7 (27.4-32.6)	0.469[U]

The data were presented as median (interquartile range). VAS=Visual Analog Scale, SPADI=Shoulder Pain and Disability Index, SSAS=SomatoSensory Amplification Scale.

[U]Mann-Whitney U test, [W]Wilcoxon test test, P* value=within-group change P value

nificantly higher in the non-responder group compared to the responder group (P<0.001), as shown in Table 2 and Fig. 2.

Table 3 demonstrated a strong positive correlation between SSAS scores and post-treatment pain and disability outcomes. Significant correlations were observed between SSAS and VAS-Rest, SSAS and VAS-Movement, and SSAS and SPADI (P<0.001 for all). No significant correlations were found between SSAS scores and demographic variables such as age and pain duration (P=0.707 and P=0.480, respectively).

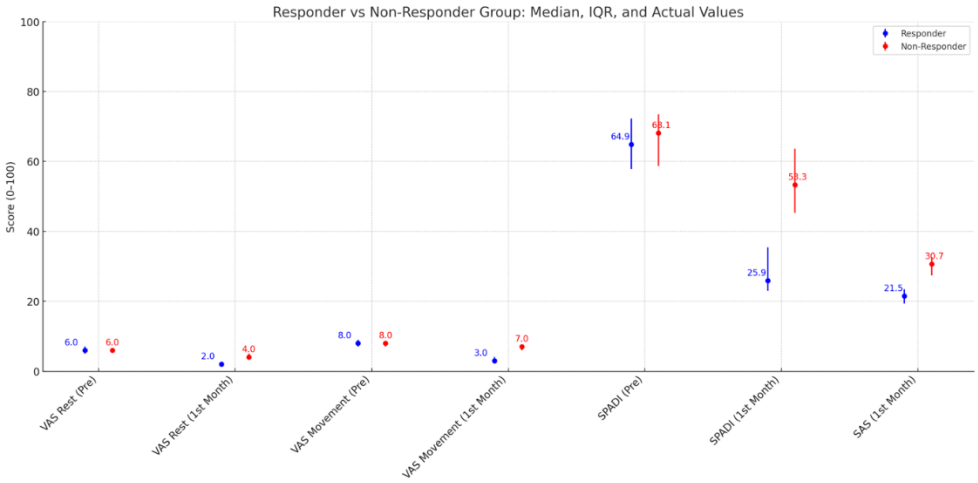

Fig. 2. Comparison of clinical outcome measures between responders and non-responders at baseline and 1-month follow-up. The figure displays the median values and interquartile ranges (IQR) for Visual Analog Scale (VAS) scores at rest and during movement, Shoulder Pain and Disability Index (SPADI), and Somatosensory Amplification Scale (SAS) scores. Data are presented separately for the responder group (blue) and non-responder group (red). Error bars indicate IQRs. Numerical labels denote median values, highlighting distinct clinical trajectories between groups.

Table 3. Correlation between somatosensory amplification scores and clinical parameters

	SSAS score*	
	rho	P value
Age (years)	0.046	0.707
Duration of pain (months)	-0.056	0.48
1st month post-injection VAS rest	0.732	**<0.001**
1st month post-injection VAS movement	0.748	**<0.001**
1st month post-injection SPADI	0.734	**<0.001**

VAS=Visual Analog Scale, SPADI=Shoulder Pain and Disability Index, SSAS=SomatoSensory Amplification Scale.
*Spearman correlation

DISCUSSION

The findings of this study provide compelling evidence that somatosensory amplification is a significant determinant of treatment outcome in patients with SIS undergoing subacromial corticosteroid injection. Non-responders, defined by <30% reduction in movement-evoked pain, exhibited significantly higher SSAS scores compared to responders. Moreover, strong positive correlations between SSAS and post-treatment VAS and SPADI scores indicate that amplified bodily perception contributes not only to pain persistence but also to functional impairment.

These findings question the adequacy of the traditional biomechanical model in explaining treatment variability in SIS. Despite the role of structural abnormalities, our results align with recent literature highlighting the influence of central pain processing and psychosomatic factors [10, 14].

Somatosensory amplification, as defined by Barsky *et al*., reflects a tendency to experience normal bodily sensations as unusually intense and distressing [12]. In chronic pain populations, high SSAS scores have been associated with increased pain perception, reduced functional capacity, and lower responsiveness to treatment, regardless of the structural pathology [13, 14, 20]. The current findings extend this association to SIS, suggesting that SSAS may serve as a prognostic indicator for suboptimal response to subacromial injections.

This interpretation is supported by findings from fibromyalgia research, a prototypical model of centrally mediated pain, where somatic amplification has consistently been linked to worse clinical outcomes [21-23]. Akay *et al*. [24] further demonstrated that elevated SSAS scores were independently associated with increased disability, greater pain intensity, and lower quality of life, even after adjusting for depression and anxiety. These observations highlight somatic amplification not merely as a byproduct of pain, but as a central factor contributing to the functional and emotional impact of chronic pain syndromes [25, 26]. Although fibromyalgia and SIS differ in anatomical scope, the shared neurocognitive features - such as affective dysregulation and amplified interoception - suggest that somatosensory amplification may exert a similarly detrimental effect in localized pain syndromes like SIS [27, 28]. Our results align with this paradigm, indicating that SSAS may reflect a broader neurophysiological vulnerability that transcends diagnostic boundaries and significantly modulates treatment response [29].

Notably, the absence of significant correlations between SSAS and age or symptom duration supports the hypothesis that somatic amplification is not merely a byproduct of chronicity, but rather a stable psychological trait influencing pain processing. This aligns with prior research proposing that somatic amplification operates independently of demographic or disease duration factors [14].

The clinical applicability of these findings lies in the potential utility of SSAS as a brief, reliable, and easy-to-administer tool for screening patients who may be at risk of suboptimal outcomes following standard interventions. Pre-treatment identification of individuals with high somatic amplification scores could facilitate early stratification and guide the integration of adjunctive therapies—such as pain neuroscience

education, cognitive-behavioral therapy, or centrally acting pharmacological agents—into routine musculoskeletal care [20, 28]. These approaches, which target the cognitive-affective components of pain perception, have been shown to improve outcomes in both generalized and regional pain syndromes exhibiting central sensitization features [30-32].

The correlations between SSAS and both VAS and SPADI emphasize the multidimensional role of somatic amplification. Despite uniform treatment, patients with higher SSAS scores experienced greater pain and disability, suggesting that perception plays a key role in treatment outcomes. Such findings support a biopsychosocial model of SIS and highlight the importance of integrating somatization-related assessments into both research protocols and daily clinical algorithms [33, 34].

From a mechanistic perspective, SSAS may indicate an increased focus on and heightened perception of bodily sensations, which are key features of altered central pain processing [12, 35]. While SSAS does not directly measure spinal or brain-level sensitization, it offers useful indirect insights into the emotional and cognitive aspects of pain perception. In clinical settings where advanced tools such as quantitative sensory testing (QST) or neuroimaging are unavailable, SSAS may serve as a practical alternative for assessing these pain-related mechanisms [36].

The results of this study, when considered as a whole, provide a compelling argument for the inclusion of SSAS in clinical workflows and research protocols targeting persistent shoulder pain. Strengths of this study include its prospective design, standardized ultrasound-guided injections, and use of validated outcome tools (SSAS, SPADI, VAS) consistent with MCID guidelines [19]. Administering the SSAS at one-month follow-up minimized expectation bias and allowed a clearer evaluation of post-treatment perceptions.

Limitations

This study has several limitations that warrant consideration. Firstly, although the SSAS provides valuable insight into heightened bodily perception, it does not directly assess the broader spectrum of central sensitization mechanisms, which encompass altered nociceptive processing, impaired descending inhibition, and widespread hyperalgesia. Given the overlapping yet distinct nature of somatic amplification and central sensitization, the absence of objective assessments - such as QST, pressure pain thresholds, or central sensitization inventory - limits the ability to delineate the relative contribution of central neuroplasticity to treatment resistance. It is possible that the observed treatment non-responsiveness represents a multifaceted interplay between peripheral input, cognitive-emotional factors, and central sensitization, which may act synergistically. Future studies incorporating both SSAS and direct central sensitization measures in tandem could provide a more comprehensive understanding of pain modulation in SIS. Secondly, the 1-month follow-up window, although suitable for detecting early treatment effects, does not provide information on long-term response sustainability. Thirdly, we did not assess key psychological constructs such as depression, pain catastrophizing, fear-avoidance, or health anxiety all of which may co-occur with somatic amplification and influence outcomes. Future studies incorporating multidimensional psychometric batteries are needed to clarify these relationships. Lastly, this single-center design may limit the generalizability of findings to broader or culturally diverse populations.

CONCLUSION

The present findings suggest that somatic amplification exerts a significant influence on pain perception and functional outcomes in patients with SIS following subacromial corticosteroid injection. Patients who failed to achieve meaningful symptom relief exhibited significantly higher SSAS scores, highlighting the role of psychosomatic mechanisms in modulating treatment response. In light of these results, it is recommended that somatic amplification be recognised and addressed in future treatment algorithms, not only to improve therapeutic sensitivity and reduce unnecessary invasive interventions, but also to guide the development of more individualised, patient-centred care models.

In cases of persistent or disproportionate pain despite technically successful procedures, psychiatric consultation should be actively considered—particularly for patients with high SSAS scores—so that underlying affective or cognitive factors contributing to pain chronification can be identified and managed

through targeted psychotherapeutic or pharmacological strategies. Integrating psychosomatic screening and interdisciplinary collaboration into musculoskeletal practice may ultimately enhance both pain resolution and functional recovery in this patient population.

Ethics Approval and Consent to Participate

The study was approved by the Başakşehir Çam and Sakura City Hospital Clinical Research Ethics Committee (Decision no.: 2024-262 and date: 16.12.2024). All procedures performed during data collection, review of patient records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Written informed consent was obtained from all individual participants included in the study.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: ÖFB, ECÇ, AÖ; Study Design: ÖFB, ECÇ, AÖ, EK; Supervision: ÖFB, Funding: N/A; Materials: ÖFB, ECÇ, EK; Data Collection and/or Processing: ÖFB, ECÇ, AÖ; Statistical Analysis and/or Data Interpretation: ÖFB, EK; Literature Review: ÖFB, ECÇ, AÖ, Manuscript Preparation: ÖFB, and Critical Review: ÖFB, ECÇ, AÖ, EK.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the author(s) and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Luime JJ, Koes BW, Hendriksen IJ, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol. 2004;33(2):73-81. doi: 10.1080/03009740310004667.
2. Garving C, Jakob S, Bauer I, Nadjar R, Brunner UH. Impingement Syndrome of the Shoulder. Dtsch Arztebl Int. 2017;114(45):765-776. doi: 10.3238/arztebl.2017.0765.
3. Linaker CH, Walker-Bone K. Shoulder disorders and occupation. Best Pract Res Clin Rheumatol. 2015;29(3):405-423. doi: 10.1016/j.berh.2015.04.001.
4. de la Serna D, Navarro-Ledesma S, Alayon F, Lopez E, Pruimboom L. A Comprehensive View of Frozen Shoulder: A Mystery Syndrome. Front Med (Lausanne). 2021;8:663703. doi: 10.3389/fmed.2021.663703.
5. Park HB, Yokota A, Gill HS, El Rassi G, McFarland EG. Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndrome. J Bone Joint Surg Am. 2005;87(7):1446-1455. doi: 10.2106/JBJS.D.02335.
6. Metin Terzibaşıoğlu A, Çınar Ç, Soluk Özdemir Y, Doran M, Bagatır N, Öneş K. Ultrasound-guided subacromial corticosteroid injections in comparison to anatomic landmark guided injections in subacromial impingement syndrome. Arch Health Sci Res. 2022;9(3):180-185. doi: 10.5152/ArcHealthSciRes.2022.22055.
7. Khan KM, Cook JL, Taunton JE, Bonar F. Overuse tendinosis, not tendinitis part 1: a new paradigm for a difficult clinical problem. Phys Sportsmed. 2000;28(5):38-48. doi: 10.3810/psm.2000.05.890.
8. Karjalainen TV, Jain NB, Page CM, et al. Subacromial decompression surgery for rotator cuff disease. Cochrane Database Syst Rev. 2019 Jan 17;1(1):CD005619. doi: 10.1002/14651858.CD005619.pub3.
9. Coronado RA, Simon CB, Valencia C, George SZ. Experimental pain responses support peripheral and central sensitization in patients with unilateral shoulder pain. Clin J Pain. 2014;30(2):143-151. doi: 10.1097/AJP.0b013e318287a2a4.
10. Sheikhzadeh A, Wertli MM, Weiner SS, Rasmussen-Barr E, Weiser S. Do psychological factors affect outcomes in musculoskeletal shoulder disorders? A systematic review. BMC Muscu-

loskelet Disord. 2021;22(1):560. doi: 10.1186/s12891-021-04359-6.
11. Martinez-Calderon J, Meeus M, Struyf F, Miguel Morales-Asencio J, Gijon-Nogueron G, Luque-Suarez A. The role of psychological factors in the perpetuation of pain intensity and disability in people with chronic shoulder pain: a systematic review. BMJ Open. 2018;8(4):e020703. doi: 10.1136/bmjopen-2017-020703.
12. Barsky AJ, Wyshak G, Klerman GL. The somatosensory amplification scale and its relationship to hypochondriasis. J Psychiatr Res. 1990;24(4):323-34. doi: 10.1016/0022-3956(90)90004-a.
13. Güleç H, Sayar K. Reliability and validity of the Turkish form of the Somatosensory Amplification Scale. Psychiatry Clin Neurosci. 2007;61(1):25-30. doi: 10.1111/j.1440-1819.2007.01606.x.
14. Nijs J, Leysen L, Vanlauwe J, et al. Treatment of central sensitization in patients with chronic pain: time for change? Expert Opin Pharmacother. 2019;20(16):1961-1970. doi: 10.1080/14656566.2019.1647166.
15. Linton SJ, Shaw WS. Impact of psychological factors in the experience of pain. Phys Ther. 2011;91(5):700-11. doi: 10.2522/ptj.20100330.
16. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken). 2011;63 Suppl 11:S240-52. doi: 10.1002/acr.20543.
17. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. Arthritis Care Res. 1991;4(4):143-149.
18. Bumin G, Tüzün EH, Tonga E. The Shoulder Pain and Disability Index (SPADI): Cross-cultural adaptation, reliability, and validity of the Turkish version. J Back and Musculoskelet Rehabil. 2008;21(1):57-62. doi: 10.3233/BMR-2008-21108.
19. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain. 2008;9(2):105-121. doi: 10.1016/j.jpain.2007.09.005.
20. Louw A, Diener I, Butler DS, Puentedura EJ. The effect of neuroscience education on pain, disability, anxiety, and stress in chronic musculoskeletal pain. Arch Phys Med Rehabil. 2011;92(12):2041-2056. doi: 10.1016/j.apmr.2011.07.198.
21. Wolfe F, Hauser W. Fibromyalgia diagnosis and diagnostic criteria. Ann Med. 2011;43(7):495-502. doi: 10.3109/07853890.2011.595734.
22. Meeus M, Nijs J. Central sensitization: a biopsychosocial explanation for chronic widespread pain in patients with fibromyalgia and chronic fatigue syndrome. Clin Rheumatol. 2007;26(4):465-473. doi: 10.1007/s10067-006-0433-9.
23. Giesecke T, Gracely RH, Williams DA, et al. The relationship between depression, clinical pain, and experimental pain in a chronic pain cohort. Arthritis Rheum. 2005;52(5):1577-1584. doi: 10.1002/art.21008.
24. Akay E, Şenormancı Ö, Eröksüz R, Şenormancı G, Demirci OO. The effect of psychopathology on quality of life and disability in patients with fibromyalgia. Eur Res J. 2022;8(1):103-110. doi: 10.18621/eurj.854956.
25. Fallon N, Li X, Chiu Y, Nurmikko T, Stancak A. Altered cortical processing of observed pain in patients with fibromyalgia syndrome. J Pain. 2015;16(8):717-726. doi: 10.1016/j.jpain.2015.04.008.
26. Barsky AJ, Wyshak G, Klerman GL. The somatosensory amplification scale and its relationship to hypochondriasis. J Psychiatr Res. 1990;24(4):323-334. doi: 10.1016/0022-3956(90)90004-a.
27. Meeus M, Nijs J, Huybrechts S, Truijen S. Evidence for generalized hyperalgesia in chronic fatigue syndrome: a case control study. Clin Rheumatol. 2010;29(4):393-398. doi: 10.1007/s10067-009-1339-0.
28. Chiarotto A, Deyo RA, Terwee CB, et al. Core outcome domains for clinical trials in non-specific low back pain. Eur Spine J. 2015;24(6):1127-1142. doi: 10.1007/s00586-015-3892-3.
29. Lumley MA, Schubiner H, Lockhart NA, et al. Emotional awareness and expression therapy, cognitive behavioral therapy, and education for fibromyalgia: a cluster-randomized controlled trial. Pain. 2017;158(12):2354-2363. doi: 10.1097/j.pain.0000000000001036.
30. Lepri B, Romani D, Storari L, Barbari V. Effectiveness of Pain Neuroscience Education in Patients with Chronic Musculoskeletal Pain and Central Sensitization: A Systematic Review. Int J Environ Res Public Health. 2023;20(5):4098. doi: 10.3390/ijerph20054098.
31. Williams ACC, Fisher E, Hearn L, Eccleston C. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev. 2020;8(8):CD007407. doi: 10.1002/14651858.CD007407.pub4.
32. Shin S, Kim H. Carryover Effects of Pain Neuroscience Education on Patients with Chronic Lower Back Pain: A Systematic Review and Meta-Analysis. Medicina (Kaunas). 2023;59(7):1268. doi: 10.3390/medicina59071268.
33. Gatchel RJ, Peng YB, Peters ML, Fuchs PN, Turk DC. The biopsychosocial approach to chronic pain: scientific advances and future directions. Psychol Bull. 2007;133(4):581-624. doi: 10.1037/0033-2909.133.4.581.
34. Nijs J, Meeus M, Cagnie B, et al. A modern neuroscience approach to chronic spinal pain: combining pain neuroscience education with cognition-targeted motor control training. Phys Ther. 2014;94(5):730-738. doi: 10.2522/ptj.20130258.
35. Woolf CJ. Central sensitization: implications for the diagnosis and treatment of pain. Pain. 2011;152(3 Suppl):S2-S15. doi: 10.1016/j.pain.2010.09.030.
36. Neblett R, Cohen H, Choi Y, et al. The Central Sensitization Inventory (CSI): establishing clinically significant values for identifying central sensitivity syndromes in an outpatient chronic pain sample. J Pain. 2013;14(5):438-445. doi: 10.1016/j.jpain.2012.11.012.

The European Research Journal 2025;11(6):1067-1077 Original Article

DOI: https://doi.org/10.18621/eurj.1768888 *Medical Oncology*

Prognostic impact of pan-immune-inflammation value and prognostic nutritional index in metastatic colorectal cancer patients treated with regorafenib

Sedat Yıldırım[1], Özkan Alan[2], Hacer Şahika Yıldız[1], Akif Doğan[3], Seval Ay Ersoy[1], Hatice Odabaş[1]

[1]Department of Medical Oncology, University of Health Sciences, Kartal Dr. Lütfi Kirdar City Hospital, İstanbul, Türkiye; [2]Department of Medical Oncology, Cerrahpaşa Faculty of Medicine, Istanbul University-Cerrahpaşa, İstanbul, Türkiye; [3]Department of Medical Oncology, University of Health Sciences, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, İstanbul, Türkiye

ABSTRACT

Objectives: Metastatic colorectal cancer (mCRC) remains a major cause of cancer-related mortality, with limited therapeutic options available after the failure of standard treatments. Regorafenib, an oral multichines inhibitor, has shown a survival advantage in treatment refractory mCRC. This research aims to evaluate the real-world effectiveness of regorafenib and to investigate the prognostic significance of inflammatory and nutritional indicators, specifically the Pan-Immune-Inflammation Value (PIV) and Prognostic Nutritional Index (PNI).

Methods: We conducted a retrospective analysis of 166 mCRC patients who received regorafenib 2014 to 2024. Demographic, clinical, and pathological data, treatment responses, and laboratory indicators were gathered. Survival outcomes were evaluated by Kaplan–Meier analysis, whereas prognostic variables were analyzed using Cox regression.

Results: The average age was 58 years, and 59% of patients were male. The overall disease control rate with regorafenib was 33%, and the objective response rate was 11%. The median progression-free survival (PFS) was 3.8 months, and the median overall survival (OS) was 9.1 months. A high PIV (≥309.1) was associated with significantly shorter overall survival (7.1 vs. 11.8 months, P=0.001), whereas a high PNI (≥47.9) was correlated with longer overall survival (10.7 vs. 6.7 months, P=0.02). Multivariate analysis confirmed PIV and PNI as independent prognostic indicators.

Conclusions: Regorafenib provides a modest survival advantage and disease stabilization in previously treated mCRC patients. PIV and PNI are independent prognostic biomarkers that may assist in patient stratification and therapy optimization. Further studies are warranted to refine predictive markers and dosing strategies.

Keywords: Regorafenib, metastatic colorectal cancer, Prognostic nutritional index, pan-immune-inflammation value, prognostic biomarkers

Received: August 25, 2025 **Accepted:** September 27, 2025 **Available Online:** October 3, 2025 **Published:** November 4, 2025

How to cite this article: Yıldırım S, Alan Ö, Yıldız HŞ, Doğan A, Ay Ersoy S, Odabaş H. Prognostic impact of pan-immune-inflammation value and prognostic nutritional index in metastatic colorectal cancer patients treated with regorafenib. Eur Res J. 2025;11(6):1067-1077. doi: 10.18621/eurj.1768888

Corresponding author: Sedat Yıldırım, MD., Phone: +90 216 458 50 00, E-mail: rezansedat@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

Colorectal cancer (CRC) is a significant cause of morbidity and mortality in Western countries. Although it is the fourth most commonly diagnosed cancer, it ranks second in cancer-related mortality [1, 2]. Approximately 15-30% of CRC cases are diagnosed at the metastatic stage, while 20–50% of patients initially diagnosed with localized or locally advanced disease develop metastases during follow-up [3]. Metastatic CRC (mCRC) is characterized by a dismal prognosis, with a five-year overall survival (OS) rate of less than 20% [4].

The standard treatment of CRC includes 5-fluorouracil (5-FU) in combination with irinotecan or oxaliplatin, and antiangiogenic therapies such as bevacizumab, aflibercept, or ramucirumab, or epidermal growth factor receptor (EGFR)-targeted antibodies such as panitumumab or cetuximab [5-8]. Second-line treatment is determined based on the agents used in the first-line setting [9]. Even after first- and second-line therapies, many patients retain good performance but have limited therapeutic options [10].

Regorafenib, a multichines inhibitor, has emerged as an important therapeutic option for patients with mCRC who have exhausted standard treatment options [11, 12]. Preclinical studies demonstrated that regorafenib can significantly inhibit tumor growth, reduce vascularization, and prevent metastasis formation, providing a strong rationale for its use in advanced CRC [13, 14]. In a randomized controlled trial comparing regorafenib with placebo, the median OS was 6.4 months in the regorafenib group versus 5.0 months in the placebo group, highlighting a survival benefit in this patient population. This study established regorafenib as the first small molecule multichines inhibitor to show a survival benefit in mCRC patients who had progressed after all standard therapies [15].

Systemic inflammation plays a critical role in cancer progression and metastasis. It is well established that cancer progression frequently correlates with systemic inflammatory response [16-22]. Hematologic parameters derived from peripheral blood have been investigated as prognostic indicator in CRC [23-27]. In this context, the Pan-Immune-Inflammation Value (PIV), a novel biomarker reflecting the host immune-inflammatory status, has been evaluated in various cancers, including colorectal cancer [23, 24, 28, 29].

Another well-established prognostic index in colorectal cancer is the Prognostic Nutritional Index (PNI) [25, 26]. PNI is calculated based on serum albumin levels and lymphocyte counts. Lymphocytes play a vital role in suppressing the development and proliferation of cancer cells and are essential for maintaining immune surveillance. A decrease in lymphocyte count creates a favorable environment for tumor cells to evade the immune system, promoting disease progression. Albumin, a major serum protein synthesized in the liver, is a marker of nutritional status and systemic inflammation. Low albumin levels often indicate poor nutritional status and increased tumor-related inflammation [30]. Thus, the PNI, derived from the combination of albumin and lymphocyte levels, reflects both inflammatory status and nutritional condition [31].

In this study, we aimed to evaluate laboratory, pathological, and clinical parameters with potential prognostic significance in patients with metastatic colorectal cancer treated with regorafenib.

METHODS

Study Population and Design

Patients diagnosed with metastatic colorectal cancer who were followed and treated in the Medical Oncology Department of Kartal Dr. Lütfi Kirdar City Hospital between January 2014 and April 2024 and who received regorafenib with available treatment response data were included in the study. Non-metastatic patients, those with a concurrent active second malignancy or individuals under 18 years of age were excluded.

Demographic variables; Eastern Cooperative Oncology Group Performance Status (ECOG PS), pathological and immunohistochemical features of the tumor, regorafenib dosing and treatment duration, recurrence dates, and survival outcomes were retrospectively extracted from medical records.

Laboratory parameters such as PIV and PNI were calculated for each patient, and cutoff values for OS were determined using receiver operating characteristic (ROC) curve analysis. Sensitivity and specificity were also assessed for each identified cutoff. Median values were used as reference for markers where cutoff values could not be determined via ROC analysis.

Table 1. Baseline demopgrahpci and clinic findings (n=166)

Variables		Data
Age (years)		58 (25-82)
Gender	Female	68 (41)
	Male	98 (59)
ECOG PS	PS 0	61 (37)
	PS 1	89 (53)
	PS ≥ 2	16 (10)
Location	Right colon	17 (10)
	Left Colon	73 (44)
	Transverse colon	8 (5)
	Cecum	6 (4)
	Rectum	62 (37)
Primary surgery	Present	127 (76)
	Absent	39 (24)
Pathologis subtype	Adenocarcinoma	153 (92)
	Mucinous adenocarcinoma	13 (8)
RAS status	Wild type	85 (51)
	Mutant type	81 (49)
MSI status	Low	106 (64,4)
	High	1 (0,6)
	Unknown	59 (36)
HER-2 status	Score 0	14 (8)
	Score 1	19 (11)
	Score 2 FISH negative	16 (10)
	Score 2 FISH positive	5 (3)
	Score 3	2 (1)
	Unkown	110 (67)
Metastatis status	Metachronous	74 (45)
	Synchronous	92 (55)
Liver metastasis	Present	128 (23)
	Absent	38 (77)
Local treatment	Yes	112 (72)
	No	44 (28)
CEA (ng/mL)		71.2 (0.6-3317)
CA 19-9 (U/mL)		65.2 (0.8-10000)

Data are shown as median (minimum-maximum) or n (%) where appropriate. CA 19-9=Carbohydrate Antigen, CEA=Carcinoma Embryonic Antigen, ECOG PS=Eastern Cooperative Oncology Group Performance Status, FISH=Fluorescence in Situ Hybridization, HER-2=Human Epidermal Growth Factor Receptor-2, MSI=Microsatellite Instability, RAS=Rat Sarkoma

The Pan-Immune-Inflammation Value (PVI) and Prognostic Nutritional Index (PNI)

PIV and PNI were calculated using neutrophil, lymphocyte, platelet, and monocyte count along with serum albumin levels, obtained from routine blood tests performed prior to the first dose of regorafenib.

PNI and PIV were calculated using the following formulas:

PNI= (10×serum albumin (g/dL)) + (0.005×total lymphocyte count),

PVI= (neutrophil count × platelet count × monocyte count) / lymphocyte count).

This single-center retrospective study was conducted after obtaining approval from the institutional ethics committee of the University of Health Sciences, Kartal Dr. Lütfi Kirdar City Hospital. The date of the ethics committee meeting was April 29, 2024, and the decision number was 2024/010.99/3/1. All procedures performed during data collection, review of patient records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments.

Statistical Analysis

Descriptive statistics included mean, standard deviation, median, minimum, maximum, frequency, and percentage values. The distribution of variables was assessed using the Kolmogorov-Smirnov and Shapiro-Wilk tests. For the analysis of non-normally distributed quantitative independent variables, the Mann-Whitney U test was used. For categorical independent variables, the chi-square test was used; when the assumptions of the chi-square test were not met, the Fisher's exact test was applied. The effect size was analyzed using univariate and multivariate logistic regression analyses. Survival analysis was performed using the Kaplan-Meier method. All analyses were conducted using SPSS software version 25.0.

OS was calculated as the time from the regorafenib initiation until death for any reason or last follow-up. Progression-free survival (PFS) was defined as the time interval in months from the start of regorafenib to disease progression, death, or last follow-up, whichever occurred first. ROC curve analysis was performed to determine the optimal cut-off value for prediction. In cases where no cut-off value could be established based on ROC analysis, the median value was used instead. Patients were then categorized into low or high PNI and low or high PIV groups according to these cut-off values.

RESULTS

Patient Characteristics and Clinical Findings

A total of 198 patients utilizing regorafenib were found in the retrospective analysis undertaken for this study. A total of 32 patients were excluded from the study due to missing data or early discontinuation of regorafenib secondary to drug intolerance. This study evaluated the efficacy and safety of regorafenib in 166

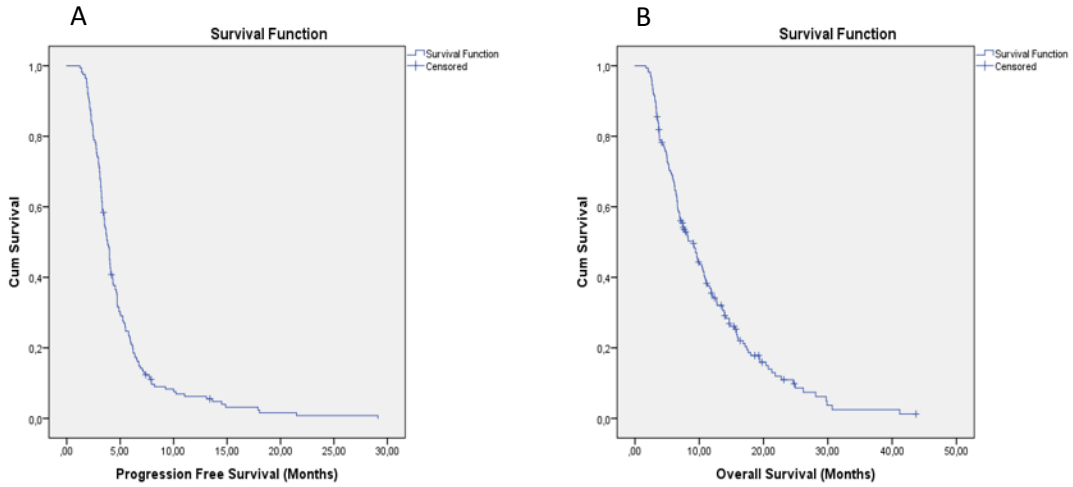

Fig. 1. (A) Kaplan Meir curve progression-free survival (PFS) in the whole cohort (mean PFS: 3.8 months (95% CI: 3.5-4.1), (B) Kaplan Meir curve overall survival (OS) was 9.1 months (95% CI: 7.3-10.9).

patients with metastatic colorectal cancer. The median age of the patients was 58 years (range: 25-82 years), and the majority were male (59%). Most patients (90%) had an ECOG PS of 0-1, indicating a generally favorable functional capacity. The anatomical distribution of the primary tumor sites demonstrated variability, with the majority located in the left colon (44%), followed by the rectum (37%) and the right colon (10%). A substantial proportion of patients (76%) had undergone surgical resection of the primary tumor. Histopathological evaluation revealed adenocarcinoma as the predominant subtype, accounting for 92% of cases. RAS (rat sarcoma) mutations were identified in 49% of patients, while low microsatellite instability (MSI-L) status was observed in 64% of the cohort. Baseline demographic and clinicopathologic findings are outlined in Table 1.

Treatment Characteristics and Outcomes

Treatment characteristics indicated that regorafenib was administered primarily as a third-line therapy in 67% of the cases, with initial doses of 80 mg/day in 43% of patients. Radiologic assessment revealed no complete responses (CR) observed within the study cohort. A partial response (PR) was achieved in 11% of patients, whereas 21% demonstrated stable disease (SD). The overall disease control rate (DCR) was 33%, and the objective response rate (ORR) was 11%.

Table 2. Treatment characteristics and survival outcomes (n=166)

Variables		Data
Treatment line	2. line	6 (4)
	3. line	111 (67)
	≥ 4. line	49 (29)
Regorafenib starting dosage	40 mg/day	13 (9)
	80 mg/day	72 (43)
	120 mg/day	66 (40)
	160 mg/day	15 (9)
Radiologic response	Complete response	0
	Partial response	18 (11)
	Stable disease	35 (21)
	Progressive disease	113 (68)
Response rate (%)	Objective response rate	11
	Disease control rate	33
Progression	Present	161 (97)
	Absent	5 (3)
Progression free survival	Median (months)	3.8 (95% CI:3.5-4.1)
	6. months (%)	21
	12. months (%)	6
Status	Alive	25 (15)
	Exitus	141 (85)
Overall survial	Median (months)	9.1 (95% CI:7.3-10.9)
	6. months (%)	55
	12. months (%)	34

Data are shown as median (95% CI) or n (%) where appropriate. CI=Confidence interval

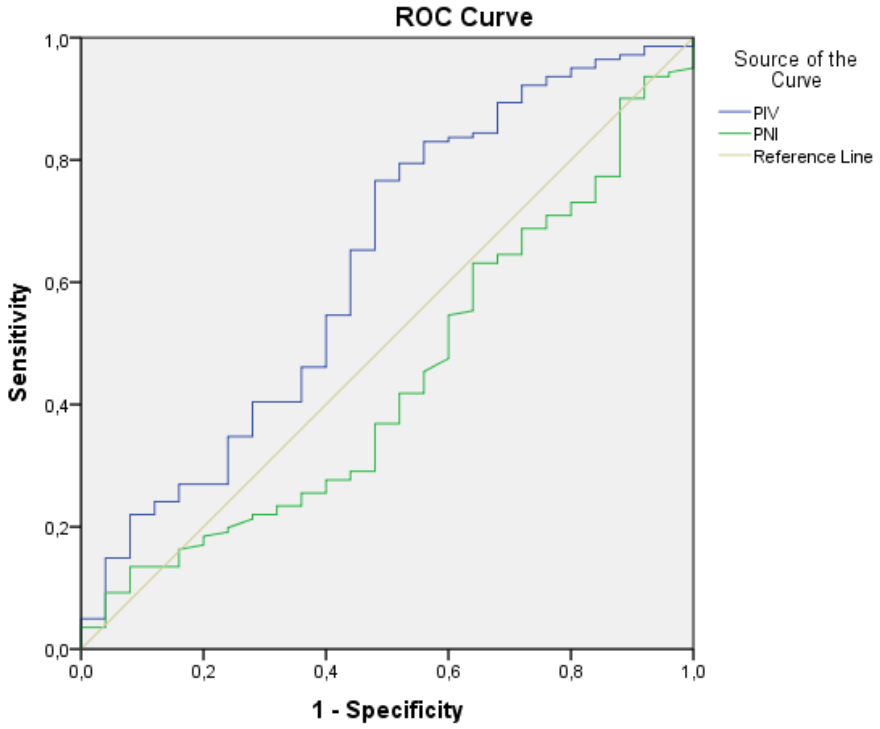

Fig. 2. ROC analysis of pan-immune-inflammation value (PIV) and prognostic nutritional index (PNI).

During the 7-month follow-up period, disease progression was observed in 161 patients (97%), and 141 patients (85%) had died. The median PFS was 3.8 months (95% CI: 3.5-4.1), with 21% of patients remaining progression-free at 6 months and 6% at 12 months. The median OS was 9.1 months (95% CI: 7.3-10.9) with 6 months and 12 months survival rates of 55% and 34%, respectively. Kaplan-Meir analysis curves of PFS and OS are shown in Figs. 1A and 1B. The treatment characteristics and survival outcomes are summarized in Table 2.

Based on the laboratory findings, the calculated median PNI was 47.9 (min: 29.3-max: 66), and the PIV was 402.7 (min: 28.4-max: 9547.9). In the ROC analysis, a significant cut off value could only be determined for PIV, with an Area Under the Curve (AUC) of 0.628 (95% CI: 0.500-0.756; P=0.001). The optimal cutoff value was identified as 309.1, yielding

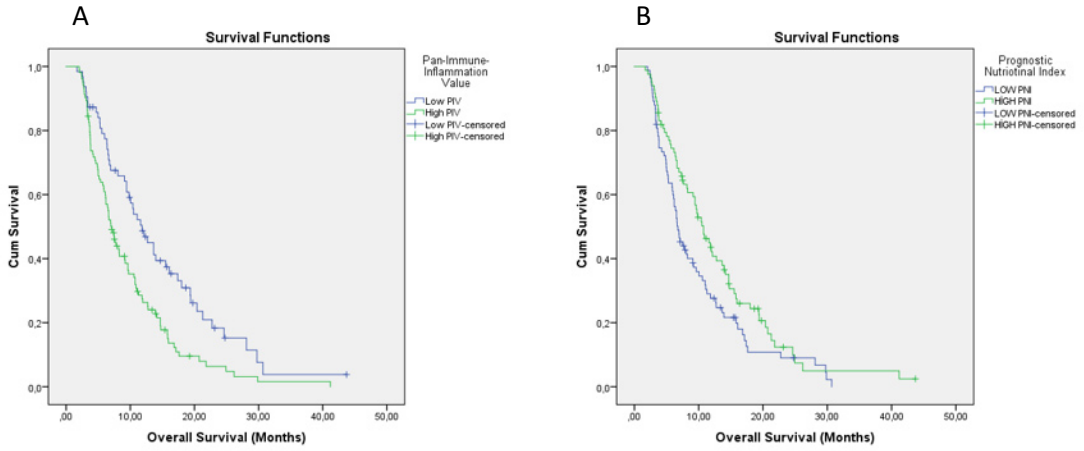

Fig. 3. (A) Overall survival according to Pan-Immune-Inflammation Value (PIV) status (P=0.001), (B) Overall survival according to Prognostic Nutritional Index (PNI) status (P=0.03).

a sensitivity of 65 and a specificity of 56. Based on this threshold, patients were stratified into low and high PIV groups. For the PNI, median values were used as reference since ROC analysis did not yield statistically significant cutoff points. The ROC analysis of PIV and PNI indices is illustrated in Fig. 2.

Cox Proportional Hazards Regression Models For Overall Survival

In the univariate analysis for OS, ECOG performance status, metastatic status, Carcinoma Embryonic Antigen (CEA), Carbohydrate Antigen (CA 19-9), PIV, and PNI were found to be statistically significant (P=0.04, P=0.04, P=0.03, P=0.005, P=0.001, and P=0.03, respectively). However, in the multivariate analysis, only the PIV and PNI indices retained statistical significance (P=0.001 and P=0.02, respectively).

Survival analysis demonstrated that a higher PIV was associated with pooer outcomes, with a median OS of 7.1 months in the high PIV group compared to 11.8 months in the low PIV group (P=0.001; Fig. 3A). Conversely, a high PNI was associated with improved OS, with a median of 10.7 months versus 6.7 months for low PNI (P=0.03; Fig. 3B). The results of the Cox regression analysis for overall survival are summarized in Table 3.

DISCUSSION

This retrospective analysis was conducted to evaluate the efficacy of regorafenib treatment in patients with

Tablo 3. Cox regression overall survival

Variables		Median OS (Months)	Univariate analysis		Multivariate analysis	
			HR (95%CI)	P value	HR (95%CI)	P value
Age (years)	< 65	9.4 (95% CI: 6.9-11.8)	1.2 (0.8-1.8)	0.24		
	≥ 65	8.1 (95% CI: 5.1-11.1)				
Gender	Female	8.2 (95% CI: 6.0-10.4)	1.1 (0.7-1.5)	0.53		
	Male	9.1 (95% CI: 6.3-12.0)				
ECOG PS	PS 0-1	9.4 (95% CI: 7.3-11.4)	1.3 (1.0-1.7)	**0.04**		
	PS≥ 2	4.8 (95% CI: 0.0-10.5)				
Metastatis status	Metachronous	9.4 (95% CI: 6.2-12.6)	0.7 (0.5-0.9)	**0.04**		
	Synchronous	8.0 (95% CI: 5.2-10.9)				
RAS status	Mutant	9.1 (95% CI: 9.1-11.1)	0.8 (0.6-1.2)	0.48		
	Wild	8.1 (95% CI: 5.6-10.7)				
Regorafenib dosing	>80 mg/day	8.2 (95% CI: 5.7-10.7)	0.7 (0.5-1.0)	0.14		
	≤ 80 mg/day	9.4 (95% CI: 6.1-12.7)				
CEA (ng/mL)	< 71.2	10.8 (95% CI: 8.7-13.0)	1.4 (1.0-2.0)	**0.03**		
	≥ 71.2	6.6 (95% CI: 5.3-7.8)				
CA 19-9 (U/mL)	< 65.2	11.8 (95% CI: 8.2-15.3)	1.6 (1.1-2.3)	**0.005**		
	≥ 65.2	6.9 (95% CI: 5.5-8.3)				
PIV	Low PIV	11.8 (95% CI:9.0-14.5)	1.7 (1.2-2.5)	**0.001**	1.9 (1.3-2.7)	**0.001**
	High PIV	7.1 (95% CI: 6.0-8.2)				
PNI	Low PNI	6.7 (95% CI: 5.6-7.8)	0.7 (0.5-0.9)	**0.03**	0.6 (0.4-0.96)	**0.02**
	High PNI	10.7 (95% CI: 8.6-12.8)				

CA 19-9=Carbohydrate Antigen, CEA=Carcinoma Embryonic Antigen, ECOG PS=Eastern Cooperative Oncology Group Performance Status, HR=Hazard Rate, OS=Overall Survival, PIV=Pan-Immune-Inflammation Value, PNI=Prognostic Nutritional Index, RAS=Rat Sarkoma

mCRC using real-world data and to identify potential prognostic biomarkers. Data obtained from 166 patients suggest that regorafenib may offer notable survival benefits in selected patient subgroups, despite its limited objective response rate, especially when administered in later lines of treatment. In our study, the median PFS was 3.8 months, and the median OS was 9.1 months- findings that are consistent with those reported in randomized controlled trials, retrospective trail and meta-analysis of retrospective series [15, 32-36].

The DCR of 33% and ORR of 11% in our study are comparable to those observed in the CORRECT (NCT01103323) and CONCUR (NCT01584830) trials, which reported DCRs of approximately 41% and 51%, respectively. These results highlight the ability of regorafenib to stabilize disease in a subset of patients, even though complete or partial responses are infrequent [15, 32].

The median PFS in our study was 3.8 months (95% CI: 3.5-4.1), with 21% of patients remaining progression-free at 6 months and 6% at 12 months. These findings are aligned with the PFS results from the CORRECT trial (median PFS of 1.9 months) and the CONCUR trial (median PFS of 3.2 months) [15, 32]. The slightly longer median PFS observed in our cohort may be attributed to differences in patient demographics or prior treatment exposures. The median OS of 9.1 months (95% CI: 7.3-10.9), with 6- and 12-month survival rates of 55% and 34%, respectively, is encouraging and consistent with those reported in the CORRECT (6.4 months) and CONCUR (8.8 months) trials [15, 32]. The relatively improved OS in our study might reflect optimized supportive care and/or favorable patient selection.

Additionally, our study found that patients with a PIV of ≥309.1 had significantly worse OS, further emphasizing the prognostic value of this marker. PIV reflects the patient's overall immune status and may also help predict responses to immunotherapy. Unlike other biomarkers, PIV integrates multiple inflammatory markers, enabling a more comprehensive assessment of the immune system and thereby enhancing prognostic prediction. This feature has been supported by a meta-analysis of over ten studies [29]. To our knowledge, no prior studies have directly evaluated the prognostic significance of PIV in mCRC patients treated with regorafenib. However, a study investigating patients with microsatellite instability-high (MSI-H) mCRC receiving immunotherapy found that higher PIV values were associated with poorer PFS and OS [37]. Interestingly, a study examining the prognostic role of PIV by tumor location in colorectal cancer reported that high PIV values correlated with worse disease-free survival (DFS) in left-sided tumors but not in right-sided tumors [38]. Furthermore, a pooled analysis including approximately 4,950 patients across various cancer types, including colorectal cancer, demonstrated that high PIV was significantly associated with increased risk of progression and death [29].

In our study, patients with a PNI of ≥47.9 had significantly longer OS compared to those with lower PNI values, consistent with findings in the literature [26]. A previous study evaluating the prognostic value of PNI in mCRC patients receiving regorafenib similarly reported prolonged OS in patients with high PNI [39]. Another investigation demonstrated that patients with high PNI had significantly longer OS [26], and our findings align with another study indicating a longer median OS in the high-PNI group [31]. In addition to colorectal malignancies, PNI is a significant prognostic indicator for all digestive cancers [40, 41]. A thorough meta-analysis of 14 studies encompassing 3,414 patients with colorectal cancer, gastric cancer, pancreatic cancer, hepatocellular carcinoma, esophageal carcinoma, and malignant pleural mesothelioma revealed that a lower PNI correlated with poorer overall survival results [40]. In patients with advanced hepatocellular carcinoma treated with regorafenib, a similarly low PNI has been identified as a significant predictor of diminished overall survival [41].

Strengths and Limitations

Regorafenib demonstrated modest clinical activity in metastatic colorectal cancer, and inflammatory indices such as PIV and PNI emerged as significant prognostic biomarkers for overall survival. These findings support integrating such inflammatory markers into clinical practice to stratify patients better and tailor treatment strategies. This study is constrained by its retrospective, single-center design, relatively brief follow-up, absence of comprehensive toxicity and dose-modification data, incomplete molecular profiling, treatment-line variability, and the singular evalu-

ation of PIV and PNI, all of which may limit the generalizability of our findings and necessitate validation in prospective multicenter studies.

CONCLUSION

In conclusion, our real-world analysis confirms that regorafenib provides modest clinical benefit in heavily pretreated metastatic colorectal cancer, with disease stabilization observed in a subset of patients. Both the PIV and PNI have been identified as independent predictive indicators for overall survival, highlighting the significant relationship between systemic inflammation and nutritional status in this context. Incorporating publicly accessible indicators into clinical decision-making may assist in identifying patients who are most likely to benefit from regorafenib. Nonetheless, due to the constraints of our retrospective, single-center methodology, these results require validation in larger prospective, biomarker-driven research.

Ethics Approval and Consent to Participate

This study was approved by the Kartal Dr. Lütfi Kırdar City Hospital Scientific Research Ethics Committee (Decision No.: 2024/010.99/3/1 and dated 29.04.2023). All procedures performed during data collection, review of patient records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Informed consent was waived because of the retrospective nature of the study and the analysis used anonymous clinical data.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: SY, HO; Study Design: SY, HŞY; Supervision: SY, AD; Funding: SY; Materials: SY; Data Collection and/or Processing: SY, HŞY, AD; Statistical Analysis and/or Data Interpretation: SY, ÖA; Literature Review: SY, SAE; Manuscript Preparation: SY, ÖA; and Critical Review: SY, ÖA.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024;74(1):12-49. doi: 10.3322/caac.21820.
2. Cervantes A, Adam R, Roselló S, et al; ESMO Guidelines Committee. Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. Ann Oncol. 2023;34(1):10-32. doi: 10.1016/j.annonc.2022.10.003.
3. Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020. CA Cancer J Clin. 2020;70(3):145-164. doi: 10.3322/caac.21601.
4. Cancer Stat Facts: Colon and Rectum Cancer. National Cancer Institute Surveillance Epidemiology and End Results Program website. 2018. https://seer.cancer.gov/statfacts/html/colorect.html
5. Hurwitz H, Fehrenbacher L, Novotny W, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. N Engl J Med. 2004;350(23):2335-2342. doi: 10.1056/NEJMoa032691.
6. Saltz LB, Clarke S, Díaz-Rubio E, et al. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. J Clin Oncol. 2008;26(12):2013-2019. doi: 10.1200/JCO.2007.14.9930.
7. Douillard JY, Oliner KS, Siena S, et al. Panitumumab-FOL-

FOX4 treatment and RAS mutations in colorectal cancer. N Engl J Med. 2013;369(11):1023-1034. doi: 10.1056/NEJMoa1305275.
8. Van Cutsem E, Köhne CH, Láng I, et al. Cetuximab plus irinotecan, fluorouracil, and leucovorin as first-line treatment for metastatic colorectal cancer: updated analysis of overall survival according to tumor KRAS and BRAF mutation status. J Clin Oncol. 2011;29(15):2011-2019. doi: 10.1200/JCO.2010.33.5091.
9. Van Cutsem E, Cervantes A, Adam R, et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. Ann Oncol. 2016;27(8):1386-1422. doi: 10.1093/annonc/mdw235.
10. Nielsen DL, Palshof JA, Larsen FO, Jensen BV, Pfeiffer P. A systematic review of salvage therapy to patients with metastatic colorectal cancer previously treated with fluorouracil, oxaliplatin and irinotecan +/- targeted therapy. Cancer Treat Rev. 2014;40(6):701-715. doi: 10.1016/j.ctrv.2014.02.006.
11. Huang J, Gong C, Jiang Z, et al. Regorafenib monotherapy as the later-line treatment for elderly patients with metastatic colorectal cancer: a multicenter real-world study. J Gastrointest Oncol. 2024;15(5):2166-2177. doi: 10.21037/jgo-24-464.
12. Ciracì P, Studiale V, Taravella A, Antoniotti C, Cremolini C. Late-line options for patients with metastatic colorectal cancer: a review and evidence-based algorithm. Nat Rev Clin Oncol. 2025;22(1):28-45. doi: 10.1038/s41571-024-00965-0.
13. Abou-Elkacem L, Arns S, Brix G, et al. Regorafenib inhibits growth, angiogenesis, and metastasis in a highly aggressive, orthotopic colon cancer model. Mol Cancer Ther. 2013;12(7):1322-1331. doi: 10.1158/1535-7163.MCT-12-1162.
14. Schmieder R, Hoffmann J, Becker M, et al. Regorafenib (BAY 73-4506): antitumor and antimetastatic activities in preclinical models of colorectal cancer. Int J Cancer. 2014;135(6):1487-1496. doi: 10.1002/ijc.28669.
15. Grothey A, Van Cutsem E, Sobrero A, et al; CORRECT Study Group. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. Lancet. 2013;381(9863):303-312. doi: 10.1016/S0140-6736(12)61900-X.
16. Bekaii-Saab TS, Barzi A, Cusnir M. Improving survival in metastatic colorectal cancer through optimized patient selection. Clin Adv Hematol Oncol. 2024;22 Suppl 4(5):1-20.
17. Aoyama T, Yukawa N, Saito A. Clinical Impact of Nutrition and Inflammation Assessment Tools in Colorectal Cancer Treatment. Anticancer Res. 2024;44(4):1335-1351. doi: 10.21873/anticanres.16930.
18. Su YL, Tsai KL, Chiu TJ, et al. Development and Validation of a Novel Serum Prognostic Marker for Patients with Metastatic Colorectal Cancer on Regorafenib Treatment. Cancers (Basel). 2021;13(20):5080. doi: 10.3390/cancers13205080.
19. Liu Y, Lyu J, Bell Burdett K, et al. Prognostic and Predictive Biomarkers in Patients with Metastatic Colorectal Cancer Receiving Regorafenib. Mol Cancer Ther. 2020;19(10):2146-2154. doi: 10.1158/1535-7163.MCT-20-0249.
20. Yan H, Liu J, Zhang Y, et al. Efficacy and safety of regorafenib in the treatment of metastatic colorectal cancer: a retrospective cohort study. J Gastrointest Oncol. 2024;15(3):987-1001. doi: 10.21037/jgo-24-180.
21. Tuomisto AE, Mäkinen MJ, Väyrynen JP. Systemic inflammation in colorectal cancer: Underlying factors, effects, and prognostic significance. World J Gastroenterol. 2019;25(31):4383-4404. doi: 10.3748/wjg.v25.i31.4383.
22. Klintrup K, Mäkinen JM, Kauppila S, et al. Inflammation and prognosis in colorectal cancer. Eur J Cancer. 2005;41(17):2645-2654. doi: 10.1016/j.ejca.2005.07.017.
23. Yu D, Liu J, Meng C, Liu B, Liao J. Pan-immune-inflammation value as a novel prognostic biomarker for digestive system cancers: a meta-analysis. World J Surg Oncol. 2024;22(1):306. doi: 10.1186/s12957-024-03595-z.
24. Cheng HW, Wang T, Yu GC, Xie LY, Shi B. Prognostic role of the systemic immune-inflammation index and pan-immune inflammation value for outcomes of breast cancer: a systematic review and meta-analysis. Eur Rev Med Pharmacol Sci. 2024;28(1):180-190. doi: 10.26355/eurrev_202401_34903.
25. Wang Z, Zhao S, Zhang X, et al. Effect of prognostic nutritional index on laboratory parameters and survival in metastatic colorectal cancer patients treated with fruquintinib: a retrospective study. PeerJ. 2024;12:e18565. doi: 10.7717/peerj.18565.
26. Keskinkilic M, Semiz HS, Ataca E, Yavuzsen T. The prognostic value of immune-nutritional status in metastatic colorectal cancer: Prognostic Nutritional Index (PNI). Support Care Cancer. 2024;32(6):374. doi: 10.1007/s00520-024-08572-6.
27. Yi J, Xue J, Yang L, Xia L, He W. Predictive value of prognostic nutritional and systemic immune-inflammation indices for patients with microsatellite instability-high metastatic colorectal cancer receiving immunotherapy. Front Nutr. 2023;10:1094189. doi: 10.3389/fnut.2023.1094189.
28. Fucà G, Guarini V, Antoniotti C, et al. The Pan-Immune-Inflammation Value is a new prognostic biomarker in metastatic colorectal cancer: results from a pooled-analysis of the Valentino and TRIBE first-line trials. Br J Cancer. 2020;123(3):403-409. doi: 10.1038/s41416-020-0894-7.
29. Guven DC, Sahin TK, Erul E, Kilickap S, Gambichler T, Aksoy S. The Association between the Pan-Immune-Inflammation Value and Cancer Prognosis: A Systematic Review and Meta-Analysis. Cancers (Basel). 2022;14(11):2675. doi: 10.3390/cancers14112675.
30. Luvián-Morales J, González-Trejo S, Carrillo JF, et al. Association of the prognostic nutritional index and overall survival in patients with colorectal cancer: A STROBE compliant retrospective cohort study. Cancer Med. 2019;8(7):3379-3388. doi: 10.1002/cam4.2212.
31. Ucar G, Ergun Y, Acikgoz Y, Uncu D. The prognostic value of the prognostic nutritional index in patients with metastatic colorectal cancer. Asia Pac J Clin Oncol. 2020;16(5):e179-e184. doi: 10.1111/ajco.13328.
32. Li J, Qin S, Xu R, et al; CONCUR Investigators. Regorafenib plus best supportive care versus placebo plus best supportive care in Asian patients with previously treated metastatic colorectal cancer (CONCUR): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol. 2015;16(6):619-629. doi: 10.1016/S1470-2045(15)70156-7.
33. Bekaii-Saab TS, Ou FS, Ahn DH, et al. Regorafenib dose-optimisation in patients with refractory metastatic colorectal cancer (ReDOS): a randomised, multicentre, open-label, phase 2 study. Lancet Oncol. 2019;20(8):1070-1082. doi: 10.1016/S1470-

2045(19)30272-4.
34. Sonbol MB, Benkhadra R, Wang Z, et al. A Systematic Review and Network Meta-Analysis of Regorafenib and TAS-102 in Refractory Metastatic Colorectal Cancer. Oncologist. 2019;24(9):1174-1179. doi: 10.1634/theoncologist.2019-0189.
35. Røed Skårderud M, Polk A, Kjeldgaard Vistisen K, Larsen FO, Nielsen DL. Efficacy and safety of regorafenib in the treatment of metastatic colorectal cancer: A systematic review. Cancer Treat Rev. 2018;62:61-73. doi: 10.1016/j.ctrv.2017.10.011.
36. Mercier J, Voutsadakis IA. A Systematic Review and Meta-analysis of Retrospective Series of Regorafenib for Treatment of Metastatic Colorectal Cancer. Anticancer Res. 2017;37(11):5925-5934. doi: 10.21873/anticanres.12039.
37. Corti F, Lonardi S, Intini R, et al. The Pan-Immune-Inflammation Value in microsatellite instability-high metastatic colorectal cancer patients treated with immune checkpoint inhibitors. Eur J Cancer. 2021; 150:155-167. doi: 10.1016/j.ejca.2021.03.043.
38. Spolverato G, Capelli G, Noel F, Steindler M, Gumbs AA. Pan-immune-inflammation in colon cancer: A prognostic biomarker and the role of tumor location in personalized care. World J Gastrointest Surg. 2025;17(4):101066. doi: 10.4240/wjgs.v17.i4.101066.
39. Erdogan B, Ozcan E, Gokmen I, et al. Relationship between prognostic nutritional index and neutrophil lymphocyte ratio with overall survival in patients with metastatic colorectal cancer receiving regorafenib. J Cancer Res Ther. 2023;19(3):762-767. doi: 10.4103/jcrt.jcrt_1620_21.
40. Sun K, Chen S, Xu J, Li G, He Y. The prognostic significance of the prognostic nutritional index in cancer: a systematic review and meta-analysis. J Cancer Res Clin Oncol. 2014;140(9):1537-1549. doi: 10.1007/s00432-014-1714-3.
41. Rimini M, Yoo C, Lonardi S, et al. Role of the prognostic nutritional index in predicting survival in advanced hepatocellular carcinoma treated with regorafenib. Hepatol Res. 2021;51(7):796-802. doi: 10.1111/hepr.13669.

The European Research Journal 2025;11(6):1078-1087 Original Article

DOI: https://doi.org/10.18621/eurj.1733954 *Medical Oncology*

Comparing nivolumab response between smokers and ex-smokers in advanced non-small cell lung cancer: It is never too late to quit smoking

Mustafa Ersoy

Department of Internal Medicine, Kütahya Health Sciences University, Faculty of Medicine, Kütahya, Türkiye

ABSTRACT

Objectives: Lung cancer is most commonly caused by smoking, and unfortunately, a significant portion of patients continue to smoke even during their treatment. Our study aimed to investigate the impact of current smoking on the treatment response in patients receiving nivolumab.

Methods: This was a retrospective cohort study that compared the treatment responses of patients who continued to smoke during nivolumab therapy and those who had a history of smoking but had quit prior to nivolumab initiation. The study included 55 patients with advanced non-small cell lung cancer who received treatment between 2019 and 2025. All participants had stage 4 disease and had progressed after initial platinum-based combination chemotherapy. The treatment responses were categorized as progressive disease, stable disease, partial response, and complete response, and the differences between current smokers and ex-smokers were analyzed, with the response evaluation conducted according to the PET Response Criteria in Solid Tumors 1.0 guidelines.

Results: The analysis revealed a statistically significant difference in treatment response between ex-smokers and current smokers (P=0.039). Ex-smokers demonstrated superior Objective Response Rates (56.0% vs. 25.0%) and Disease Control Rates (76.0% vs. 33.3%) compared to current smokers, with the difference in Disease Control Rate reaching statistical significance (P=0.042). Furthermore, multivariate logistic regression indicated that current smokers were 3.64 times less likely to achieve an objective response to nivolumab than ex-smokers, a finding that, while borderline significant, suggests a clinically meaningful trend (P=0.084).

Conclusions: Our study demonstrated that continued smoking during nivolumab therapy may negatively impact the treatment response. While more prospective data is needed, the current results and existing literature suggest that smoking cessation is crucial for patients receiving nivolumab, and clinicians should be more vigilant in addressing this issue.

Keywords: Lung cancer, nivolumab, smoking, ex-smoker

Lung cancer remains the leading cause of cancer-related deaths worldwide, attributable to its high incidence and mortality rates [1]. In recent years, immunotherapeutic agents, including the monoclonal antibody nivolumab, have emerged as increasingly integral components of non-small cell lung

Received: July 3, 2025 **Accepted:** August 12, 2025 **Available Online:** August 26, 2025 **Published:** November 4, 2025

How to cite this article: Ersoy M. Comparing nivolumab response between smokers and ex-smokers in advanced non-small cell lung cancer: It is never too late to quit smoking. Eur Res J. 2025;11(6):1078-1087. doi: 10.18621/eurj.1733954

Corresponding author: Mustafa Ersoy, MD., Assist. Prof., Phone: +90 274 260 00 43, E-mail: mustafa.ersoy@ksbu.edu.tr

Available Online at https://dergipark.org.tr/en/pub/eurj

cancer (NSCLC) treatment, demonstrating the potential to prolong patient survival [2-4]. Nivolumab, a fully human IgG4 monoclonal antibody, functions by disrupting the Programmed cell death-1 (PD-1) mediated signaling pathway, thereby restoring the previously suppressed antitumor immune responses [5].

Extensive studies have underscored the dominant role of chronic cigarette smoking in the development of lung cancer, with approximately 87% of cases linked to prolonged smoking [6]. Smoking is a chronic condition driven by both physical nicotine dependence and learned behavioral patterns [7]. Approximately 70% of smokers express a desire to quit, yet the majority of adult smokers attempting cessation often relapse [8]. Disturbingly, studies indicate that more than 50% of cancer patients continue to smoke even after their diagnosis [9, 10]. Persisting smoking during cancer treatment is associated with more severe symptoms, an elevated risk of developing additional primary tumors, and poorer survival outcomes [11, 12]. Furthermore, it may diminish the efficacy of chemotherapy and radiation therapy [13]. Given the potential for smoking to exert an immunosuppressive effect and compromise anti-tumor immunity, the relationship between smoking and immunotherapy warrants close examination, particularly as the increasing incorporation of immunotherapy into cancer treatment regimens has coincided with a concerning trend of patients continuing to smoke while undergoing these novel therapies [14]. Recent studies have demonstrated that cigarette smoke can induce polarization of macrophages toward the M2 phenotype, contributing to an immunosuppressive tumor microenvironment that may hinder the efficacy of immune checkpoint inhibitors such as PD-1/ programmed cell death ligand 1 (PD-L1) blockers [15]. Moreover, tobacco-related mutational signatures have been associated with altered immune infiltration and interferon-γ signaling, suggesting that smoking may modulate both the immunogenicity and responsiveness to immunotherapy in a tumor-type-dependent manner [16]. This study aims to compare the response rates of non-small cell lung cancer patients receiving nivolumab therapy, contrasting those who currently smoke versus those who had quit before treatment.

METHODS

Study Design and Participants

This retrospective study enrolled 55 patients with advanced NSCLC who received treatment at the Kütahya Education and Research Hospital or Kütahya City Hospital between 2019 and 2025. All participants had stage 4 disease and had experienced disease progression following initial platinum-based combination chemotherapy. Patients received nivolumab at a dose of 3 mg/kg every 14 days. Tumor response was evaluated after 8-12 weeks of treatment in both groups. The number of nivolumab cycles administered and the cumulative dose (mg/kg) were recorded for each patient. Similar doses and treatment cycles were applied to both groups.

The inclusion criteria were: (1) being at least 18 years old; (2) receiving regular medical care and follow-up visits; (3) undergoing positron emission tomography-computed tomography imaging (PET-CT) studies conducted and reported at the participating hospitals; (4) having a baseline PET-CT scan before receiving nivolumab; (5) continuing treatment until a PET-CT scan was performed for response evaluation; (6) having their Immune-related adverse events (irAEs) documented and recorded in the side effects forms every 14 days; (7) having their smoking status recorded in the oncology records for each treatment protocol; and (8) being received nivolumab treatment for at least 8 weeks with 4 doses.

The exclusion criteria were: (1) presence of symptomatic brain metastases; (2) having activating mutations in Epidermal growth factor receptor (EGFR), anaplastic lymphoma kinase (ALK), or c-ros oncogene 1 (ROS1); (3) presence of uncontrolled autoimmune disease; (4) never-smokers; and (5) resumed smoking during the treatment process despite declaring that they had quit smoking.

Smoking Status

The patients were categorized into two groups based on their smoking status: current smokers and ex-smokers. The ex-smoker group had their cumulative smoking history (in pack-years) recorded, but no statistical analysis was performed regarding the quantity of smoking. Patients who had quit smoking at least 6 weeks before nivolumab initiation were categorized as ex-smokers, based on evidence indicating that the

immunological effects of smoking begin to attenuate within this period [17]. The current smoking group was further subdivided into six categories based on the average number of cigarettes smoked per day (1-4, 5-14, 15-24, 25-34, 35-44, and ≥45) to enable a more nuanced characterization of smoking intensity among participants. This stratification aimed to reflect the heterogeneity within the current smoker population and potentially explore dose-dependent associations with immunotherapy outcomes. However, due to the limited sample size and the uneven distribution of participants across these subgroups, the statistical power required to perform meaningful comparative analyses within the current smoker strata was not met. Therefore, subgroup-level comparisons were not conducted, and statistical analyses focused on broader comparisons between current smokers and ex-smokers. The current smoker and ex-smoker groups were comparable in terms of patient demographics and clinical characteristics, including age, sex, pathological type, and history of lung radiotherapy.

Pathological Assessment

All biopsy specimens were reviewed by the institutional pathology department to confirm a diagnosis of NSCLC. Although PD-L1 expression is a well-established predictive biomarker for response to nivolumab, it was not required for patient inclusion and was not incorporated into the current analysis. This exclusion was primarily due to limitations in data availability and inconsistency in immunohistochemical testing platforms across contributing pathology units. While the inclusion of PD-L1 status could have enhanced the biological interpretability of treatment outcomes, standardization and comparability of PD-L1 data were not feasible in this retrospective setting. Future prospective research incorporating harmonized PD-L1 assessment protocols may further elucidate its prognostic utility in this population.

Immune-related Adverse Events (irAEs) Assessment

At each hospital visit, patients were questioned about the presence of irAEs, which were documented in their oncology records using standardized reporting forms. The patients' adverse events were evaluated and graded according to the guidelines of the Common Terminology Criteria for Adverse Events, with severity ranging from 1 to 5.

Radiological Assessment

The patients' treatment responses were evaluated using PET-CT imaging. Three nuclear medicine specialists at the institution classified the treatment outcomes into four distinct categories: progressive disease, stable disease, partial response, and complete response. The assessment was based on the PET Response Criteria in Solid Tumors 1.0 guidelines. A complete response was defined as the disappearance of dynamic 2-deoxy-2-fluoro-D-glucose (FDG) uptake in all lesions, while a partial response was a ≥30% decrease and >0.8 unit decrease in the peak standardized uptake value corrected for lean body mass. Stable disease was defined as changes that did not meet the criteria for partial response or progressive disease, and progressive disease was a ≥30% increase and >0.8 unit increase in the peak standardized uptake value corrected for lean body mass, or the appearance of new FDG-avid lesions.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, Version 25.0. Descriptive statistics, including frequencies and percentages, were used to summarize categorical variables. The normality of the data was assessed, and due to non-normal distribution, non-parametric tests were applied. For comparisons between smoking status groups, the Linear-by-Linear Association chi-square test was used for ordinal variables, while the Pearson chi-square test was applied for nominal variables. Objective Response Rate (ORR) and Disease Control Rate (DCR) were calculated and compared across groups. Multivariate logistic regression analysis was conducted to identify independent predictors of objective response, incorporating variables such as Eastern Cooperative Oncology Group (ECOG) performance status, metastatic sites, pathological diagnosis, lung radiotherapy, and smoking status. A P-value of <0.05 was considered statistically significant.

RESULTS

As shown in Table 1, demographic and clinical characteristics were stratified by smoking status. The ex-smoker group (n = 31) exhibited a slightly higher proportion of male patients (90.3%) compared to cur-

Table 1. Demographic and clinical characteristics of the study population stratified by smoking status

Variables	Ex-smoker (n=31)	Current smoker (n=24)
Gender (Male)	28 (90.3%)	21 (87.5%)
ECOG performance status (0/1/2)	6/19/6	6/15/3
Brain metastasis (yes)	13 (41.9%)	9 (37.5%)
Liver metastasis (yest)	12 (38.7%)	12 (50.0%)
Bone metastasis (yes)	13 (41.9%)	10 (41.7%)
Pleural metastasis (yes)	19 (61.3%)	16 (66.7%)
Pathological diagnosis		
AD	14 (45.2%)	11 (45.8%)
SCC	17 (54.8%)	13 (54.2%)
Lung radiotherapy (yes)	11 (35.5%)	8 (33.3%)

Data are shown as n (%). AD=Adenocarcinoma, SCC=Squamous Cell Carcinoma, PD=Progressive Disease, CR=Complete Response, ECOG=Eastern Cooperative Oncology Group

rent smokers (87.5%). ECOG performance status distributions were similar across groups, with the majority scoring 1. Brain metastases were present in 41.9% of ex-smokers and 37.5% of current smokers, while liver and bone metastases showed comparable rates between groups. Pleural metastases were more common among current smokers (66.7%) than ex-smokers (61.3%). Histopathological evaluation revealed a predominance of squamous cell carcinoma (SCC) over adenocarcinoma (AD) in both cohorts. Lung radiotherapy was administered to 35.5% of ex-smokers and 33.3% of current smokers.

The relationship between smoking status and treatment response is presented in Table 2. Among ex-smokers, 12 had progressive disease, 5 had stable disease, 8 had partial response, and 6 had complete response. In the current smoker group, 16 had progressive disease, 2 had stable disease, 5 had partial response, and 1 had complete response. The difference in treatment response between ex-smokers and current smokers was statistically significant (P=0.039).

Table 3 delineates the ORR and DCR stratified by smoking status. Among the ex-smoker cohort, an ORR of 56.0% (14/25 patients) and a DCR of 76.0% (19/25 patients) were observed. Conversely, the current smoker group exhibited an ORR of 25.0% (6/24 patients) and a DCR of 33.3% (8/24 patients). A statistically significant difference was identified in DCR between the two groups (P=0.042), whereas the difference in ORR did not achieve statistical significance (P=0.127).

Multivariate logistic regression was performed to identify independent predictors of objective response to nivolumab. After adjusting for ECOG score, metastatic sites (brain and liver), histological subtype, and history of curative lung radiotherapy, smoking sta-

Table 2. Relationship between smoking status and treatment response

Smoking status	PD	SD	PR	CR	Total (n)
Ex-smoker	12	5	8	6	31
Current smoker	16	2	5	1	24
P value*					**0.039**

PD=Progressive Disease, SD=Stable Disease, PR=Partial Response, CR=Complete Response

P-value calculated using the linear-by-linear association chi-square test. Statistical significance was set at P<0.05

Table 3. Comparison of ORR and DCR by smoking status

Smoking status	ORR (PR+CR)	DCR (SD+PR+CR)
Ex-smoker	14 (56.0%)	19 (76.0%)
Current smoker	6 (25.0%)	8 (33.3%)
P-value*	0.127	**0.042**

Data are shown as n (%). PD=Progressive Disease, SD=Stable Disease, PR=Partial Response, CR=Complete Response, ORR=Objective Response Rate, DCR=Disease Control Rate

P-value calculated using the linear-by-linear association chi-square test. Statistical significance was set at P<0.05

tus demonstrated a borderline significant association with treatment response. Current smokers were 3.64 times less likely to achieve an objective response compared to ex-smokers (odds ratio [OR]: 3.641; 95% confidence interval [CI]: 0.842-15.742; P=0.084). Although this did not meet the conventional threshold for statistical significance, the direction and magnitude of the association suggest a clinically meaningful trend that warrants further exploration in larger prospective cohorts. These findings are shown in Table 4.

irAEs were observed in a total of 10 patients. Among these, the most common irAEs included colitis (n=4), dermatitis (n=3), thyroiditis (n=1), fever (n=1), and pneumonitis (n=1). Of the 10 patients with irAEs, 6 were ex-smokers and 4 were current smokers. When the overall incidence of irAEs was compared between ex-smokers and current smokers, no statistically significant difference was found (P=0.798). Due to the limited number of events per category, statistical comparisons for individual adverse events were not performed. These findings are summarized in Table 5.

DISCUSSION

The literature has long recognized that a significant proportion of patients struggle to quit smoking following a cancer diagnosis and during subsequent treatments [18]. Our study examined patients with stage 4 NSCLC who had received one round of platinum-based chemotherapy, experienced disease progression, and then received an average of 6 doses of nivolumab every 2 weeks. At each nivolumab administration, the patients were queried about their smoking status, and their continued use of cigarettes was documented. The patients in the current smoking cohort persisted in smoking throughout the lengthy treatment period and during the entirety of their nivolumab therapy, despite clinicians being aware of their smoking habits. Upon reviewing the patient records, it was observed that none of the patients had been referred to dedicated smoking cessation clinics or to expert pulmonologists and psychiatrists at the study hospitals, even though such resources were available to provide the necessary

Table 4. Multivariate logistic regression analysis evaluating factors associated with objective response

Predictor	B	SE	Wald	df	P value	Exp(B)	95% CI for Exp(B)
ECOG Score	–0.905	0.610	2.198	1	0.138	0.405	0.122-1.338
Brain metastasis (yes)	–0.149	0.729	0.042	1	0.838	0.862	0.207-3.593
Liver metastasis (yes)	–0.964	0.716	1.815	1	0.178	0.381	0.094-1.550
Pathological diagnosis (SCC)	–0.441	0.720	0.374	1	0.541	0.643	0.157-2.641
Curative lung RT (yes)	–0.482	0.726	0.441	1	0.507	0.617	0.149-2.564
Smoking status (current)	1.292	0.747	2.993	1	0.084	3.641	0.842-15.742
Constant	0.488	1.191	0.168	1	0.682	1.630	-

SCC=Squamous Cell Carcinoma, RT=Radiotherapy, SE=Standard Error, CI=Confidence Interval

Statistical significance was set at P<0.05

Table 5. Immune-related adverse events by smoking status

Adverse event	Total (n)	Ex-smoker (n)	Active smoker (n)	P value
Colitis	4	3	1	-
Dermatitis	3	2	1	-
Thyroiditis	1	0	1	-
Fever	1	1	0	-
Pneumonitis	1	0	1	-
Total irAE cases	10	6	4	0.798

irAE=immune-related adverse event.

Total irAE frequency between groups was compared using the Pearson chi-square test (P=0.798). Statistical significance was set at P<0.05.

psychological and, if required, pharmacological support. The findings suggest that clinicians should take a more proactive role in addressing the issue of smoking cessation, in addition to monitoring the treatment course and managing medication side effects, when caring for patients who continue to smoke during their cancer treatments.

In our study, treatment responses were found to be significantly poorer in patients who continued to smoke compared to those who did not. Smoking can accelerate cancer progression through numerous mechanisms [19]. Hypoxia-inducible factor-1α (HIF-1α) promotes the proliferation, metastasis, angiogenesis, and drug resistance of tumors [20]. It also participates in the psychologic stress-driven progression of tumors [21]. Cigarette smoking inhibits the ubiquitination degradation of HIF-1α, elevating its levels [22]. The N6-methyladenosine (m6A) modification participates in the progression of malignant tumors, including lung cancer [23]. Furthermore, the elevation of methyltransferase 3(METTL3), a regulator of m6A modification, contributes to the development and progression of lung cancers [24]. After exposure to cigarette smoke extract, HIF-1α, via METTL3 regulation of the m6A modification, drives smoking-induced progression of NSCLC through promoting cell proliferation [25]. Additionally, the tumor-promoting activities of nicotine involve the inactivation of tumor suppressors such as retinoblastoma tumor suppressor protein and p53; tobacco smoke components can also induce oncogenic proteins such as c-Myc and K-Ras [26]. Smoking's impact on lung cancer proliferation may be one of the factors that negatively influence treatment response.

A study of 237 non-small-cell lung cancer patients revealed that early-stage non-smoking patients had significantly higher two-year overall survival rates compared to smokers upon initiating chemoradiation [27]. The nicotinic acetylcholine receptor subunit α5 was identified as a contributor to radiation resistance. Researchers found that the addition of nicotine decreased radiosensitivity. Furthermore, they demonstrated that the α5 nicotinic acetylcholine receptor subunit activated E2F transcription factor-mediated pathways, promoting proliferation and survival, thereby conferring resistance to radiation [28]. Collectively, these findings suggest that exposure to tobacco smoke or its constituents, such as nicotine, can enhance resistance to radiotherapy, and that smoking cessation may improve sensitivity to radiation [29]. Cancer stem cells, commonly referred to as tumor-initiating cells, exhibit elevated levels of survival proteins, embryonic stem cell transcription factors, and drug efflux pumps [30, 31]. Studies have found that exposure to nicotine enhances the self-renewal capacity of stem-like cells derived from NSCLC cell lines. Additionally, research has demonstrated that nicotine induces a stem-like phenotype, increasing the proportion of side-population cells with stem-like characteristics and conferring resistance to chemotherapeutic agents [32]. These findings suggest that the induction of a stem-like state is one of the mechanisms by which nicotine confers resistance to chemotherapy drugs [13]. Another proposed mechanism by which smoking affects the efficacy of chemotherapy involves the alteration of chemotherapy agent metabolism, thereby

modulating their systemic levels [33].

The relationship between immunotherapy and NSCLC is more complex than the mechanisms described above that contribute to treatment resistance[34]. Checkpoint inhibitors are known to be significantly more effective in tumors with higher mutational burden [35]. Smoking is known to induce DNA damage and mutations in various oncogenes and tumor suppressor genes, resulting in a higher mutational burden [36, 37]. Therefore, smoking and the resulting higher mutational burden might lead to a better response to immunotherapy [38]. A case-control study comparing the efficacy of first-line chemotherapy and immunotherapy in smokers and never smokers with NSCLC found that never-smokers had a higher risk of disease progression when treated with Pembrolizumab, while never smokers treated with chemotherapy had longer progression-free survival. In contrast, smokers eligible for immunotherapy had better progression-free survival, especially in earlier stages. These findings suggest that smoking status impacts treatment response, with never smokers benefiting more from chemotherapy and smokers showing better outcomes with immunotherapy [34]. Therefore, our study included participants with a history of smoking, either current or former, to demonstrate the potential harm of current smoking, and did not include those who had never smoked.

Although the relationship between smoking history and immunotherapy is known, there is unfortunately insufficient data on the impact of smoking during treatment. While the negative effect of continued smoking in our study patients cannot be definitively attributed to a direct blockade of the nivolumab effect, it is clear that the underlying mechanisms are complex and multifaceted. Nivolumab is a human monoclonal antibody and is not metabolized by cytochrome P450 enzymes or other drug-metabolizing enzymes. Therefore, the induction of these enzymes by smoking, as seen with chemotherapy, is not expected to affect the pharmacokinetics of nivolumab [39]. However, the immunological consequences of smoking may contribute to cancer development and progression through both pro-inflammatory and immunosuppressive effects [40, 41]. The pro-inflammatory and immunosuppressive effects of smoking, including increased mutational load and alterations to the tumor immune microenvironment, have been strongly associated with the tumor response to immunotherapy [42]. Clinical data indicate that smokers with lung cancer often have higher response rates to immune checkpoint inhibitor therapies, likely due to the increased mutational burden and subsequent heightened antigenicity of the tumors [43]. In contrast, smokers with head and neck squamous cell carcinoma tend to have lower response rates to immunotherapy, suggesting that the immunosuppressive effects of smoking may predominate in certain tumor types [44]. It is hypothesized that the dominant effect of smoking in lung cancer is to increase the tumor mutational burden, which can enhance the treatment response to immunotherapies. However, the continued smoking during treatment may unmask the immunosuppressive effects on the tumor immune microenvironment, similar to what is observed in head and neck squamous cell carcinoma. This immunosuppressive effect may then negatively impact the response to nivolumab treatment, highlighting the complex interplay between smoking, tumor biology, and the efficacy of immunotherapies. Further research is needed to fully elucidate these mechanisms and guide personalized treatment strategies for smokers with lung cancer.

Limitations

This single-center retrospective cohort study inherently carries limitations related to sample size and generalizability. Although current smokers and ex-smokers were compared in terms of initial treatment response, the subgroup analysis based on smoking intensity (number of cigarettes per day) was not performed due to insufficient statistical power. Additionally, the study focused solely on early treatment outcomes following immunotherapy initiation and did not include long-term parameters such as progression-free survival or overall survival. The absence of PD-L1 stratification and longer-term clinical endpoints may limit the breadth of conclusions; thus, future multi-center prospective studies with larger cohorts and standardized biomarker assessments are warranted.

CONCLUSION

This study highlights the persistent negative impact of smoking on treatment outcomes in patients with non-small cell lung cancer receiving immunotherapy. De-

spite the transformative role of immune checkpoint inhibitors in improving survival among patients without targetable mutations, continued smoking remains a significant clinical challenge. Although the extent of smoking was not quantitatively assessed, its detrimental influence was evident. Notably, the absence of molecular markers such as PD-L1, HIF-1α, and METTL3 limits mechanistic interpretation and underscores the need for future biomarker-integrated studies. These molecules may play a role in modulating immune response and treatment efficacy. Therefore, smoking cessation should be prioritized as part of routine oncologic care, and clinicians should actively incorporate cessation strategies into treatment planning and follow-up.

Ethics Approval and Consent to Participate

This study was approved by the Kütahya Health Sciences University Non-Interventional Clinical Research Ethics Committee (Decision No: 2025/04-39; date: 11.03.2025). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Since this was a retrospective study, informed consent was not obtained from the participants.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: ME; Study Design: ME; Supervision: ME; Funding: ME; Materials: ME; Data Collection and/or Processing: ME; Statistical Analysis and/or Data Interpretation: ME; Literature Review: ME; Manuscript Preparation: ME; and Critical Review: ME.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

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(3):209-249. doi: 10.3322/caac.21660.
2. Herbst RS, Baas P, Kim DW, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet. 2016;387(10027):1540-1550. doi: 10.1016/S0140-6736(15)01281-7.
3. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. N Engl J Med. 2015;373(2):123-135. doi: 10.1056/NEJMoa1504627.
4. Rittmeyer A, Barlesi F, Waterkamp D, et al; OAK Study Group. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. Lancet. 2017;389(10066):255-265. doi: 10.1016/S0140-6736(16)32517-X.
5. Wang C, Thudium KB, Han M, et al. In vitro characterization of the anti-PD-1 antibody nivolumab, BMS-936558, and in vivo toxicology in non-human primates. Cancer Immunol Res. 2014;2(9):846-856. doi: 10.1158/2326-6066.CIR-14-0040.
6. Walser T, Cui X, Yanagawa J, et al. Smoking and lung cancer: the role of inflammation. Proc Am Thorac Soc. 2008;5(8):811-815. doi: 10.1513/pats.200809-100TH.
7. Babb S, Malarcher A, Schauer G, Asman K, Jamal A. Quitting Smoking Among Adults - United States, 2000-2015. MMWR Morb Mortal Wkly Rep. 2017;65(52):1457-1464. doi: 10.15585/mmwr.mm6552a1.
8. Balbani AP, Montovani JC. Methods for smoking cessation and treatment of nicotine dependence. Braz J Otorhinolaryngol.

2005;71(6):820-827. doi: 10.1016/s1808-8694(15)31254-4.
9. Cataldo JK, Dubey S, Prochaska JJ. Smoking cessation: an integral part of lung cancer treatment. Oncology. 2010;78(5-6):289-301. doi: 10.1159/000319937.
10. Duffy SA, Khan MJ, Ronis DL, et al. Health behaviors of head and neck cancer patients the first year after diagnosis. Head Neck. 2008;30(1):93-102. doi: 10.1002/hed.20665.
11. Peppone LJ, Mustian KM, Morrow GR, et al. The effect of cigarette smoking on cancer treatment-related side effects. Oncologist. 2011;16(12):1784-1792. doi: 10.1634/theoncologist.2011-0169.
12. Videtic GM, Stitt LW, Dar AR, et al. Continued cigarette smoking by patients receiving concurrent chemoradiotherapy for limited-stage small-cell lung cancer is associated with decreased survival. J Clin Oncol. 2003;21(8):1544-1549. doi: 10.1200/JCO.2003.10.089.
13. Chellappan S. Smoking Cessation after Cancer Diagnosis and Enhanced Therapy Response: Mechanisms and Significance. Curr Oncol. 2022;29(12):9956-9969. doi: 10.3390/curroncol29120782.
14. Luo D, Yang D, Cao D, et al. Effect of smoking status on immunotherapy for lung cancer: a systematic review and meta-analysis. Front Oncol. 2024;14:1422160. doi: 10.3389/fonc.2024.1422160.
15. Bianchi F, Le Noci V, Bernardo G, et al. Cigarette smoke sustains immunosuppressive microenvironment inducing M2 macrophage polarization and viability in lung cancer settings. PLoS One. 2024;19(5):e0303875. doi: 10.1371/journal.pone.0303875.
16. Desrichard A, Kuo F, Chowell D, et al. Tobacco Smoking-Associated Alterations in the Immune Microenvironment of Squamous Cell Carcinomas. J Natl Cancer Inst. 2018;110(12):1386-1392. doi: 10.1093/jnci/djy060.
17. Mehta H, Nazzal K, Sadikot RT. Cigarette smoking and innate immunity. Inflamm Res. 2008;57(11):497-503. doi: 10.1007/s00011-008-8078-6.
18. Sardari Nia P, Weyler J, Colpaert C, Vermeulen P, Van Marck E, Van Schil P. Prognostic value of smoking status in operated non-small cell lung cancer. Lung Cancer. 2005;47(3):351-359. doi: 10.1016/j.lungcan.2004.08.011.
19. Schaal C, Chellappan SP. Nicotine-mediated cell proliferation and tumor progression in smoking-related cancers. Mol Cancer Res. 2014;12(1):14-23. doi: 10.1158/1541-7786.MCR-13-0541.
20. Mennerich D, Kubaichuk K, Kietzmann T. DUBs, Hypoxia, and Cancer. Trends Cancer. 2019;5(10):632-653. doi: 10.1016/j.trecan.2019.08.005.
21. Liu H, Yang J, Zhang Y, et al. Psychologic Stress Drives Progression of Malignant Tumors via DRD2/HIF1α Signaling. Cancer Res. 2021;81(20):5353-5365. doi: 10.1158/0008-5472.CAN-21-1043.
22. Xu H, Ling M, Xue J, et al. Exosomal microRNA-21 derived from bronchial epithelial cells is involved in aberrant epithelium-fibroblast cross-talk in COPD induced by cigarette smoking. Theranostics. 2018;8(19):5419-5433. doi: 10.7150/thno.27876.
23. Jiang X, Liu B, Nie Z, et al. The role of m6A modification in the biological functions and diseases. Signal Transduct Target Ther. 2021;6(1):74. doi: 10.1038/s41392-020-00450-x.
24. Lin S, Choe J, Du P, Triboulet R, Gregory RI. The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. Mol Cell. 2016;62(3):335-345. doi: 10.1016/j.molcel.2016.03.021.
25. Yang Y, Cheng C, He B, et al. Cigarette smoking, by accelerating the cell cycle, promotes the progression of non-small cell lung cancer through an HIF-1α-METTL3-m6A/CDK2AP2 axis. J Hazard Mater. 2023;455:131556. doi: 10.1016/j.jhazmat.2023.131556.
26. Pezzuto A, Citarella F, Croghan I, Tonini G. The effects of cigarette smoking extracts on cell cycle and tumor spread: novel evidence. Future Sci OA. 2019;5(5):FSO394. doi: 10.2144/fsoa-2019-0017.
27. Fox JL, Rosenzweig KE, Ostroff JS. The effect of smoking status on survival following radiation therapy for non-small cell lung cancer. Lung Cancer. 2004;44(3):287-293. doi: 10.1016/j.lungcan.2003.11.012.
28. Lin CH, Lee HH, Kuei CH, et al. Nicotinic Acetylcholine Receptor Subunit Alpha-5 Promotes Radioresistance via Recruiting E2F Activity in Oral Squamous Cell Carcinoma. J Clin Med. 2019;8(9):1454. doi: 10.3390/jcm8091454.
29. Shitara K, Matsuo K, Hatooka S, et al. Heavy smoking history interacts with chemoradiotherapy for esophageal cancer prognosis: a retrospective study. Cancer Sci. 2010;101(4):1001-6. doi: 10.1111/j.1349-7006.2009.01466.x.
30. Phi LTH, Sari IN, Yang YG, et al. Cancer Stem Cells (CSCs) in Drug Resistance and their Therapeutic Implications in Cancer Treatment. Stem Cells Int. 2018;2018:5416923. doi: 10.1155/2018/5416923.
31. Perumal D, Pillai S, Nguyen J, Schaal C, Coppola D, Chellappan SP. Nicotinic acetylcholine receptors induce c-Kit ligand/Stem Cell Factor and promote stemness in an ARRB1/β-arrestin-1 dependent manner in NSCLC. Oncotarget. 2014;5(21):10486-10502. doi: 10.18632/oncotarget.2395.
32. Guha P, Bandyopadhyaya G, Polumuri SK, et al. Nicotine promotes apoptosis resistance of breast cancer cells and enrichment of side population cells with cancer stem cell-like properties via a signaling cascade involving galectin-3, α9 nicotinic acetylcholine receptor and STAT3. Breast Cancer Res Treat. 2014;145(1):5-22. doi: 10.1007/s10549-014-2912-z.
33. O'Malley M, King AN, Conte M, Ellingrod VL, Ramnath N. Effects of cigarette smoking on metabolism and effectiveness of systemic therapy for lung cancer. J Thorac Oncol. 2014;9(7):917-926. doi: 10.1097/JTO.0000000000000191.
34. Cortellini A, De Giglio A, Cannita K, et al. Smoking status during first-line immunotherapy and chemotherapy in NSCLC patients: A case-control matched analysis from a large multicenter study. Thorac Cancer. 2021;12(6):880-889. doi: 10.1111/1759-7714.13852.
35. Ritterhouse LL, Gogakos T. Molecular Biomarkers of Response to Cancer Immunotherapy. Clin Lab Med. 2022;42(3):469-484. doi: 10.1016/j.cll.2022.05.004.
36. Husgafvel-Pursiainen K, Ridanpää M, Anttila S, Vainio H. p53 and ras gene mutations in lung cancer: implications for smoking and occupational exposures. J Occup Environ Med. 1995;37(1):69-76. doi: 10.1097/00043764-199501000-00010.
37. Wang X, Christiani DC, Wiencke JK, et al. Mutations in the p53 gene in lung cancer are associated with cigarette smoking

and asbestos exposure. Cancer Epidemiol Biomarkers Prev. 1995;4(5):543-548.
38. Hellmann M, Rizvi N, Wolchok JD, Chan TA. Genomic profile, smoking, and response to anti-PD-1 therapy in non-small cell lung carcinoma. Mol Cell Oncol. 2015;3(1):e1048929. doi: 10.1080/23723556.2015.1048929.
39. Rendon A, Rayi A. Nivolumab. 2024 Feb 28. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan–.
40. Hodge G, Nairn J, Holmes M, Reynolds PN, Hodge S. Increased intracellular T helper 1 proinflammatory cytokine production in peripheral blood, bronchoalveolar lavage and intraepithelial T cells of COPD subjects. Clin Exp Immunol. 2007;150(1):22-29. doi: 10.1111/j.1365-2249.2007.03451.x.
41. Hernandez CP, Morrow K, Velasco C, Wyczechowska DD, Naura AS, Rodriguez PC. Effects of cigarette smoke extract on primary activated T cells. Cell Immunol. 2013;282(1):38-43. doi: 10.1016/j.cellimm.2013.04.005.
42. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science. 2015;348(6230):124-128. doi: 10.1126/science.aaa1348.
43. Carbone DP, Reck M, Paz-Ares L, et al; CheckMate 026 Investigators. First-Line Nivolumab in Stage IV or Recurrent Non-Small-Cell Lung Cancer. N Engl J Med. 2017;376(25):2415-2426. doi: 10.1056/NEJMoa1613493.
44. Ferris RL, Blumenschein G Jr, Fayette J, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2016;375(19):1856-1867. doi: 10.1056/NEJMoa1602252.

The European Research Journal 2025;11(6):1088-1097 Original Article

DOI: https://doi.org/10.18621/eurj.1652950 *Pediatric Neurology*

The side effects of clobazam as add-on therapy in pediatric epilepsy patients

Serap Bilge[1], Sema Nur Taşkın[2], Nevzat Başkaya[3]

[1]Department of Pediatric Neurology, Ministry of Health, Diyarbakır Children's Hospital, Diyarbakır, Türkiye; [2]Department of Pediatric Rheumatology, Ministry of Health, Diyarbakır Children's Hospital, Diyarbakır, Türkiye; [3]Department of Pediatric Allergy and Immunology, Ministry of Health, Diyarbakır Children's Hospital, Diyarbakır, Türkiye

ABSTRACT

Objectives: Epilepsy is a long-term cerebral disorder that accompanies a lifelong predisposition to epileptic seizures, resulting in neurobiological, cognitive, psychological, and social problems. Clobazam is considered a relatively safe and effective anticonvulsant, frequently prescribed for the management of pediatric epilepsy. This study was designed to rigorously evaluate the safety profile of clobazam in pediatric epilepsy, with a particular focus on the incidence and nature of treatment-emergent side effects.

Methods: Patients aged 2 to 17 years, who received clobazam as an adjunct to valproic acid, levetiracetam, or carbamazepine therapy, who were referred to Diyarbakır Children's Hospital between December 2021 and March 2023 were included. We evaluated the safety profile of clobazam in pediatric epilepsy, with a particular focus on the incidence and nature of treatment-emergent side effects.

Results: The study included 100 patients using clobazam as an add-on anti-seizure medication. The mean age was 11.0±4.0 years, with males comprising 48% of the cohort. The average duration of clobazam use was 16.3±7.2 months. Side effects were reported in 30 patients (30%), with the most common being insomnia (10%), agitation (9%), and somnolence (6%). Allergic reactions were the least frequent. No patients experienced enuresis, ataxia, hypertension, hypotension, or constipation in our study. Clobazam was discontinued by 12% of the patients due to side effects.

Conclusions: Clobazam is regarded as safe in pediatric epilepsy patients, with minimal concerns regarding drug interactions and adverse reactions. Although it may cause some side effects, initiating low-dose treatment and gradually increasing the dosage can enhance treatment success.

Keywords: Clobazam, epilepsy, pediatric patients, side effects

Epilepsy is one of the most common neurological disorders that affect the world population. Clobazam is a widely used benzodiazepine indicated as adjunctive therapy for pediatric epilepsy with proven efficacies in reducing seizure frequency in a range of epilepsy syndromes including Lennox-Gastaut syndrome, focal epilepsy [1-3], and generalized epilepsies. However, as with other anti-seizure drugs (ASDs), clobazam use is associated with a range of side effects. Clobazam undergoes extensive hepatic metabolism via both cytochrome P450 (CYP) and non-CYP pathways. Its major metabolite, N-

Received: March 8, 2025 **Accepted:** April 15, 2025 **Available Online:** June 11, 2025 **Published:** November 4, 2025

How to cite this article: Bilge S, Taşkın SN, Başkaya N. The side effects of clobazam as add-on therapy in pediatric epilepsy patients. Eur Res J. 2025;11(6):1088-1097. doi: 10.18621/eurj.1652950

Corresponding author: Serap Bilge, MD., Phone: +90 412 224 57 51, E-mail: sarabsmustafa@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

desmethyl clobazam (nor clobazam), has similar activity as clobazam at Gamma-Aminobutyric Acid (GABAA) receptors and acts as an active antiseizure drug [3, 4].

During prolonged use after 1-month being on clobazam, concentrations of norclobazam, the active metabolite of clobazam, are 8-20-fold greater than clobazam levels, with little clobazam remaining, and fully responsible for seizure control during chronic treatment [5-7]. Common side effects of clobazam in children include somnolence, fatigue, and dizziness. These effects are probably associated with its mechanism of action, which increases GABAergic inhibition in the central nervous system, leading to sedation. Some children have also shown behavioral changes including irritability, aggression, and hyperactivity [1, 6]. Clobazam can occasionally lead to serious adverse events, including severe dermatological reactions and respiratory depression, especially in children with chronic respiratory diseases, or taking other central nervous system (CNS) depressants at the same time. Between 1994 and 2004, only five cases of severe outcomes such as hepatic failure, drug-induced status epilepticus, or death have been reported in association with clobazam use [7-10], reflecting the rarity of such adverse events.

Chronic use gives rise to physical dependence, withdrawal, and tolerance, requiring dose tapering on cessation. Another rare but important side effect is the paradoxical reaction, which leads to increased seizure activity, agitation, or worsened behavior problems. Clobazam is commonly used in combination with other ASDs, which can potentiate its side effects by both pharmacodynamic and pharmacokinetic interactions. For example, co-administration with valproate or lamotrigine may enhance sedation, whereas it is metabolized through CYP3A4, as are many other drugs that could influence drug levels and side effect profiles if they are either inducers or inhibitors. Upcoming studies will prove its efficacy in use by clinicians early as an adjuvant therapy in the treatment of refractory epilepsy and may even be considered as monotherapy in a broad spectrum of epilepsy syndromes [4-7]. The purpose of this study is to investigate clobazam-related adverse effects in children with epilepsy and the parameters that may play a role in initiating these effects.

While clobazam is an effective anticonvulsant for managing epilepsy, it is frequently prescribed alongside other ASDs to enhance therapeutic outcomes. However, the combination of clobazam with these medications can lead to a broader spectrum of side effects, some of which may be exacerbated or compounded. Valproate, for example, is known for its side effects, including weight gain, tremors, gastrointestinal discomfort, sedation, hepatotoxicity, and thrombocytopenia. Levetiracetam, though widely favored for its pharmacokinetic profile, can cause behavioral changes such as irritability, aggression, anxiety, and depression. Similarly, carbamazepine may result in dizziness, drowsiness, ataxia, hyponatremia, and, in rare cases, severe dermatological reactions like Stevens-Johnson syndrome. Many of these side effects overlap with those observed in clobazam, particularly sedation, dizziness, ataxia, behavioral disturbances, and gastrointestinal complaints [11]. Understanding these cumulative side effects is crucial in optimizing patient care, especially in pediatric populations where the management of side effects is essential for long-term treatment adherence and quality of life.

METHODS

The medical files of patients aged 2 to 17 years who were admitted to Diyarbakır Children's Hospital between December 2021 and March 2023 were evaluated retrospectively. Patients were identified by the following ICD-10 codes: G40. 0 (Epilepsy), G40. 1 (Epilepsy), G41. 2 (Complex Partial Epilepsy), G41. 8 (Other Epilepsy), and G41. 9 (Unspecified Epilepsy). Inclusion criteria were: Patients with one of these diagnoses were required to meet the medical criteria of having been prescribed clobazam an add-on ASDs [in addition to valproic acid (VPA), levetiracetam (LEV), and carbamazepine (CMZP) with a reliable seizure record]. The demographic and clinical information obtained included age, gender, weight, blood pressure, seizure type, etiology, and anti-seizure medication duration. Hematological and biochemical parameters were determined at the initiation of clobazam treatment and after one year or at cessation due to side effects (which is the routine protocol of starting any ASDs in our clinics). Magnetic Resonance Imaging (MRI) and electroencephalogram (EEG) evaluations were conducted on all participants as part

of the standard diagnostic workup to assess brain structure and electrical activity, ensuring comprehensive clinical assessment. Only those with normal hemogram and biochemical test results at the start of clobazam therapy were included. Exclusion criteria encompassed any modification in the anti-seizure medication regimen within the past year, the concomitant use of more than one additional drug alongside clobazam, the absence of side effects related to the primary medication, and the presence of systemic or psychiatric disorders. At the end of the selection, patients were divided into two groups based on their response to clobazam. Group 1 (No Side Effects): Patients who tolerated clobazam well and continued treatment without developing significant side effects. Group 2 (With Side Effects): Patients who developed side effects that led to dose adjustment or discontinuation of clobazam therapy. Based on medical records, clobazam dosing was documented as the medium dose (0.5 mg/kg/day) and high dose (1 mg/kg/day or 2×20 mg/day). The classification of epilepsy type and etiology was based on the guidelines of the International League Against Epilepsy (ILAE). The study was approved by the University of Health Sciences Gazi Yaşargil Training and Research Hospital Ethics Committee (Approval Date: 17.05.2023 & Approval No: 417).

Statistical Analysis

All analyses were performed with the IBM SPSS Statistics version 20.0 software package. Categorical variables were expressed as numbers and percentages, whereas continuous variables were summarized as means and standard deviations, with medians and interquartile ranges (IQRs) when appropriate. To compare categorical variables between adverse event groups, either the Pearson chi-square test or Fisher's exact test was used, depending on whether the expected value problem arose. The normality of distribution for continuous variables was confirmed using the Shapiro-Wilk test. The Student's t-test or Mann-Whitney U test was used to compare continuous variables between adverse event groups, depending on whether or not statistical hypotheses were met. To compare baseline and 1-year follow-up measurements, either the paired-sample t-test or the Wilcoxon signed-rank test was used, depending on whether or not the statistical hypotheses were met. The statistical significance level for all tests was set at 0.05.

RESULTS

A total of 100 pediatric patients receiving clobazam as an add-on anti-seizure medication were included in the

Table 1. Variables in patients on clobazam anti-seizure medication

Variables	Data
Age (years)	11.0±4.0 12 (8-14)
Gender, n (%)	
Male	48 (48)
Female	52 (52)
Weight (kg)	34.4±14.9
Clobazam usage period (months)	16.3±7.2 16 (13-21.5)
Dose, n (%)	
Low-Medium	67 (67)
High	33 (33)
First drug, n (%)	
CMZP	50 (50)
LEV	25 (25)
VPA	25 (25)
Semiology, n (%)	
Focal	62 (62)
Generalized	38 (38)
Birth history, n (%)	
Normal	94 (94)
Event	6 (6)
MRI, n (%)	
Normal	84 (84)
Abnormal/Lesion	16 (16)
MRI-lesion, n (%)	
Leucomalasia	9 (57)
Arknoid kist	2 (12)
Arnold Chiari	2 (12)
Thin korpus callosum	2 (12)
Atrophy	1 (6)
Type of epilepsy, n (%)	
Symptomatic-cryptogenic	50 (50)
Idiopathic	37 (37)
Structural	13 (13)
Systolic blood pressure (mmHg)	101.3±13.9
Diastolic blood pressure (mmHg)	62.4±10.7

Data are shown as mean±standard deviation or median (IQR) or n (%). CMZP=carbamazepine, LEV=levetiracetam, MRI=magneting resonance imaging, VPA=valproic acid.

study. The mean age of the patients was 11.0±4.0 years. The cohort was almost evenly distributed by gender, with 48 males (48%) and 52 females (52%).The average duration of clobazam use was 16.3±7.2 months, with a median of 16 months. Regarding dosing, 67 patients (67%) were on a low-to-medium dose, while 33 patients (33%) were receiving a high dose. As for the first-line anti-seizure drugs used in combination with clobazam, 50 patients (50%) were on carbamazepine (CMZP), 25 (25%) on levetiracetam (LEV), and 25 (25%) on valproic acid (VPA). Sixty-two patients (62%) presented with focal seizures, while 38 (38%) had generalized seizures. Birth history was normal in 94% of the patients, with perinatal events reported in 6%. MRI findings were normal in 84% of the cohort, while 16% had abnormal findings or structural lesions (Table 1). At the 1-year follow-up, WBC counts significantly increased, while neutrophil, lymphocyte, and monocyte percentages showed a notable decrease. Other hematological and biochemical parameters, including RBC, hemoglobin, liver enzymes, and thyroid markers, remained stable over time (Table 2). Side effects associated with clobazam were observed in 30% of the pediatric epilepsy patients, with the most common being insomnia (10%), agitation (9%), and somnolence (6%). Less frequent side effects included allergic reactions (1%), visual disturbances (2%), and increased frequency of illness or pyrexia (2%). Insomnia and somnolence were most frequently reported in patients on CMZP,

Table 2. Effect of clobazam on hemogram and biochemical parameters at baseline and 1-year follow-up

Variables	Baseline	1-year follow up	P value
WBC (× 10^9/L)	7.51±2.07	8.69±2.74	**<0.001**
Neutrophil (%)	49.8±12.9	49.0±13.3	**<0.001**
Lymphocyte (%)	40.8±13.4	39.5±13.2	**<0.001**
Monocyte (%)	8.36±2.75	7.58±2.57	**<0.001**
Eosinophil (%)	3.3±2.46 2.6 (1.6-4.5)	3.37±3.03 2.35 (1.3-4.65)	**<0.001**
Basophil (%)	0.467±0.323 0.4 (0.2-0.6)	0.469±0.324 0.4 (0.2-0.6)	0.862
RBC (× 10^{12}/L)	4.67±0.7	4.7±0.73	0.855
RDW-CV (%)	14.1±1.9	14.2±2	0.933
PLT (× 10^9/L)	326.8±119.7	322.4±123.3	0.949
MPV (fL)	9.8±1.9	9.8±1.8	0.945
Hb (g/L)	12.4±1.7	12.4±1.7	0.968
Ferritin (ng/dL)	41.9±13.3	43.7±14.3	0.313
TSH (mlU/L)	2.19±0.79	2.19±0.79	0.986
T4 (µg/dL)	1.26±0.21	1.26±0.21	1.000
CK (U/L)	155.1±84.6 150.5 (78-198)	159.9±82.7 156 (89-211)	0.958
Glucose (mg/dL)	79.6±10.8	79.4±11.1	0.881
ALT (U/L)	20.3±8.1 21 (13-23)	20.1±8.1 19 (13-23)	0.743
Albumin (g/dL)	40.3±2.9	40.3±2.9	0.930

Data are shown as mean±standard deviation or median (IQR). WBC=White blood cells, RBC=Red blood cell, RDW-CV=Red cell distribution width-coefficient of variation, PLT=Platelet, MPV=Mean platelet volume, Hb=Hemoglobin, CK=Creatine kinase, ALT=Alanine transaminase, AST=Aspartate transferase, TSH=Thyroid Stimulating Hormone, T4=Thyroxine

Table 3. Percentage and variation of side effects of clobazam

Variables	Data
Side Effect, n (%)	30 (30)
First drug	
CMZP	12 (40)
LEV	8 (27)
VPA	10 (33)
Agitation, n (%)	9 (9)
First drug	
CMZP	1 (11)
LEV	2 (22)
VPA	6 (66)
Insomnia, n (%)	10 (10)
First drug	
CMZP	5 (50)
LEV	2 (20)
VPA	3 (30)
Allergy, n (%)	1 (1)
First drug	
LEV	1 (100)
Visual (Accommodation), n (%)	2 (2)
First drug	
CMZP	2 (100)
Pyrexia/Increased frequency of illness, n (%)	2 (2)
First drug	
CMZP	1 (5)
LEV	1 (50)
Somnolence, n (%)	6 (6)
First drug	
CMZP	3 (50)
LEV	2 (33)
VPA	1 (17)
Stop side effect, n (%)	12 (12)
Agitation	4 (4)
Insomnia	4 (33)
Increased frequency of illness	1 (8)
Somnolence	3 (25)

Data are shown as n (%). CMZP=carbamazepine, LEV=levetiracetam, VPA=valproic acid.

while agitation was predominantly observed in those receiving VPA, accounting for 66% of cases. Notably, 12% of the cohort discontinued clobazam due to side effects, with insomnia and agitation being the primary reasons for cessation (Table 3). A comparison between Group 1 (patients without clobazam-related side effects) and Group 2 (patients experiencing side effects) revealed that those in Group 2 had significantly higher mean body weight (P<0.001). No significant differences were observed between the groups regarding age, gender, clobazam dose, seizure semiology, epilepsy type, birth history, or MRI findings (Table 4). A comparison of hematological and biochemical parameters at baseline and after 1-year follow-up in relation to side effects of clobazam revealed no significant differences in most blood parameters, including WBC, neutrophils, RBC, and ferritin between the groups. However, an increase in ALT was observed (P<0.001) (Table 5).

DISCUSSION

Clobazam, a benzodiazepine frequently prescribed for epilepsy and anxiety disorders, can lead to several side effects, particularly those related to its sedative properties. As a central nervous system (CNS) depressant, clobazam enhances the effects of GABA, a neurotransmitter that inhibits neural activity. This action results in sedative and hypnotic effects, commonly leading to somnolence (excessive sleepiness), which may manifest as daytime drowsiness, lethargy, and a decline in alertness. These effects can impair daily functioning and cognitive performance, especially during the initiation of therapy or when adjusting the dose. In the long term, tolerance to the sedative effects of clobazam may develop, meaning that the calming effect decreases over time. This can lead to an escalation in dosage and an increased risk of dependence. Conversely, some individuals may experience insomnia as a paradoxical side effect, particularly after they have developed tolerance to the drug's sedative properties. Moreover, rebound insomnia characterized by difficulty falling asleep, frequent awakenings, and reduced sleep quality can occur when clobazam is suddenly discontinued or its dose is rapidly reduced. This is due to the neuroadaptive changes that occur in the brain with chronic use, where the cessation of the drug creates a temporary imbalance in neurotransmitter activity [7].

In a study by Uzunhan *et al*. [12], side effects of clobazam were observed in 18 patients (45%). Among

Table 4. Effects of variables on the side effects of clobazam

	Side effects		P value
	No (Group 1)	Yes (Group 2)	
Age (years)	10.8±3.5 12 (8-14)	11.5±4.9 14 (6-16)	0.477
Gender, n (%)			0.541
Male	35 (%)	13 (43)	
Female	35 (50)	17 (57)	
Weight (kg)	32.4±14.8	39.1±14.4	**0.040**
Clobazam period (months)	17.9±5.4 16.5 (14-22)	12.5±9.4 13.5 (1-18)	**0.022**
Dose, n (%)			0.150
Low-Medium	50 (71)	17 (57)	
High	20 (29)	13 (43)	
First drug, n (%)			0.351
KMZP	38 (54)	12 (40)	
LEV	17 (24)	8 (27)	
VPA	15 (21)	10 (33)	
Semiology, n (%)			0.787
Focal	44 (63)	18 (60)	
Generalized	26 (37)	12 (40)	
Birth history, n (%)			0.361
Normal	67 (96)	27 (90)	
Event	3 (4)	3 (10)	
MRI, n (%)			0.771
Normal	58 (83)	26 (87)	
Lesion	12 (17)	4 (13)	
MRI-lesion, n (%)			0.155
Leucomalasia	8 (67)	1 (25)	
Arknoid kist	2 (17)	0 (0)	
Arnold Chiari	0 (0)	2 (50)	
Thin korpus kallosum	1 (8)	1 (25)	
Atrophy	1 (8)	0 (0)	
Type of epilepsy, n (%)			0.350
Symptomatic-cryptogenic	32 (46)	18 (60)	
Idiopathic	29 (41)	8 (27)	
Structural	9 (13)	4 (13)	
Systolic blood pressure (mmHg)	100.3±13.3	103.8±15.1	0.240
Diastolic blood pressure (mmHg)	63.7±10.7	59.4±10.2	0.066
Stop side effect, n (%)			**<0.001**
No	70 (100)	18 (60)	
Yes	0 (0)	12 (40)	

Data are shown as mean±standard deviation or median (IQR) or n (%). CMZP=carbamazepine, LEV=levetiracetam, MRI=magneting resonance imaging, VPA=valproic acid.

Table 5. Hematological and biochemical variables at baseline and 1-year follow-up

	Baseline values			1-year follow-up values		
	Side effects		P value	Side effects		P value
	No (Group 1)	Yes (Group 2)		No (Group 1)	Yes (Group 2)	
WBC (× 10^9/L)	7.59±2.19	7.32±1.77	0.556	8.77±2.99	8.52±2.07	0.675
Neutrophil (%)	48.9±11.6	51.9±15.6	0.289	48.1±11.9	51.3±16.2	0.281
Lymphocyte (%)	41.8±11.9	38.4±16.4	0.243	40.4±11.7	37.3±16.2	0.283
Monocyte (%)	8.28±2.8	8.54±2.66	0.672	7.5±2.77	7.77±2.04	0.633
Eosinophil (%)	3.21±2.44 2.6 (0.1-12.5)	3.51±2.56 2.6 (0.7-10.2)	0.892	3.29±3.1 2.3 (0.1-14.6)	3.57±2.91 3.05 (0.6-11.9)	0.649
Basophil (%)	0.47±0.33 0.4 (0.02-1.4)	0.46±0.31 0.4 (0.1-1.4)	0.909	0.46±0.3 0.4 (0.02-1.4)	0.5±0.38 0.3 (0.2-1.4)	0.973
RBC (× 10^{12}/L)	4.73±0.6	4.51±0.87	0.152	4.68±0.58	4.74±1.01	0.670
RDW-CV (%)	13.9±1.4	14.7±2.8	0.151	14.1±1.8	14.2±2.3	0.853
PLT (× 10^9/L)	330±114.9	319.7±131.5	0.699	330.4±120.9	303.7±128.8	0.324
MPV (fL)	9.93±2.18	9.42±0.86	0.099	9.82±2	9.61±1.08	0.586
Hb (g/L)	12.4±1.8	12.3±1.3	0.659	12.4±1.7	12.5±1.7	0.704
Ferritin (ng/dL)	42.9±13.2	39.6±13.3	0.266	43.6±14.9	44±13	0.876
TSH (mlU/L)	2.11±0.78	2.39±0.79	0.102	2.16±0.79	2.25±0.8	0.609
T4 (µg/dL)	1.24±0.2	1.31±0.22	0.100	1.26±0.23	1.26±0.17	0.942
CK (U/L)	163.29±86.04 160 (45-453)	136.1±79.16 117 (54-345)	0.115	164.9±85.4 160 (54-453)	148.4±76.3 150.5 (45-324)	0.402
Glucose (mg/dL)	79.8±11.7	79±8.7	0.742	80.8±11.4	76.1±9.7	0.053
ALT (U/L)	20.87±8.25 21.5 (9-45)	19.03±7.64 17.5 (11-45)	0.224	18.9±7.3 17 (9-44)	23.0±9.2 23 (12-45)	**0.026**
Albumin (g/dL)	40,5±2,6	40±3,5	0,421	40.6±1.4	39.4±4.8	0.178

Data are shown as mean±standard deviation or median (IQR). WBC=White blood cells, RBC=Red blood cell, RDW-CV=Red cell distribution width- oefficient of variation, PLT=Platelet, MPV=Mean platelet volume, Hb=Hemoglobin, CK=Creatine kinase, ALT=Alanine transaminase, AST=Aspartate transferase, TSH=Thyroid Stimulating Hormone, T4=Thyroxine

them, only six (33%) discontinued the drug due to adverse effects. The most commonly reported side effect was hyperactivity, followed by sedation. The remaining 22 patients (55%) experienced no side effects. Reasons for discontinuation included sedation, refusal to take the drug due to its taste, irritability, hypersalivation (in two patients), and malaise in one patient [12]. In our study, 30% of the patients had side effects, 10% insomnia, 6 % Somnolence, and 9% agitation. Twelve percent of patients discontinued clobazam due to side effects. (These adverse effects were dose-dependent, and approximately 18% of patients experienced them in mild to moderate severity, while 12% experienced them in heavy severity that the family and doctors had to stop clobazam). Initiating high doses led to expressing more frequent or heavier side effects. No cases of clobazam-induced ataxia, pyrexia, hypertension, worsening of seizures, and nocturnal, enuresis in our pediatric epilepsy study have been seen. Our results were in line with the findings of the study by Uzunhan *et al*. [12].

Although clobazam has a safe profile, it can in-

duce side effects affecting visual functions but is less frequently reported than other side effects such as sedation. Some patients may also experience visual disturbances. These side effects include blurred vision, decreased vision, light sensitivity (photophobia), and eye fatigue. One study reported a four-year-old boy with symptomatic generalized epilepsy who had also been treated with clobazam at a dosage of 0.75 mg/kg/day, developed eye-rolling with episodes of ataxia and back arching, but these were of non-epileptic origin [13-15]. Our two patients experienced blurred vision/accommodation disorders which were reversible while using clobazam, this may be secondary to the muscle relaxant properties of clobazam as a benzodiazepine. Like other drugs in this class, clobazam increases the effect of GABA, the brain's main inhibitory neurotransmitter. This results in sedation, anxiolysis, anticonvulsant effects , and relaxation of muscle. None of our patients experienced eye-rolling as a side effect.

Hematologic adverse effects may occur in patients using clobazam. While infrequent, these adverse impacts could entail changes in blood cell counts, as well as potential modifications in hematological parameters. Joshi *et al*. [16] witnessed an increase in WBC and hemoglobin and a decrease in all other blood count parameters but they were not statistically significant, which needs more extensive evaluation and further monitoring guidelines. Our study observed an increase in WBC, monocyte, and eosinophile, and a decrease in neutrophile, and lymphocyte, but despite both changes being statistically significant, both values were in the normal range. Basophile, RBC, Hg, and PLT values did not change. Pyrexia and an increased frequency of illness were observed in only 2% of the patients. While these changes in neutrophil and lymphocyte counts may suggest a potential association with a higher incidence of illness, further investigation through larger studies is required to confirm this relationship. Patients who reported side effects have higher values of ALT at the end of 1-year follow-up control, this can raise the suspicion of abnormal metabolism of clobazam through cytochrome CYP450 which can be due to genetic factors.

Skin change can be seen in most anti-seizure drugs. Drug-induced hypersensitivity syndrome (DiHS)/drug reaction with eosinophilia and systemic symptoms (DRESS) is reported while using clobazam. It is estimated that more than 750,000 prescriptions of clobazam were filled in 2022 in the United States. Exploring cases of DiHS/DRESS related to clobazam in the Food and Drug Administration Adverse Event Reporting System (FAERS) database and current literature yielded 10 cases associated with clobazam [17, 18]. In addition, four cases of pedal, pitting nonpainful edema, with one of the patients developing anasarca were reported due to clobazam use. Edema was completely regressed 6 weeks after stopping clobazam in all four patients. Another case of edema in the lower extremities was reported by İncecik *et al*. [19]. In our study, allergy was seen in 1% of the patients which was reversible, we think that DiHS/DRESS due to clobazam is under-reported.

Clobazam is also an ASD that is relatively safe and mainly has sedative and anxiolytic effects [15]. However, paradoxical reactions have been reported including agitation, aggression, irritability, and hyperactivity. These effects are more frequent in children, older patients, or people with previous psychiatric disorders. Mechanisms are unclear but may involve GABAergic dysregulation, genetic factors, or dose-dependent excitatory effects. Agitation can also occur as a result of tolerance, withdrawal, or drug interactions [20-22]. Management consists of dose adjustment, slow titration, and alternative therapies. Uzunhan TA *et al*. [12] reported that four patients (10%) had their clobazam dosage reduced due to hyperactivity and behavioral disorders, and irritability was seen as a cause for clobazam cessation in that study. In our study, 6% of the patients experienced agitation. Further research is needed to understand individual susceptibility and optimize treatment strategies to mitigate these adverse effects. In our study, it was found that MRI, EEG findings, age gender, final given dose of clobazam, associated drug (CMZP, VPA, LEV), semiology, birth history, type of epilepsy, didn't play a role in creating side effects. Weight had a role in initiating side effects, which might be due to starting a high dose of clobazam from the first day can lead to having more side effects than starting a low dose and titering it slowly till reaching a high dose over some time.

Limitations

Our study is retrospective. Agitation, insomnia,

and somnolence disturbance were based on family and patient reports and could not be assessed by a valid and reliable scale.

CONCLUSION

In conclusion, while clobazam is generally effective and well-tolerated in pediatric epilepsy, common side effects such as agitation, insomnia, and somnolence may still occur. To mitigate the risk of adverse effects and optimize therapeutic outcomes, it is advisable to initiate clobazam therapy at a lower dose, followed by a gradual dose titration. This cautious approach should be adopted as standard practice when incorporating clobazam into pediatric epilepsy management.

The side effect was seen in people with higher weight. Clobazam was also used for a shorter period in those who had side effects (naturally due to the side effects).

Ethics Approval and Consent to Participate

The study was approved by the University of Health Sciences Gazi Yaşargil Training and Research Hospital Clinical Research Ethics Committee (Decision no.: 417 and date: 26.05.2023). It was conducted in accordance with the ethical standards established in the Declaration of Helsinki and its later amendments or comparable ethical standards. Written informed consent was obtained from all the participants and/or legal guardians.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: SB; Study Design: SB; Supervision: SB; Funding: SB, NB, SNT; Materials: SB, NB, SNT; Data Collection and/or Processing: NB, SNT; Statistical Analysis and/or Data Interpretation: NB, SNT; Literature Review: SB, NB, SNT; Manuscript Preparation: SB and Critical Review: SB, NB, SNT.

Conflict of interest

The author disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the author(s) and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Faulkner MA. Comprehensive overview: efficacy, tolerability, and cost-effectiveness of clobazam in Lennox-Gastaut syndrome. Ther Clin Risk Manag. 2015;11:905-914. doi: 10.2147/TCRM.S55930.
2. Gauthier AC, Mattson RH. Clobazam: A Safe, Efficacious, and Newly Rediscovered Therapeutic for Epilepsy. CNS Neurosci Ther. 2015;21(7):543-548. doi: 10.1111/cns.12399.
3. Ng YT, Collins SD. Clobazam. Neurotherapeutics. 2007;4(1):138-144. doi: 10.1016/j.nurt.2006.11.002.
4. Nagarajan E, Lynch TM, Frawley B, Bunch ME. Tolerability of clobazam as add-on therapy in patients aged 50years and older with drug-resistant epilepsy. Neurol Sci. 2023;44(8):2883-2888. doi: 10.1007/s10072-023-06765-1.
5. Dean L. Clobazam Therapy and CYP2C19 Genotype. 2019 Sep 23. In: Pratt VM, Scott SA, Pirmohamed M, Esquivel B, Kattman BL, Malheiro AJ, editors. Medical Genetics Summaries [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2012–.
6. An P, Liu X, Zhang B. Safety profile of clobazam in the real world: an analysis of FAERS database and systematic review of

case reports. Expert Opin Drug Saf. 2024;23(1):119-128. doi: 10.1080/14740338.2023.2204227.
7. Giarratano M, Standley K, Benbadis SR. Clobazam for the treatment of epilepsy. Expert Opin Pharmacother. 2012;13(2):227-233. doi: 10.1517/14656566.2012.647686.
8. Arya R, Giridharan N, Anand V, Garg SK. Clobazam monotherapy for focal or generalized seizures. Cochrane Database Syst Rev. 2018;7(7):CD009258. doi: 10.1002/14651858.CD009258.pub3.
9. Thurman DJ, Beghi E, Begley CE, et al; ILAE Commission on Epidemiology. Standards for epidemiologic studies and surveillance of epilepsy. Epilepsia. 2011;52 Suppl 7:2-26. doi: 10.1111/j.1528-1167.2011.03121.x.
10-Sankar R. GABA(A) receptor physiology and its relationship to the mechanism of action of the 1,5-benzodiazepine clobazam. CNS Drugs. 2012;26(3):229-244. doi: 10.2165/11599020-000000000-00000.
11. Kaplan YC, Demir O. Use of Phenytoin, Phenobarbital Carbamazepine, Levetiracetam Lamotrigine and Valproate in Pregnancy and Breastfeeding: Risk of Major Malformations, Dose-dependency, Monotherapy vs Polytherapy, Pharmacokinetics and Clinical Implications. Curr Neuropharmacol. 2021;19(11):1805-1824. doi: 10.2174/1570159X19666210211150856.
12. Uzunhan TA, Gor Z. Efficacy and Side Effect Profile of Clobazam in Children with Different Etiologies of Epilepsy from a Single Center. Sisli Etfal Hastan Tip Bul. 2020;54(2):236-244. doi: 10.14744/SEMB.2020.60252.
13. Ng YT, Conry JA, Drummond R, Stolle J, Weinberg MA; OV-1012 Study Investigators. Randomized, phase III study results of clobazam in Lennox-Gastaut syndrome. Neurology. 2011;77(15):1473-1481. doi: 10.1212/WNL.0b013e318232de76.
14. Aylett SE, Cross H, Berry D. Eye rolling as a manifestation of clobazam toxicity in a child with epilepsy. Dev Med Child Neurol. 2006;48(7):612-615. doi: 10.1017/S0012162206001289.
15. Humayun MJ, Samanta D, Carson RP. Clobazam. 2023 Dec 3. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan–.
16. Joshi P, verma S, Sui A, Kuo Y-H, Johnson V, Sultan R. Hematologic Profile Associated With Newer-Generation Anti-Epileptic Medications in the Pediatric Population (P3.283). Neurology. 2014;82(10_Suppl):P3.283. doi: 10.1212/WNL.82.10_supplement.P3.283
17. Kichili NR, Hansen A, Kagithala D, Gonzalez AM, Wilkerson M. Incidence of drug-induced hypersensitivity syndrome associated with levetiracetam and clobazam is likely underreported. Ann Allergy Asthma Immunol. 2024;132(5):652-653. doi: 10.1016/j.anai.2024.02.003.
18. Mathew T, D'Souza D, Nadimpally US, Nadig R. Clobazam-induced pedal edema: "An unrecognized side effect of a common antiepileptic drug". Epilepsia. 2016;57(3):524-525. doi: 10.1111/epi.13316.
19. Incecik F, Ozcanyüz DG. Unusual side effects due to clobazam: a case report with edema of the extremities. Acta Neurol Belg. 2018;118(3):521-522. doi: 10.1007/s13760-018-0901-4.
20. Gimigliano F. Is clobazam monotherapy effective and safe in people with focal or generalized seizures? A Cochrane Review summary with commentary. Dev Med Child Neurol. 2020;62(6):670-672. doi: 10.1111/dmcn.14539.
21. Tolbert D, Larsen F. A Comprehensive Overview of the Clinical Pharmacokinetics of Clobazam. J Clin Pharmacol. 2019;59(1):7-19. doi: 10.1002/jcph.1313.
22. Strzelczyk A, Schubert-Bast S. Psychobehavioural and Cognitive Adverse Events of Anti-Seizure Medications for the Treatment of Developmental and Epileptic Encephalopathies. CNS Drugs. 2022;36(10):1079-1111. doi: 10.1007/s40263-022-00955-9.

The European Research Journal 2025;11(6):1098-1107 Original Article

DOI: https://doi.org/10.18621/eurj.1670072 *Cardiovascular Surgery*

Knowledge levels of obstetricians and gynecologists on deep vein thrombosis

Abdullah Burak Balcı[1], Ömer Faruk Rahman[2], Emrah Dağdeviren[3]

[1]Department of Cardiovascular Surgery, Başakşehir Çam and Sakura City Hospital, İstanbul, Turkiye; [2]Department of Cardiovascular Surgery, Faculty of Medicine, İzmir Bakırçay University, İzmir, Türkiye; [3]Department of Obstetrics and Gynecology, Başakşehir Çam and Sakura City Hospital, Istanbul, Türkiye

ABSTRACT

Objectives: The aim of this study was to evaluate the level of knowledge of obstetricians and gynecologists about deep vein thrombosis (DVT) in pregnancy.

Methods: A cross-sectional questionnaire comprising 12 items was administered to obstetricians and gynecologists employed at a tertiary care hospital between January 2024 and March 2024. Following a reliability analysis, a DVT questionnaire scale was developed, which was subsequently evaluated across three sub-dimensions: (1) General Disease Knowledge (F1), (2) Medical Treatment Knowledge (General) (F2), and (3) Anticoagulant Knowledge Level (F3).

Results: A total of 163 participants were evaluated in the study, with a mean age of 32.33±5.75 years. Among these individuals, 107 (65.6%) were female and 56 (34.3%) were male. Additionally, 72 (44.2%) of the participants were classified as resident physicians. The scores for the sub-dimensions of the DVT scale were as follows: 2.25±1.03 for (F1), 1.33±0.72 for (F2), and 2.53±1.09 for (F3). Notably, (F3) of the resident physicians was found to be significantly lower than that of the specialist/faculty member group, with a P-value of less than 0.017.

Conclusions: The study demonstrated that the knowledge levels of obstetricians and gynecologists fell below the established proficiency threshold. This deficiency is believed to stem from inadequate understanding of the differentiation between superficial and deep vein thrombosis, the management of anticoagulant therapy during pregnancy, and post-thrombotic syndrome. Enhancing awareness in these areas may improve patient outcomes, reduce reliance on cardiovascular surgery consultations, and alleviate clinical workload.

Keywords: Deep vein thrombosis, obstetrics and gynecology, anticoagulants, knowledge assessment, healthcare surveys

Deep vein thrombosis (DVT) involves the formation of clots within deep veins, most frequently affecting the lower limbs, and poses serious health risks due to its potential to cause pulmonary embolism (PE) and post-thrombotic syndrome (PTS) [1]. Both conditions contribute to significant patient morbidity and mortality, highlighting the need for early recognition and treatment [2]. While DVT is more commonly associated with cardiovascular care, it is also a pertinent issue in obstetrics and gynecology

Received: April 4, 2025 **Accepted:** May 16, 2025 **Available Online:** June 13, 2025 **Published:** November 4, 2025

How to cite this article: Balcı AB, Rahman ÖF, Dağdeviren E. Knowledge levels of obstetricians and gynecologists on deep vein thrombosis. Eur Res J. 2025;11(6):1098-1107. doi: 10.18621/eurj.1670072

Corresponding author: Abdullah Burak Balcı, MD., Phone: +90 212 909 60 00, E-mail: burakmed188@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

due to patient exposure to unique risk factors. For instance, pregnancy increases blood clot risks through physiological changes, while treatments like oral contraceptives, hormone replacement therapies, and certain surgical procedures further elevate the likelihood of venous thromboembolism (VTE) [3]. Given these considerations, it is essential for obstetricians and gynecologists to have a thorough understanding of DVT risk factors, diagnostic approaches, and effective treatment protocols. By examining these areas within a specialty that frequently encounters patients with heightened clotting risks, this research strives to enhance patient safety and improve clinical outcomes through more informed DVT awareness and practices [4, 5]. This study aims to assess the current knowledge levels among obstetrics and gynecology specialists and residents concerning DVT management, identifying key gaps that could benefit from targeted educational interventions.

METHODS

Ethical Approval and Methodology

This study was designed as a prospective survey utilizing a two-point Likert-type scale to assess gynecologists' knowledge of DVT. Responses were categorized as correct or incorrect. Ethical approval for the study was obtained from the ethics committee on January 15, 2024 (Ethics Approval No: 2024/00345). A structured questionnaire comprising 12 questions was administered to participants concerning DVT. The study was conducted between January 20, 2024, and March 15, 2024. The questionnaire items were categorized into three domains: general knowledge about DVT, general medical treatment information, and anticoagulant treatment information. The questions were developed in collaboration with cardiovascular surgeons and obstetricians, adhering to the national vascular guidelines published in 2021. The questionnaire was distributed online to obstetricians and gynecologists employed at a tertiary care hospital via Google Forms (Google LLC, Mountain View, CA, USA). The specific questions included in the survey are presented in Table 1.

Data Collection Tools

In this study, a structured, prospective questionnaire was utilized to evaluate the knowledge levels of obstetricians and gynecologists regarding DVT. The questionnaire comprises two principal components: demographic information and an assessment of knowledge pertaining to DVT.

Scale Development

The DVT scale questionnaire consists of 12 questions in total. As a result of the KR-20 reliability analysis, since the item distinctiveness indices (biserial) of the 1st and 3rd questions were below 20%, they were removed from the scale, and their reliability was examined again. The KR-20 coefficient was obtained as 0.586. The scale is of low reliability but can be used as a prototype. The results of the reliability analysis for the DVT scale are given in Table 2. The first-level multifactorial structure of the DVT questionnaire, which consists of 3 sub-dimensions and a total of 10 items, was tested by confirmatory factor analysis (CFA) using the RSP library on the R Project. Since the data did not satisfy the assumption of multiple normality, the construct validity of the model fit was tested with the diagonal-weighted least squares (DWLS) estimation technique. The results of CFA are presented in Fig. 1. As a result of the goodness of fit values obtained (χ^2= 26.641, df=32, P=0.735, RMSEA= 0, CFI=1, SRMR=0.053, χ^2/df=0.833), it was shown that the proposed three-factor model was well compatible with the data and was acceptable. It was found that the three-factor structure for the DVT questionnaire was confirmed. In the DVT questionnaire, the F1 dimension includes questions about general disease knowledge, the F2 dimension includes questions in the medical treatment knowledge (general) category, and the F3 dimension includes questions about anticoagulant treatment knowledge. The questionnaire does not have an overall score and is evaluated on three dimensions. The F1 dimension consists of Q2, Q4, Q5 and Q6 items and a score of 0-4 is obtained from this section. The F2 dimension consists of Q7 and Q8 items, and a score of 0-2 is obtained from this section. The F3 dimension consists of Q9, Q10, Q11 and Q12 items, and a score of 0-4 is obtained from this section. In the survey evaluation, the limit for the level of proficiency was set at 75% of the total score (>3 points for F1, 1.5> points for F2 and >3 points for F3, respectively).

Table 1. Questions asked to participants and their numbers

Question number	Questions
1	Pregnancy is a known risk factor for DVT.
2	Elevated D-dimer levels are a definitive diagnostic criterion for DVT.
3	DVT is a rare cause of pulmonary embolism.
4	The most common chronic complication of DVT is post-thrombotic syndrome.
5	DVT involvement is most frequently observed at the iliac vein level.
6	Thrombosis of the great saphenous and small saphenous veins is considered within the scope of DVT.
7	In cases of pregnancy-induced DVT, anticoagulant therapy should be planned for life.
8	Antiplatelet drugs are preferred in the standard treatment of DVT.
9	Apixaban, dabigatran, edoxaban, and rivaroxaban are new-generation anticoagulant drugs that can be used in the treatment of DVT during pregnancy.
10	Both heparin and low-molecular-weight heparins can be used in the treatment of DVT in pregnant women.
11	Warfarin, a vitamin K antagonist, is an anticoagulant drug that can be used in thetreatment of DVT during pregnancy.
12	For patients receiving low-molecular-weight heparin treatment for DVT, INR tests should be performed regularly to evaluate treatment efficacy.

DVT=Deep Vein Thrombosis, INR=International Normalized Ratio

Table 2. Reliability analysis results for DVT questionnaire

Item	Item Difficulty	Biserial
Q2	0.767	0.631
Q4	0.822	0.539
Q5	0.644	0.617
Q6	0.325	0.603
Q7	0.779	0.643
Q8	0.546	0.675
Q9	0.374	0.712
Q10	0.779	0.331
Q11	0.755	0.574
Q12	0.626	0.719
Number of participants	163	
Number of items	10	
KR-20	0.586	

DVT=Deep Vein Thrombosis, KR=Kuder Richarson coefficient

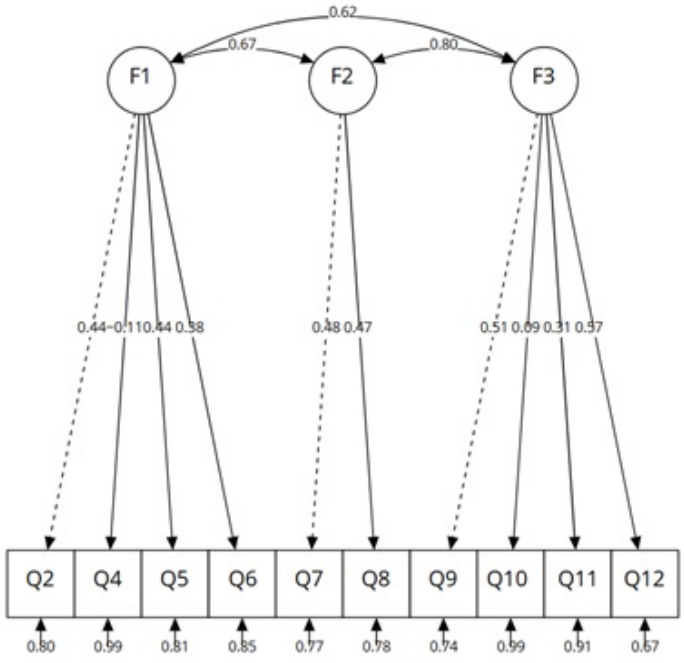

Fig. 1. Confirmatory factor analysis graph of the DVT survey.

Statistical Analysis

The internal consistency and reliability of the DVT questionnaire were evaluated using the KR-20 coefficient. To assess construct validity, CFA was performed utilizing the DWLS estimation technique. The assumption of normality was assessed through kurtosis and skewness coefficients; data were considered to conform to a normal distribution if the kurtosis and skewness coefficients were within the range of ±3. Two independent sample t-tests and one-way analysis of variance (ANOVA) were employed to compare normally distributed survey scores across different groups. The relationship between age, which was normally distributed, and the questionnaire scores was examined using the Pearson correlation coefficient. The analysis findings were obtained using R software (R Core Team, 2024), with results derived from the RSP package included in the R software. The results of the analysis are presented as mean ± standard deviation, median (minimum–maximum), and frequency (percentage). A significance level of $P<0.05$ was established.

RESULTS

A total of 163 physicians participated in the study. Among the participants, 107 (65.6%) were female and 56 (34.4%) were male, with a mean age of 32.33±5.75 years. The median duration of employment at the institution was determined to be 3 years. Of the participants, 81 (49.7%) were specialist physicians, 72 (44.2%) were resident physicians, and 10 (6.1%) were faculty members. Descriptive statistics of the demographic data are presented in Table 3. The question q6 posed to the participants had the lowest response rate at 32.52%. The five questions with the lowest response rates were Q6 (32.52%), Q9 (37.42%), Q4 (51.53%), Q8 (54.60%), and Q12 (62.58%), respectively. The twelve questions included in the questionnaire and their corresponding accuracy rates are illustrated in Fig. 2. Upon examining the mean scores of the participants across the sub-dimensions of the DVT questionnaire scale, the mean score for the general disease knowledge dimension was 2.25, the mean score for the medical treatment information (general) dimension was 1.33, and the mean score for the anticoagulant treatment information dimension was 2.53. No statistically significant relationship was found between the DVT questionnaire sub-dimension scores and the demographic characteristics of the participants ($P>0.05$) (Table 4). No statistically significant relationship was identified between the medical treatment general knowledge dimension score and other demographic characteristics of the participants ($P>0.05$) (Table 5). Additionally, a statistically significant difference was found in the mean scores of the anticoagulant treatment knowledge dimension according to the titles of the participants ($P=0.017$), with resident physicians scoring lower than specialists or faculty members. No statistically significant relationship was noted between the anticoagulant knowledge dimension score and other demographic characteristics of the participants ($P>0.05$) (Table 5).

DISCUSSION

The primary objective of this study was to systematically evaluate the knowledge levels of obstetricians

Table 3. Descriptive statistics on demographic characteristics

	Data (n=163)
Age (years)	32.33±5.75
Gender	
Male	56 (34.4%)
Female	107 (65.6%)
Occupation	
Resident physicians	72 (44.2%)
Specialist	91 (55.8%)
Titles of participants	
Resident physicians	72 (44.2%)
Faculty members	10 (6.1%)
Specialist	81 (49.7%)
Duration of employment in the institution	
0-1 years	34 (20.9%)
1-2 years	31 (19%)
2-3 years	38 (23.3%)
>3 years	60 (36.8%)

Data are shown as mean± standard deviation or n (%)

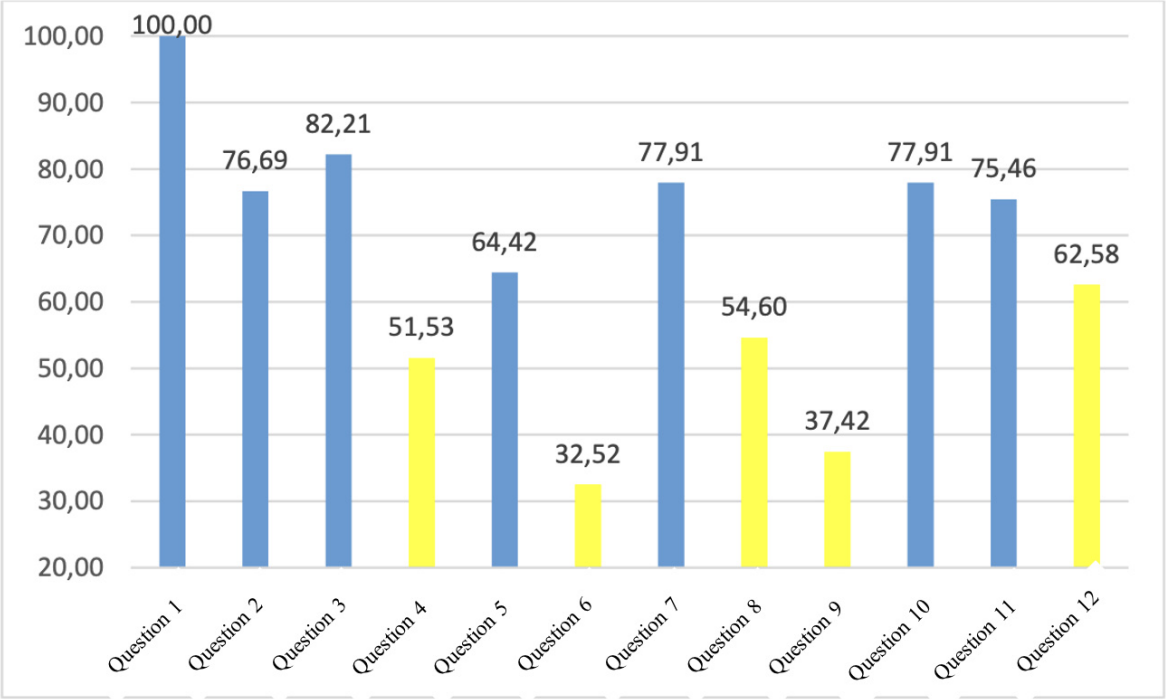

Fig. 2. **Score distribution of questions in the study.**

and gynecologists regarding DVT, with a specific focus on noncardiovascular specialties that frequently encounter DVT cases. This research addresses a critical gap in the extant literature by rigorously assessing knowledge related to the diagnosis, risk factors, and management protocols of DVT within a domain where awareness is paramount yet often insufficiently addressed. The findings of the study revealed several encouraging insights: participants demonstrated a comprehensive understanding of prevalent risk factors and essential treatment modalities, particularly for high-risk populations, including pregnant women, and exhibited significant proficiency in recognizing general diagnostic criteria. In addition, this study is important because it is the first survey conducted on DVT to healthcare professionals in our country. These findings provide a robust foundation, indicating that with targeted educational interventions, these healthcare professionals can further augment their competencies in the effective management of DVT-related complications.

When the average scores of the DVT scale, which consists of three sub-dimensions, were examined in the study, it was seen that all three sub-dimensions were below the determined proficiency level. When the results of the sub-dimension of general disease knowledge were considered, it was thought that the insufficiency here was due to superficial vein thrombosis (Q4) and postthrombotic syndrome (Q6). Our investigation elucidated considerable deficiencies in the comprehension of the differentiation between superficial vein thrombosis (SVT) and DVT. While SVT

Table 4. **Descriptive statistics of DVT questionnaire subdimension scores**

	n=163	
General information	2.25±1.03	2 (0-4)
Medical treatment general information	1.33±0.72	1 (0-2)
Anticoagulant information	2.53±1.09	3 (0-4)

Data are shown as mean± standard deviation or median (minimum-maximum). DVT=Deep Vein Thrombosis

Table 5. Investigation of the relationship between the sub-dimensions of the DVT questionnaire and demographic characteristics

	General information	Medical treatment general information	Anticoagulant information
Gender			
Male	2.3±1.03	1.41±0.68	2.52±1.14
Female	2.22±1.04	1.28±0.74	2.54±1.07
Test stat.	0.464	1.100	-0.134
P[t] value	0.643	0.273	0.893
Occupation			
Resident physicians	2.26±1.02	1.29±0.76	2.31±1.15
Specialist	2.24±1.05	1.35±0.69	2.71±1.01
Test stat.	0.135	-0.528	-2.412
P[t] value	0.892	0.598	**0.017**
Duration of employment in the institution			
0-1 years	2.5±1.13	1.26±0.71	2.5±1.19
1-2 years	2.39±0.95	1.26±0.73	2.55±1.12
2-3 years	2.24±1	1.37±0.79	2.63±1.13
>3 years	2.05±1.02	1.37±0.69	2.48±1.02
Test stat.	1.617	0.279	0.155
P[t] value	0.187	0.841	0.926
Age			
r	-0.062	-0.005	0.106
P value	0.432	0.946	0.180

Data are shown as mean± standard deviation. DVT=Deep Vein Thrombosis
[t]Independent sample t test, [k]One way variance test, [r]Pearson correlation coefficient

is not as acutely life-threatening as DVT, it can nonetheless lead to significant morbidity and typically necessitates a shorter therapeutic regimen. Therefore, the precise distinction between SVT and DVT is of paramount clinical significance to avert unnecessary interventions and to alleviate the burden of consultations [6]. The research conducted by Turton *et al*. [7] in the United Kingdom and Ireland underscored that accurate differentiation between SVT and DVT can substantially diminish clinical workload and optimize treatment management. Their findings indicated that the mismanagement of SVT could exacerbate clinical burdens, particularly in high-demand settings such as metropolitan hospitals, where precise differentiation is critical to prevent unwarranted consultations. Furthermore, our study identified substantial gaps in knowledge regarding the administration of novel oral anticoagulants (NOACs) during pregnancy. It is well-documented that NOACs are generally contraindicated in pregnant populations due to safety concerns, potential teratogenic effects, and their restricted use within this demographic.

The FRONTLINE survey conducted by Kakkar *et al*. [8] elucidated the limited utilization of NOACs in patients with cancer and underscored the associated risks of their use during pregnancy due to potential teratogenic effects. There exists a consensus within the medical community that low-molecular-weight heparin (LMWH) represents the safest therapeutic option for the management of DVT during pregnancy. Our study

corroborated these findings, revealing that the restricted application of NOACs during pregnancy contributes to a pervasive lack of knowledge regarding their use. This highlights the imperative for enhanced educational initiatives concerning NOACs and the necessity for updates to clinical guidelines of their administration in pregnant populations. Furthermore, we identified a significant deficiency in awareness regarding PTS, a chronic complication that can result in permanent damage and markedly diminish quality of life if not addressed promptly. Consequently, early diagnosis and intervention for PTS are of paramount importance.

The National Venous and Arterial Disease Guidelines (2021) elucidated that PTS possesses the potential to induce chronic morbidity, and that early therapeutic intervention may substantially mitigate this risk. The existing body of literature indicates that PTS is more prevalent among patients situated in intensive care units and within high-risk populations, with a failure to initiate timely treatment potentially culminating in severe adverse outcomes [9]. Cook *et al*. [10] accentuated that PTS constitutes a prevalent complication in intensive care settings, which may result in lifelong morbidity if not addressed with alacrity. The findings of the present study corroborate this perspective, underscoring the imperative for heightened awareness of PTS and the necessity for comprehensive educational initiatives aimed at healthcare professionals regarding this condition.

In the evaluation made specifically for the sub-dimension of medical treatment knowledge (general), it was determined that the potential cause of the detected insufficiency was the question about the antiplatelet (Q8). Furthermore, our investigation identified pervasive misconceptions regarding the application of antiplatelet therapy in the management of deep vein thrombosis (DVT). In the study conducted by Kakkar *et al*. [8], it was observed that the utilization of antiplatelet agents, such as aspirin, for the management of DVT is not supported by a robust evidence-based framework, particularly in Eastern Europe, where the prevalence of inappropriate treatment practices is notable. Furthermore, Cook *et al*. [10] emphasized that antiplatelet therapy does not constitute an effective strategy for the management of DVT. Our research corroborated these assertions, concluding that antiplatelet therapy is ineffective in the treatment of DVT, thereby underscoring the imperative for the adoption of appropriate therapeutic interventions. There exists a significant necessity to address the existing knowledge gaps pertaining to antiplatelet therapy and to enhance the educational initiatives aimed at physicians regarding this critical issue.

The reason for the low level of anticoagulant treatment knowledge, which is the third sub-dimension of the scale, was evaluated as the question about INR (Q12). Knowledge deficiencies were observed in this study regarding the necessity of monitoring INR in patients treated with LMWH. While the effects of LMWH on INR are limited, regular monitoring of INR during treatment is important to prevent complications. In the study by Kakkar *et al*. [8] it was emphasized that INR monitoring should be performed alongside LMWH treatment, especially in high-risk groups. Similarly, in this study, it was found that INR monitoring was not adequately performed in patients receiving LMWH, indicating a lack of awareness. This underscores the importance of INR monitoring in preventing complications. Future educational programs should focus on raising awareness about this issue. In addition, in our study, it was observed that the level of anticoagulant therapy knowledge of resident physicians was found to be significantly lower than that of specialist/academic physicians.

The association between iliac vein involvement and the risk of pulmonary embolism was misinterpreted in our study. The research conducted by Cook *et al*. [10] indicated that the risk of pulmonary embolism is markedly elevated in patients exhibiting iliac vein involvement, thereby necessitating meticulous monitoring of these individuals. This study identified a significant gap in knowledge regarding the correlation between iliac vein involvement and the risk of pulmonary embolism, highlighting the imperative for more rigorous management of such cases. It is essential for physicians to be thoroughly informed about the increased risk of pulmonary embolism in patients with iliac vein involvement, as this factor is pivotal in the effective management of these clinical scenarios.
Based on the findings of our study, information sessions were conducted for obstetricians and gynecologists. In comparison to the study by Zhao *et al*. [11], which evaluated nurses' knowledge of DVT, both studies underscored the necessity of addressing knowledge deficiencies related to DVT.

In a study conducted by Alyousef *et al*. [12] in

Saudi Arabia, it was noted that knowledge gaps regarding DVT increased healthcare costs, and such information sessions could improve patient outcomes.

This study demonstrates that obstetricians and gynecologists exhibit significant knowledge deficiencies concerning DVT indicating a pressing need for the implementation of educational programs to address these gaps. Future research should focus on evaluating the impact of such educational initiatives on clinical outcomes.

This study makes significant contributions to the existing literature by highlighting the knowledge deficiencies of obstetricians and gynecologists regarding DVT and emphasizing the areas requiring improvement. The findings indicate that the implementation of appropriate treatment practices could diminish the necessity for cardiovascular surgery consultations. A key conclusion of the study is the imperative for training and awareness programs, particularly in high-demand healthcare settings such as urban hospitals.

Our study shows some similarities and differences compared to other studies in the literature. In the study by Kakkar *et al*. [8] it was emphasized that antiplatelet therapy has no place in DVT management and that evidence-based treatment methods such as LMWH should be preferred. This finding aligns with our study, where antiplatelet therapies were misunderstood in the management of DVT. While antiplatelet therapies are known to be ineffective in DVT treatment, our study observed that the knowledge level regarding these therapies was low. Additionally, in the study by Cook *et al*. [10] it was stated that early treatment and appropriate anticoagulation are vital in DVT management, which is also consistent with our findings. It is known that incorrect treatment practices in DVT increase morbidity and mortality risks.

In the study by Zhao *et al*. [11] it was stated that if PTS is not treated early, it can lead to permanent damage and significantly impact quality of life. Similarly, this study found that the importance of post-thrombotic syndrome was not sufficiently understood. In our study, it was emphasized that post-thrombotic syndrome should be addressed as a permanent complication and that early treatment can prevent the development of this syndrome. The National Venous and Arterial Disease Guidelines (2021) also suggest that more attention should be given to PTS in DVT treatment [9].

Among the risk factors for DVT during pregnancy are the use of oral contraceptives, immobility, and genetic factors. In the study by Atilgan *et al*. [13] which evaluates the incidence of DVT based on age and gender in the Central Anatolia region in light of regional factors, the effects of age and gender on DVT development were examined. It was particularly noted that DVT incidence increases in older individuals and that post-menopausal women have a DVT incidence similar to men. This study is consistent with our findings, as it was observed that obstetricians and gynecologists tend to misjudge the risk of DVT during pregnancy and have a lack of knowledge about new oral anticoagulants. Our study also highlighted the need to prefer treatment methods such as LMWH during pregnancy. In the study by Kakkar *et al*. [8] it was stated that the risk of pulmonary embolism is higher in patients with iliac vein involvement. Similarly, in the study by Cook *et al*. [10] it was noted that the risk of pulmonary embolism is higher in patients with iliac vein thrombosis, and this can lead to fatal outcomes. In this study, it was found that the risk of pulmonary embolism in patients with iliac vein involvement is not sufficiently known. Although this relationship is clearly stated in the literature, our study observed that physicians do not have enough knowledge about it. This indicates that patients with iliac vein thrombosis should be closely monitored.

This study makes a significant contribution to the existing body of knowledge by revealing that obstetricians and gynecologists possess a limited understanding of deep vein thrombosis (DVT). It suggests the implementation of educational programs to address these knowledge gaps. Similarly, Zhao *et al*. [11] emphasized the importance of training programs in DVT management. These findings are consistent with our study's results, indicating that physicians in high-demand settings, such as urban hospitals, require additional training to manage thrombosis cases more effectively.

The findings of this study are largely consistent with the literature and show that the knowledge deficiencies of obstetricians and gynecologists in DVT management need to be addressed. The literature indicates that early diagnosis and treatment of DVT are crucial for preventing complications.

These studies can serve as a foundation for future awareness initiatives and educational programs in major

healthcare institutions, such as urban hospitals. As highlighted by Cayley *et al*. [14], the appropriate implementation of DVT prophylaxis can significantly mitigate the risk of both fatal and non-fatal complications.

Limitations

This study has several limitations. Firstly, it relied on a self-reported questionnaire format, which may be subject to response bias and may not fully capture the actual clinical knowledge or decision-making practices of participants. Additionally, the study was conducted primarily in urban, high-volume centers, potentially limiting the generalizability of findings to rural or lower-volume healthcare settings. The study also focused on knowledge levels rather than clinical practice outcomes, leaving the impact of knowledge gaps on patient outcomes unmeasured. Future research could address these limitations by incorporating broader sample populations and evaluating how knowledge levels directly influence clinical practices and patient outcomes.

CONCLUSION

This study sought to evaluate the knowledge levels of obstetricians and gynecologists regarding DVT, particularly within high-demand healthcare settings such as tertiary centers. The results identified significant knowledge deficiencies, especially in differentiating between superficial and deep vein thrombosis, the application of anticoagulants during pregnancy, and the management of post-thrombotic syndrome. Addressing these gaps through targeted educational initiatives could improve patient outcomes, reduce the reliance on cardiovascular surgery consultations, and ease the clinical burden in high-volume centers. This study highlights the imperative of enhancing awareness and providing adequate training for healthcare professionals in DVT management.

Ethics Approval and Consent to Participate

This study was approved by the Başakşehir Çam and Sakura City Hospital Clinical Research Ethics Committee (Decision no.: 2024/00345, date: 15.01.2024). All procedures performed during data collection, review of patient records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Informed consent was waived because of the retrospective nature of the study and the analysis used anonymous clinical data.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: ABB; Study Design: ÖFR; Supervision: ABB; Funding: ABB; Materials: ED; Data Collection and/or Processing: ED; Statistical Analysis and/or Data Interpretation: ABB; Literature Review: ABB; Manuscript Preparation: ABB and Critical Review: ABB.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Anderson FA Jr, Spencer FA. Risk factors for venous thromboembolism. Circulation. 2003;107(23 Suppl 1):I9-16. doi: 10.1161/01.CIR.0000078469.07362.E6.
2. Heit JA. Epidemiology of venous thromboembolism. Nat Rev Cardiol. 2015;12(8):464-474. doi: 10.1038/nrcardio.2015.83.
3. James AH. Venous thromboembolism in pregnancy. Arterioscler Thromb Vasc Biol. 2009;29(3):326-331. doi: 10.1161/ATVBAHA.109.184127.
4. Pomp ER, Lenselink AM, Rosendaal FR, Doggen CJ. Pregnancy, the postpartum period and prothrombotic defects: risk of venous thrombosis in the MEGA study. J Thromb Haemost. 2008;6(4):632-637. doi: 10.1111/j.1538-7836.2008.02921.x.
5. Skeith L. Prevention and management of venous thromboembolism in pregnancy: cutting through the practice variation. Hematology Am Soc Hematol Educ Program. 2021;2021(1):559-569. doi: 10.1182/hematology.2021000291.
6. Zullo F, Di Mascio D, Berghella V. Evidence-based labor management: postpartum care after vaginal delivery (part 6). Am J Obstet Gynecol MFM. 2023;5(7):100977. doi: 10.1016/j.ajogmf.2023.100977.
7. Turton EP, Coughlin PA, Berridge DC, Mercer KG. A survey of deep venous thrombosis management by consultant vascular surgeons in the United Kingdom and Ireland. Eur J Vasc Endovasc Surg. 2001;21(6):558-563. doi: 10.1053/ejvs.2001.1317.
8. Kakkar AK, Levine M, Pinedo HM, Wolff R, Wong J. Venous thrombosis in cancer patients: insights from the FRONTLINE survey. Oncologist. 2003;8(4):381-388. doi: 10.1634/theoncologist.8-4-381.
9. National Treatment Guidelines for Peripheral Arterial and Venous Diseases. Editor: A. Kürşat Bozkurt. Bayçınar Medical Publishing and Advertising Trade Ltd. Co.:İstanbul. 2021.
10. Cook D, Meade M, Guyatt G, et al.; Canadian Critical Care Trials Group. Clinically important deep vein thrombosis in the intensive care unit: a survey of intensivists. Crit Care. 2004;8(3):R145-52. doi: 10.1186/cc2859.
11. Zhao Y, Li B, Qin H, et al. Knowledge of deep vein thrombosis among nursing interns: A cross-sectional study. Nurse Educ Pract. 2024;75:103915. doi: 10.1016/j.nepr.2024.103915.
12. Alyousef HA, Badawi SEA, Elghoneimy YA, Alameri RA, Almutairi AM. Nurses' Knowledge and Practices Regarding the Prevention of Deep Vein Thrombosis in Saudi Arabia: Descriptive Cross-Sectional Study. Med Arch. 2022;76(6):447-453. doi: 10.5455/medarh.2022.76.447-453.
13. Atılgan K, Demirdaş E, Er CZ, Çiçekçioğlu F. Evaluation of deep vein thrombosis incidence with respect to age and gender in light of regional factors in central Anatolia: A population-based study. J Surg Med. 2020;4(1):62-65. doi: 10.28982/josam.643503.
14. Cayley WE Jr. Preventing deep vein thrombosis in hospital inpatients. BMJ. 2007;335(7611):147-151. doi: 10.1136/bmj.39247.542477.AE

The European Research Journal 2025;11(6):1108-1116 Original Article

DOI: https://doi.org/10.18621/eurj.1741967 *Perinatology*

Prenatal diagnosis and postnatal outcomes of absent pulmonary valve syndrome: A case series with genetic and hemodynamic insights

Tuğçe Arslanoğlu[1], Verda Alpay[2], İsa Özyılmaz[3]

[1]Department of Perinatology, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye; [2]Department of Perinatology, Başakşehir Çam and Sakura City Hospital, İstanbul, Türkiye; [3]Department of Pediatric Cardiology, Başakşehir Çam and Sakura City Hospital, İstanbul, Türkiye

ABSTRACT

Objectives: This study therefore aims to determine the perinatal prognosis and delineate the key risk factors associated with outcomes in fetuses with a prenatal diagnosis of absence of pulmonary valve syndrome (APVS), with particular emphasis on Doppler ultrasound parameters, the presence of extracardiac anomalies, and comprehensive genetic findings - including rare monogenic mutations - as significant contributors to the observed perinatal course.

Methods: This retrospective study included eight fetuses diagnosed with absent pulmonary valve syndrome (APVS) between 2020 and 2024 at a tertiary perinatology referral center. One patient with major extracardiac anomalies was electively terminated and excluded from the outcome analysis. For the remaining seven fetuses, detailed fetal echocardiographic assessments—including cardiac anatomy and Doppler hemodynamic parameters - were evaluated alongside genetic testing results (prenatal and/or postnatal), associated extracardiac anomalies, and postnatal clinical and surgical outcomes.

Results: Among eight fetuses prenatally diagnosed with APVS, one case was electively terminated due to major extracardiac anomalies and excluded from further analysis. All of the remaining seven cases resulted in live births. Four neonates underwent surgical intervention, three of whom survived postoperatively, yielding a surgical survival rate of 75%. Two fetuses that developed hydrops fetalis died in the early postnatal period before surgery could be performed. The overall perinatal mortality rate was 57.1%. Clinically significant genetic anomalies, including trisomy 21, 22q11.2 deletion, and a novel ABAT gene mutation detected via prenatal whole-exome sequencing, were identified in three patients (42.9%). Nonsurvivors were more likely to present with an absent ductus arteriosus and severely dilated pulmonary arteries.

Conclusions: Our study highlights that prognosis is more strongly influenced by prenatal hemodynamic markers - such as pulmonary artery velocities, ductus arteriosus status, and hydrops - than by anatomic subtype. The identification of both common chromosomal anomalies and novel ABAT gene mutations underscores the value of comprehensive genetic evaluation.

Keywords: Absent pulmonary valve syndrome, Fallot tetralogy, fetal echocardiography, ABAT gene mutation, hydrops fetalis

Received: July 14, 2025 **Accepted:** August 18, 2025 **Available Online:** August 28, 2025 **Published:** November 4, 2025

How to cite this article: Arslanoğlu T, Alpay V, Özyılmaz İ. Prenatal diagnosis and postnatal outcomes of absent pulmonary valve syndrome: A case series with genetic and hemodynamic insights. Eur Res J. 2025;11(6):1108-1116. doi: 10.18621/eurj.1741967

Corresponding author: Tuğçe Arslanoğlu, MD., Assist. Prof., Phone: +90 212 404 15 00, E-mail: drtugcetunc@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

Absent pulmonary valve syndrome (APVS) is a rare anomaly constituting approximately 0.2–0.4% of congenital heart diseases, and its incidence in live births is reported to be 1 in 20,000–33,000 [1]. Although it is most commonly associated with tetralogy of Fallot (TOF), it may also occur with other complex cardiac anomalies or rarely in isolation [2]. The distinctive features of the disease are the rudimentary or complete absence of pulmonary valve leaflets. This leads to free pulmonary regurgitation and various degrees of pulmonary stenosis. As a result, marked pulmonary artery dilatation may cause airway obstruction, polyhydramnios, hydrops fetalis and high perinatal mortality by creating pressure on the tracheobronchial system [2, 3].

Prenatal diagnosis of APVS is critically important in terms of family counseling, risk stratification and postnatal management planning. Although the diagnosis is usually made in the second trimester, earlier diagnosis has become possible with the development of high-resolution ultrasound and fetal echocardiography technologies [3]. Although surgical and intensive care practices have improved in the postnatal period, the prognosis of this syndrome is still poor, and high mortality rates have been reported [4].

A strong association has been found between APVS and genetic disorders, including 22q11.2 deletion (DiGeorge syndrome) and trisomy 21. Therefore, recommending genetic testing in the prenatal period is important both diagnostically and prognostically. Detailed fetal evaluation is not only a diagnostic tool but also a prognostic tool [5].

This study aimed to evaluate the clinical features, associated cardiac and extracardiac anomalies, genetic findings and postnatal outcomes of fetuses prenatally diagnosed with APVS in a tertiary referral center. Absent pulmonary valve syndrome is a very rare disease and has been reported in the literature, usually as small series or case reports [6]. In this context, with this retrospective analysis, we aimed to contribute to the knowledge on APVS and to draw attention to the effect of systematic prenatal evaluation on perinatal prognosis.

METHODS

This retrospective case series study was conducted in a tertiary referral hospital in Istanbul, Turkey, using the joint databases of the Perinatology and Pediatric Cardiology clinics. The study included fetuses prenatally diagnosed with absent pulmonary valve syndrome (APVS) between January 2020 and June 2024.

The patients included in the study were selected from those diagnosed prenatally with APVS by an experienced perinatologist via fetal echocardiography and confirmed postnatally by a pediatric cardiologist. The inclusion criteria were as follows: (i) complete ultrasonography and fetal echocardiography data and (ii) postnatal follow-up results on file. Patients with uncertain diagnoses or missing data were excluded. All the data from the hospital information management system were retrospectively reviewed. Demographic information, including maternal age, gravidity, parity, body mass index, gestational week at diagnosis, amniotic fluid volume, Doppler parameters, associated anomalies, and postnatal outcomes, was recorded. Echocardiographic examinations included measurements of the main pulmonary artery (MPA), right pulmonary artery (RPA), and left pulmonary artery (LPA) diameters, peak systolic velocity (PSV), pulmonary regurgitation, and stenosis findings. Associated cardiac anomalies (e.g., TOF, double outlet right ventricle [DORV], and arch anomalies) were recorded. Fetal biometric measurements and Doppler examinations were also performed.

In the postnatal period, birth weight, 5-minute Apgar scores, the need for intensive care, the need for mechanical ventilation, and surgical interventions were analyzed; postoperative survival and mortality outcomes were followed from birth to discharge or death.

All ultrasonographic and Doppler evaluations were performed jointly by the same experienced perinatologist and pediatric cardiologist. A Fujifilm Arietta 850 (Fujifilm Healthcare, Tokyo, Japan) ultrasonography system and a convex abdominal transducer (LISENDO™ C251) were used for imaging. Fetal biometry (biparietal diameter [BPD], head circumference [HC], abdominal circumference [AC], femur length [FL]) was performed as part of routine evaluation, and estimated fetal weight (EFW) was calculated via the Hadlock formula. The amniotic fluid index (AFI) was measured, and a detailed fetal anatomical scan was performed according to the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) guidelines.

All cardiac assessments were performed jointly by two separate experts experienced in perinatology and pediatric cardiology. The following basic cardiac sections were systematically obtained and analyzed:

(1) Four-chamber view: Right ventricular hypertrophy and ventricular septal defects (VSDs) were investigated.

(2) Right ventricular outflow tract (RVOT): Failure to visualize the pulmonary valve and aneurysmatic dilatation of the pulmonary artery were evaluated.

(3) Three-vessel trachea (3VT) section: Posterior compression of the trachea by dilated pulmonary arteries was investigated.

(4) Color Doppler: A "to-and-fro" flow pattern was observed at the pulmonary artery level.

The main, right and left pulmonary artery diameters were measured, and z scores were calculated according to gestational week. Pulmonary artery PSV was recorded, and pulmonary flow hemodynamics were evaluated.

Invasive prenatal genetic diagnosis was recommended for all patients. The methods used included conventional karyotyping after amniocentesis or cordocentesis, Fluorescence in situ hybridization (FISH) analysis for 22q11.2 deletion and chromosomal microARRAY (CMA). Next-generation sequencing (NGS)-based analysis of a large panel of genes, including the 4-aminobutyrate aminotransferase (ABAT) gene, was performed in selected cases.

The study was conducted in accordance with the principles of the Declaration of Helsinki, and ethical approval was obtained from the Clinical Research Ethics Committee of the Başakşehir Çam and Sakura City Hospital 1 (Decision No: 2024-63). Written informed consent was obtained from the parents of all patients to use the data for scientific purposes.

Statistical Analysis

Statistical analyses were performed via IBM SPSS Statistics v26.0 (IBM Corp, Armonk, NY, USA). Continuous variables are presented as the means ± standard deviations or medians (min–max); categorical variables are presented as frequencies and percentages (%). Owing to the limited sample size, descriptive analyses were performed, and comparative statistical tests (e.g., P values) were not included.

RESULTS

In this retrospective series, a total of 8 pregnant women with a prenatal diagnosis of APVS were evaluated. One

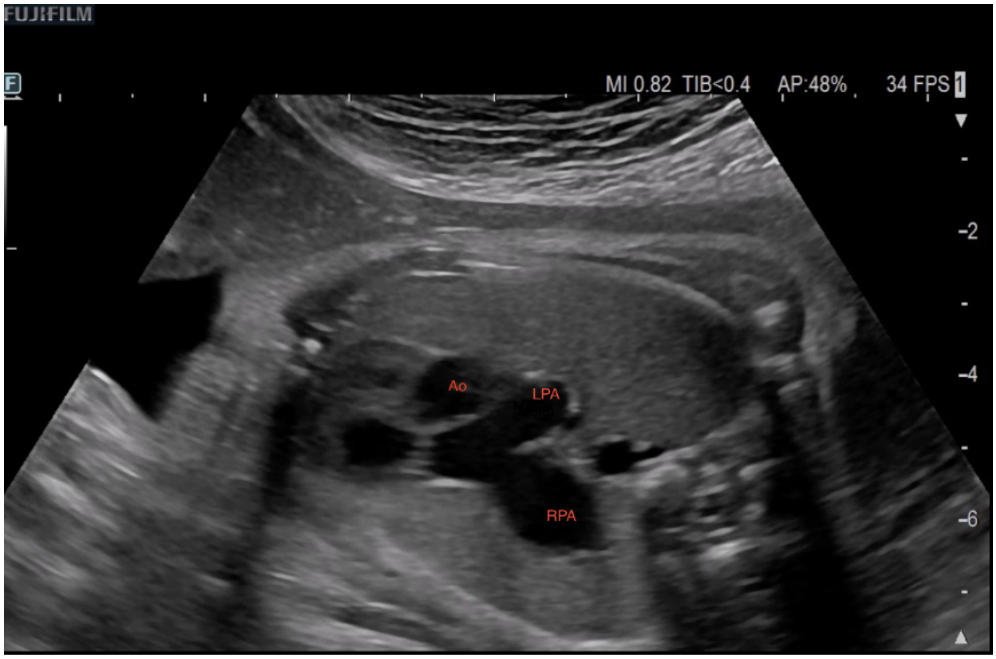

Fig. 1. Dilated right and left pulmonary arteries (RPA-LPA) and the aorta (Ao) in a fetus with pulmonary valve absence (Ao=Aorta, RPA=Right Pulmonary Artery, LPA=Left Pulmonary Artery).

of these cases was terminated in the second trimester because of associated extracardiac anomalies (acrania and omphalocele). Therefore, the clinical findings and outcome analyses were based on 7 live births.

The mean age of the pregnant women was 28.25±6.64 years, and the age range was 21-39 years. The mean body mass index (BMI) was 30.11±4.75 kg/m², and the values varied between 22.7 and 36.3. The mean gravida was 2.87±1.88, the mean parity was 1.12±0.99 (range: 0-2), and the mean abortion rate was 0.75±1.38 (range: 0-4).

With respect to the mode of delivery, 6 (85.7%) patients had cesarean section, and 1 (14.3%) patient had normal delivery. When the fetal sex distribution was analyzed, 5 (71.4%) patients were male, and 2 (28.6%) patients were female. Fetal birth weights ranged between 1200 and 3150 grams, with a mean of 2178.33±780.13 grams. When the gestational weeks of the prenatally diagnosed cases were analyzed, the earliest gestational week at the time of diagnosis was recorded as 18+2 (128 days), and the latest was recorded as 28+5 (201 days). The mean gestational age at presentation was 24+5 weeks (173 days) ±3.5 weeks.

When the postnatal Apgar scores were examined, the 1-minute Apgar values ranged between a minimum of 1 and a maximum of 8, with a mean of 4.00±2.37. The 5-minute Apgar scores ranged from a minimum of 5 to a maximum of 8, with a mean of 5.71±1.38. Postnatal oxygen saturation values ranged from a minimum of 70% to a maximum of 88%, with a median of 83%. In two (28.6%) patients, postnatal respiratory distress and oxygen desaturation necessitated endotracheal intubation. Amniotic fluid assessment revealed polyhydramnios in four (57.1%) patients, normal fluid volume in two (28.6%) patients, and oligohydramnios in one (14.3%) patient.

When the gestational weeks were analyzed, the earliest gestational week was 30+3 (day 213), and the latest gestational week was 38+5 (day 271), after one patient who was excluded from the study because of termination. The mean number of gestational weeks was 31+2 (219th day), and the median number of gestational weeks was 32+1 (225th day). The standard deviation was ±6.45 weeks (~45.2 days).

In one patient with prenatal acrania and omphalocele, the pregnancy was terminated in the second trimester, and this patient was excluded from the study. In the evaluation of the remaining seven live fetuses, extracardiac structural anomalies were detected in two (28.6%) cases. One of these patients had esophageal atresia and skeletal dysplasia, and the other had bilateral pyelectasis and agenesis of the corpus callosum. The other five (71.4%) fetuses had no extracardiac structural abnormalities.

In terms of cardiac anatomy, 6 (85.7%) patients lacked a pulmonary valve with TOF, and 1 patient (14.3%) lacked a pulmonary valve with a DORV. A right aortic arch anomaly was found in two patients (33.3%) with TOF and left pulmonary artery hypoplasia, and direct origin of the left pulmonary artery from the aortic arch (AOLPA) was observed in one patient (16.7%).

Pulmonary artery diameters were analyzed as follows:

(1) MPA diameter ranged from 6.5-10.3 mm, with a mean of 8.53±1.38 mm.

(2) LPA diameter ranged from 1.5-10.6 mm, with a mean of 7.65±3.24 mm. AOLPA was present in one patient with an LPA diameter of 1.5 mm.

(3) RPA diameter ranged from 6.8-12.7 mm, with a mean of 8.6±2.08 mm.

Fig. 1 shows the markedly dilated right and left pulmonary arteries along the aorta in a thoracic axial three-vessel view of a fetus diagnosed with absent pulmonary valve syndrome. This image supports the typical prenatal echocardiographic findings associated with the condition.

In fetuses with pulmonary regurgitation, PSV values ranged between 127 and 210 cm/sec. The mean PSV was 178.8±36.1 cm/sec, and the median value was 200.0 cm/sec. In patients with pulmonary stenosis, the PSV ranged between 120.0 and 260.0 cm/sec, with a mean of 184.0±50.8 cm/sec and a median of 187.0 cm/sec.

Postnatal cardiac evaluation revealed TOF in 6 of the 7 patients (85.7%), and 1 patient (14.3%) had a DORV associated with this condition. Pulmonary regurgitation was observed in all patients (100%); 3 (42.9%) had free regurgitation, and 4 (57.1%) had moderate to significant regurgitation. Pulmonary stenosis was present in 6 patients (85.7%), of whom 5 (71.4%) had valvular stenosis of moderate to severe severity.

Among the postnatal echocardiography controls, patency at the atrial level was found in 5 cases (71.4%); 3 (42.9%) of these cases were defined as PFO, and 2 (28.6%) were defined as wide secondum

type ASDs. Patent ductus arteriosus (PDA) was observed in 1 patient (14.3%). The right aortic arch was confirmed postnatally in 2 patients (28.6%).

When the clinical outcomes of 7 live-born APVS patients were evaluated at the postnatal follow-up, 4 patients (57.1%) underwent surgery. Three of these patients (42.9%) survived postoperatively- one followed up to 19 months, one to 6 months, and one in the early postoperative period. One (14.3%) patient died in the second postoperative month during intensive care unit follow-up. Two (28.6%) patients with severe cardiac failure and hydrops fetalis died in the early neonatal period before surgery. Another patient (14.3%) died on postnatal day 40 before reaching surgery.

Fig. 2 provides a visual summary of the perinatal course and postnatal outcomes of the seven live-born cases, including surgical interventions, survival status, and causes of mortality. This diagram enhances the understanding of the clinical prognosis of APVS patients.

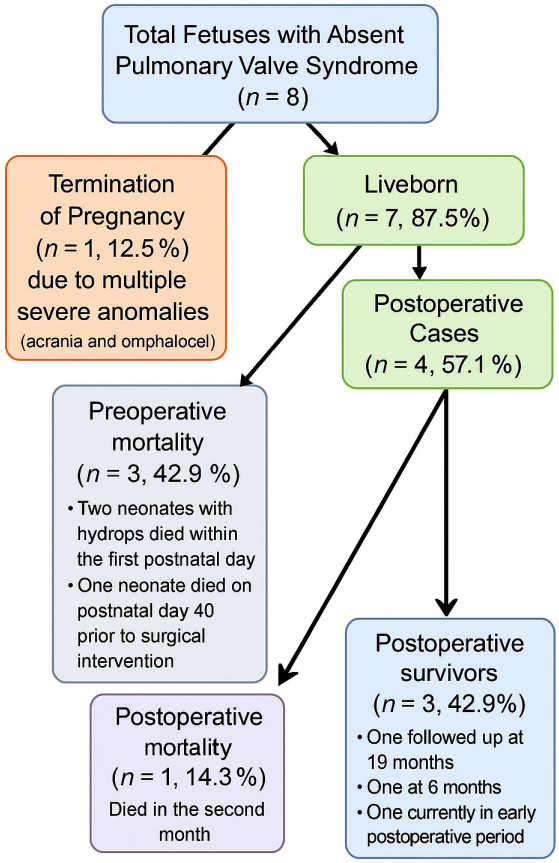

Fig. 2. Clinical outcomes of fetuses with APVS: From rermination to postoperative survival (APVS=Absent Pulmonary Valve Syndrome).

In the series of 7 cases, genetic analysis was performed prenatally in three cases and postnatally in two cases. One of the three patients tested prenatally was found to have ABAT gene mutations, and the other two were reported to be normal. Among the two postnatal cases, one was compatible with trisomy 21 (Down syndrome), and the other had a 22q11.2 deletion (DiGeorge syndrome). Accordingly, genetic anomalies were found in a total of three cases (42.8%) in the whole series: (1) One case (14.3%) of ABAT gene mutation (by prenatal WES); (2) 1 patient (14.3%) with trisomy 21; and (3) One patient (14.3%) had a 22q11.2 deletion.

In two of the remaining four cases, prenatal tests were reported as normal, and genetic testing was not performed in two cases.

DISCUSSION

Absent pulmonary valve syndrome (APVS) is a rare congenital heart defect characterized by high perinatal mortality and complex structural anomalies that pose significant challenges for surgical repair, necessitating accurate prenatal diagnosis, detailed genetic evaluation, and coordinated postnatal multidisciplinary care [6]. Despite its clinical importance, large-scale case series on APVS remain scarce in the literature, with most information derived from isolated case reports and limited retrospective analyses [7]. In this context, our study provides a comprehensive evaluation of seven live-born fetuses among eight prenatally diagnosed APVS cases, focusing on the timing of diagnosis, associated cardiac and extracardiac anomalies, Doppler findings, genetic results, and postnatal surgical outcomes.

In our series, 87.5% of fetuses had tetralogy of Fallot, 28.6% had extracardiac anomalies, 57.1% experienced perinatal mortality, and 42.9% presented with genetic abnormalities. The presence of the ductus arteriosus and hydrops fetalis are key predictors of postnatal survival. These findings highlight the prognostic value of early diagnosis and stratification, and our comparative analysis aims to deepen the understanding of this rare and complex condition.

One of the key prenatal indicators of APVS is the "to-and-fro" flow pattern observed via color Doppler ultrasonography, which is characterized by antegrade

systolic flow into the pulmonary arteries and diastolic backflow into the right ventricle due to the absence or rudimentary pulmonary valve leaflets. This bidirectional mosaic flow at pulmonary artery outflow supports the diagnosis and aids in distinguishing APVS from other congenital heart defects, such as TOF and pulmonary atresia [8, 9]. In our series, this pattern was observed in all patients and served as the principal diagnostic marker. It was also associated with severe valvular regurgitation and pulmonary artery dilatation. Notably, these Doppler findings may predict postnatal complications, including hemodynamic overload and tracheobronchial compression. Thus, careful evaluation during fetal echocardiography is essential for accurate differential diagnosis of outflow tract anomalies.

Prenatal diagnosis of APVS typically occurs between 18-24 weeks of gestation, and the mean gestational age at diagnosis in our series was 24+5 weeks, which is consistent with the 23-week average reported in the literature [10]. First-trimester diagnoses are rare and generally associated with poor prognosis [11]. The absence of such early cases in our series may reflect both the rarity of early-onset forms and the imaging limitations during the first trimester.

Hydrops fetalis was identified antenatally in two fetuses, both of which were born alive but died in the early postnatal period, resulting in a hydrops-related perinatal mortality rate of 100%, which is in line with previous reports [8].

In one of these cases, postnatal evaluation revealed a patent ductus arteriosus (PDA), suggesting its contribution to perinatal decompensation via increased right ventricular volume load and worsened pulmonary regurgitation. These findings support the notion that PDA may exacerbate the hemodynamic burden in select cases. The interplay between PDA-related volume overload, pulmonary regurgitation, and right ventricular dilatation may lead to pulmonary vascular congestion and hydrops, increasing the risk of mortality. Therefore, the presence of hydrops and their hemodynamic contributors, including PDA, should be carefully assessed in the prenatal evaluation of APVS.

Careful assessment of the ductus arteriosus via fetal echocardiography is essential, as its presence has been associated with increased perinatal mortality, warranting appropriate genetic and prognostic counseling. Conversely, some studies suggest that the absence of PDA may contribute to hemodynamic stability and a more favorable prognosis in APVS patients [10, 11]. Thus, PDA should be regarded not only as an anatomical variant but also as a dynamic factor influencing clinical outcomes.

Pulmonary artery dilatation has been linked to polyhydramnios, hydrops fetalis, and postnatal respiratory failure due to tracheobronchial compression [12, 13]. In our series, the APA (8.53±1.38 mm), LPA (7.65±3.24 mm), and RPA (8.6±2.08 mm) diameters exceeded the gestational age-specific reference values, indicating a risk for airway obstruction. These findings were corroborated by indirect signs such as polyhydramnios and mediastinal shift during the prenatal period.

Therefore, pulmonary artery diameters in APVSs should be evaluated not only for diagnostic purposes but also for their prognostic implications in predicting postnatal respiratory morbidity. Quantitative assessment during fetal echocardiography provides valuable guidance for prenatal counseling and postnatal management [14].

The extracardiac anomalies observed in our series involved multiple systems, including fetal growth restriction (FGR), esophageal atresia, skeletal dysplasia, acrania, omphalocele, bilateral pyelectasis, and agenesis of the corpus callosum. These findings suggest that APVS may often present as part of syndromic or monogenic conditions rather than as an isolated cardiac defect. The most commonly affected systems in the literature are the central nervous, gastrointestinal, and genitourinary systems [14, 15].

The incidence of extracardiac anomalies in our study was 28.6%, which aligns with the reported range of 21.6–33.8% in previous studies [16]. This rate may reflect both the extent of systematic prenatal screening and the thoroughness of individual case evaluations.

Among the seven cases in our series, genetic analysis was performed in five—three prenatally and two postnatally. Clinically significant genetic anomalies were identified in three patients (42.9%): an ABAT gene mutation was detected prenatally, whereas postnatal analyses revealed trisomy 21 (Down syndrome) and 22q11.2 microdeletion (DiGeorge syndrome). Trisomy 21 and 22q11.2 deletions are well-documented syndromic associations with APVS, whereas the ABAT mutation represents a novel finding, potentially implicating a new genetic etiology.

The genetic diagnosis rate in our cohort (42.9%)

exceeded the 39.3% reported in the meta-analysis by Recker *et al*. [1, 7], possibly due to our small sample size and the application of advanced, phenotype-guided analyses such as prenatal whole exome sequencing (WES). Notably, the ABAT mutation was identified via prenatal WES in a fetus presenting with FGR, esophageal atresia, and limb shortness, highlighting the diagnostic value of phenotype-driven testing. Therefore, in addition to conventional chromosomal anomalies, specific monogenic syndromes should be considered in all fetuses with APVS, as genetic abnormalities may impact both cardiac and extracardiac development as well as perinatal prognosis [15, 16]. Stratified, phenotype-oriented screening approaches may thus enhance prenatal counseling and postnatal management.

Over the past 10-15 years, advancements in cardiac surgery, neonatal intensive care, and multidisciplinary management have significantly improved APVS outcomes. Reports from high-volume centers demonstrate high survival among patients undergoing surgery [16-18]. For example, the Toronto SickKids group reported 5- and 10-year survival rates of 93% and 87%, respectively, in 62 surgical patients - most of whom required prolonged neonatal ventilation (mean: 36 days) [19]. Similarly, data from CHOP indicate notable improvements in postnatal survival compared with earlier decades [20].

In our cohort, one of the eight cases was excluded due to pregnancy termination for severe extracardiac anomalies. Among the remaining seven live-born individuals, four underwent surgery; three survived, and one died in the second postnatal month, yielding a surgical survival rate of 75% and a mortality rate of 25%. The overall perinatal mortality rate among live births was 57.1% (4/7). These outcomes approach the upper limit of the 50-64% survival range reported in the literature and may reflect the benefits of early diagnosis, proactive postnatal evaluation, and multidisciplinary care at our center.

The pulmonary artery PSV is a key Doppler parameter for evaluating pulmonary outflow in APVSs. In healthy fetuses, second- and third-trimester values typically range from 120-160 cm/sec; values >200 cm/sec may indicate severe regurgitation and increased volume load [15-17]. In our series, the mean PSV was 178.7 cm/sec (range: 140-260). The highest PSV (260 cm/sec) occurred in a patient who died postnatally due to respiratory failure and tracheobronchial compression. Survivors had a PSV <200 cm/sec and underwent timely surgery, suggesting that a higher PSV may predict adverse outcomes.

These findings highlight the prognostic value of Doppler-based hemodynamic assessment. Although fetal MRI was not performed, indirect echocardiographic signs - such as marked pulmonary artery dilatation, polyhydramnios, and narrowed tracheal diameter on the 3VT view - supported the presence of airway compromise. Fetal MRI has been reported as a complementary tool for assessing tracheal compression and predicting postnatal respiratory morbidity in APVSs [12]. However, current evidence does not support its routine use. Instead, MRI should be selectively considered in patients with severe pulmonary artery dilatation, mediastinal shift, or marked polyhydramnios.

Most APVS cases exhibit TOF morphology, with reported rates of 84-93% [9, 10, 15]. Our series aligns with this trend but also includes a rare variant involving the DORV, characterized by both great arteries originating from the right ventricle, a large perimembranous VSD, and complete pulmonary valve aplasia. This highlights that APVSs may present with a broader morphologic spectrum than previously recognized, underscoring the need for careful differential diagnosis. By reporting both typical and rare anatomical variants, our study contributes novel insights to the literature.

CONCLUSION

This study represents one of the few prenatally diagnosed case series of APVS, highlighting its morphological and genetic heterogeneity and its association with high perinatal mortality. Among the seven live-born patients, genetic anomalies - including Down syndrome, 22q11.2 deletion, and a novel ABAT gene mutation - were identified in 42.9% of the patients, emphasizing the syndromic nature of APVS. A surgical survival rate of 75% was achieved among operated patients, while all patients with hydrops fetalis died before surgery, underscoring the prognostic severity of this complication. These findings confirm the importance of detailed fetal echocardiography, genetic evaluation, and hemodynamic monitoring in prenatal

care. Individualized, multidisciplinary management plans based on early risk stratification may improve both postnatal outcomes and the quality of prenatal counseling. Larger, prospective studies are needed to refine the diagnostic pathways and therapeutic strategies for APVS.

Ethics Approval and Consent to Participate

This study was approved by the Başakşehir Çam and Sakura City Hospital Clinical Research Ethics Committee (Decision No: 2024-63; date: 06.05.2024). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Written informed consent was obtained from the parents of all patients to use the data for scientific purposes.

Data Availability

All data generated or analyzed during this study are included in this published article. Additional information and supporting data related to the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contribution

Study Conception: TA; Study Design: VAT; Supervision: TA, VA; Funding: N/A; Materials: TA, İÖ; Data Collection and/or Processing: TA; Statistical Analysis and/or Data Interpretation: TA, İÖ; Literature Review: TA; Manuscript Preparation: TA; and Critical Review: TA.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the author(s) and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Recker F, Weber EC, Strizek B, Geipel A, Berg C, Gembruch U. Management and outcome of prenatal absent pulmonary valve syndrome. Arch Gynecol Obstet. 2022;306(5):1449-1454. doi: 10.1007/s00404-022-06397-4.
2. Piacentini G, Mastromoro G, Romano V, Riccardi R, Orfeo L. Fetal echocardiographic features of absent pulmonary valve syndrome. Am J Obstet Gynecol. 2022;227(2):331-332. doi: 10.1016/j.ajog.2022.02.023.
3. Gottschalk I, Jehle C, Herberg U, et al. Prenatal diagnosis of absent pulmonary valve syndrome from first trimester onward: novel insights into pathophysiology, associated conditions and outcome. Ultrasound Obstet Gynecol. 2017;49(5):637-642. doi: 10.1002/uog.15977.
4. Torok K, Brettle E, Desai T, et al. Long-term outcomes in children with absent pulmonary valve syndrome: it is not just fixing the heart. Arch Dis Child. 2021;106(9):877-881. doi: 10.1136/archdischild-2020-320219.
5. Chelliah A, Moon-Grady AJ, Peyvandi S, et al. Contemporary Outcomes in Tetralogy of Fallot With Absent Pulmonary Valve After Fetal Diagnosis. J Am Heart Assoc. 2021;10(12):e019713. doi: 10.1161/JAHA.120.019713.
6. Szwast A, Tian Z, McCann M, et al. Anatomic variability and outcome in prenatally diagnosed absent pulmonary valve syndrome. Ann Thorac Surg. 2014;98(1):152-158. doi: 10.1016/j.athoracsur.2014.03.002.
7. Yang Z, Zhou L. Right aortic arch with mirror image branching accompanied by absent pulmonary valve syndrome and tricuspid stenosis: Prenatal echocardiographic diagnosis of an unusual congenital heart defect. Echocardiography. 2019;36(10):1952-1955. doi: 10.1111/echo.14466.
8. Toyokawa T, Inamura N, Kawazu Y, Kayatani F. Circular shunt in fetal absent pulmonary valve with tricuspid stenosis. Pediatr Int. 2023;65(1):e15480. doi: 10.1111/ped.15480.
9. Swaminathan S, Agarwal A, Infante JC, Rosenkranz E. Tetralogy of Fallot With Absent Pulmonary Valve and Nonconfluent Pulmonary Arteries: A Management Conundrum. World J Pediatr Congenit Heart Surg. 2020;11(4):NP168-NP171. doi: 10.1177/2150135118775661.
10. Sourour W, Powell SK. A Rare Case of Tricuspid Atresia Absent Pulmonary Valve Diagnosed on Fetal Echocardiography. CASE

(Phila). 2023;7(12):487-491. doi: 10.1016/j.case.2023.09.002.
11. Song Y, Zou YF, Ru YH, Qiu J, Yin H. Absent pulmonary valve syndrome with tetralogy of fallot and patent ductus arteriosus at 14 weeks of gestation and follow-up 2 weeks later: Case report and review of literature. Echocardiography. 2021;38(3):484-487. doi: 10.1111/echo.14936.
12. Song J, Xu Y, Jiang Y, et al. The Role of First-Trimester Ultrasound in Detecting Aortic and Pulmonary Valve Agenesis: A Rare Case of Trisomy 13. J Clin Ultrasound. 2025 May 12. doi: 10.1002/jcu.24039.
13. Rao S, Najm HK, Stewart RD, Ahmad M, Erenberg F, Yaman M. Tetralogy of Fallot with absent pulmonary valve-When the ductus is present: A case of isolated branch pulmonary artery and review of literature. Echocardiography. 2019;36(5):996-1000. doi: 10.1111/echo.14334.
14. Qasim A, Johnson CB, Aly MA, Aly AM. Prenatal Diagnosis and Successful Palliation of Absent Aortic Valve with Hypoplastic Left Heart Syndrome: A Case Report and Review of Literature. AJP Rep. 2019;9(2):e121-e126. doi: 10.1055/s-0038-1677480.
15. Murakami T, Lin L, Ishiodori T, Takeuchi S, Shiono J, Horigome H. Prenatal diagnosis of congenital absence of aortic valve associated with restrictive foramen ovale: Hemodynamic features and clinical outcome. J Clin Ultrasound. 2019;47(2):104-106. doi: 10.1002/jcu.22636.
16. Monacci F, Bondi T, Canessa C, Chiappa E. 'Absent' pulmonary valve with intact ventricular septum mimicking tricuspid valve atresia: Prenatal diagnosis and postnatal course. J Obstet Gynecol Res. 2019;45(3):714-718. doi: 10.1111/jog.13878.
17. Moleiro ML, Guedes-Martins L. Prenatal diagnosis of absent pulmonary valve syndrome. BMJ Case Rep. 2021;14(1):e240567. doi: 10.1136/bcr-2020-240567.
18. Inamura N, Takada N, Marutani S. The prenatal diagnosis of a rare circular shunt with absent pulmonary valve syndrome. J Clin Ultrasound. 2022;50(1):86-89. doi: 10.1002/jcu.23031.
19. Wertaschnigg D, Jaeggi M, Chitayat D, et al. Prenatal diagnosis and outcome of absent pulmonary valve syndrome: contemporary single-center experience and review of the literature. Ultrasound Obstet Gynecol. 2013;41(2):162-167. doi: 10.1002/uog.11193.
20. Nair AK, Haranal M, Elkhatim IM, Dillon J, Hew CC, Sivalingam S. Surgical outcomes of absent pulmonary valve syndrome: An institutional experience. Ann Pediatr Cardiol. 2020;13(3):212-219. doi: 10.4103/apc.APC_111_19.

The European Research Journal 2025;11(6):1117-1122 Original Article

DOI: https://doi.org/10.18621/eurj.1669429 *Dermatology*

Left atrial volume index in patients with seborrheic dermatitis

Mehmet Semih Çelik[1], Abdulkadir Arpa[2]

[1]Department of Dermatology, University of Health Sciences, Diyarbakır Gazi Yaşargil Training and Research Hospital, Diyarbakır, Türkiye; [2]Department of Cardiology, Bismil State Hospital, Diyarbakır, Türkiye

ABSTRACT

Objectives: Seborrheic dermatitis is a chronic inflammatory skin condition that primarily affects the scalp, face, and body folds. Emerging research indicates a potential link between seborrheic dermatitis and cardiac parameters. This study aims to further investigate this relationship.

Methods: Fifty individuals diagnosed with seborrheic dermatitis who presented to the dermatology outpatient clinic were included, alongside a control group of healthy participants (n=30) matched by age and sex. Echocardiographic assessments were carried out for both groups, and the data were recorded using standardized forms.

Results: The left atrial volume index (LAVI) was recorded as 28.3±1.24 in seborrheic dermatitis patients and 26.2±1.47 in the control group, with this difference reaching statistical significance (P<0.001). Similarly, the E/e' ratio was found to be 8.52±0.777 in the patient group and 7.16±0.706 in the control group, demonstrating a statistically significant difference (P<0.001).

Conclusions: The observed increase in LAVI and E/e' ratio among seborrheic dermatitis patients suggests a potential association with an elevated risk of cardiovascular disease and atrial fibrillation. Given these findings, long-term monitoring of individuals with seborrheic dermatitis may be warranted.

Keywords: Left atrial volume index, E/e' ratio, seborrheic dermatitis

Seborrheic dermatitis is a frequently encountered dermatological condition that affects individuals across different age groups, including infants, adolescents, and adults. It is characterized by symptoms such as erythema, scaling, and pruritus, predominantly occurring on the scalp, face, chest, back, axillae, and groin. Although the exact pathogenesis of seborrheic dermatitis remains unclear, it is recognized as a common inflammatory skin disorder [1]. Recent research has explored potential associations between seborrheic dermatitis and various cardiac parameters [2].

One of the components of metabolic syndrome is inflammation, and many inflammatory cytokines play a critical role in the disease. In a study conducted on patients with seborrheic dermatitis, metabolic syndrome was found to be significantly more prevalent, and seborrheic dermatitis was suggested to be a potential predictive factor for metabolic syndrome. Similarly, another study observed that hypertension was significantly more common in patients with seborrheic dermatitis [3,4]. Among these, the left atrial volume index (LAVI) and the E/e' ratio are widely utilized indicators for assessing diastolic dysfunction [5].

In this study, we aim to expand on existing research

Received: April 3, 2025 **Accepted:** July 17, 2025 **Available Online:** July 29, 2025 **Published:** November 4, 2025

How to cite this article: Çelik MS, Arpa A. Left atrial volume index in patients with seborrheic dermatitis. Eur Res J. 2025;11(6):1117-1122. doi: 10.18621/eurj.1669429

Corresponding author: Mehmet Semih Çelik, MD., Phone: +90 412 258 00 60, E-mail: drmsemihcelik@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

by evaluating LAVI and E/e' ratio measurements in patients diagnosed with seborrheic dermatitis.

METHODS

The ethics committee of the University of Health and Sciences, Gazi Yaşargil Training and Research Hospital, provided ethical approval for this research. The study's execution was aligned with the principles set forth in the Helsinki Declaration. A total of 50 patients diagnosed with seborrheic dermatitis, who sought care at the dermatology outpatient clinic, were enrolled. Individuals with atrial fibrillation, structural heart valve disease, coronary artery disease, or atherosclerotic risk factors such as hypertension, diabetes, and smoking were excluded from the study. The severity of seborrheic dermatitis was assessed via the Seborrheic Dermatitis Area and Severity Index (SEDASI). SEDASI was used to evaluate disease severity. This scoring system assesses four anatomical regions and rates each region on four criteria: extent of involvement (0-6), lesion pattern (0-3), erythema intensity (0-3), and scaling severity (0-3). The total score is the sum of all regional subscores. Severity categories are defined as 1-14 (mild), 15-29 (moderate), 30-44 (severe), and >45 (very severe). Patients scoring below 29 were categorized as having mild to moderate disease, while those with scores exceeding 29 were classified as severe cases. Key echocardiographic parameters, including LAVI, E/e' ratio, left ventricular end-diastolic pressure, and left ventricular ejection fraction, were recorded using standardized forms. Standard echocardiographic imaging was performed for all participants. The ejection fraction was determined via the modified biplane Simpson's method. LAVI was measured using the biplane disk summation technique at the end of systole and indexed to body surface area, yielding values in mL/m². Mitral inflow dynamics were analyzed using pulsed-wave Doppler in apical four-chamber and two-chamber views, with early (E) filling velocity recorded. Additionally, early (e') mitral annular velocity was assessed using tissue Doppler imaging in the apical four-chamber position, allowing for the calculation of the E/e' ratio. For comparison, 30 healthy volunteers were included in the study, and their echocardiographic findings were evaluated alongside those of the patient group.

Statistical Analysis

The data underwent statistical processing with SPSS version 27 and JAMOVI software. Frequency and percentage distributions summarized categorical data, whereas continuous data were characterized by their mean, standard deviation, minimum, maximum, and median. Subgroup comparisons for categorical variables were conducted using the chi-squared test. The Student's t-test served to analyze differences in continuous variables between independent groups, depending on their normality. A P-value below 0.05 was the criterion for statistical significance in all analyses.

RESULTS

The study included 50 patients diagnosed with seborrheic dermatitis and 30 healthy controls. Among all participants, 42 (52%) were male, and 38 (48%) were female. Disease severity in the patient group assessment of the condition was assessed via the SEDASI. Based on their SEDASI scores, patients scoring below 29 were categorized as having mild to moderate disease, while those with scores exceeding 29 were classified as severe. According to this classification, 30 (60%) patients were identified as having severe seborrheic dermatitis, whereas 20 (40%) patients had mild to moderate disease. The average age of the patients included in the study was 27.8±6.55 years; for those in the control group, the mean age was 29.0±5.03 years. The two groups did not show a statistically significant difference (P=0.384).

LAVI was measured at 28.3±1.24 in patients and 26.2±1.47 in controls, with this difference reaching statistical significance (P<0.001). Similarly, compared to the control, the patient group exhibited an elevated E/e' ratio (8.52±0.777) compared to controls (7.16±0.706) (P<0.001). In contrast, left ventricular end-diastolic pressure was similar between patients (42.8±1.75) and controls (42.8±1.65), showing no significant difference (P=0.96). Likewise, LVEF was comparable between groups, with values of 63.3±2.39 in patients and 63.0±2.49 in controls (P=0.595) (Table 1).

LAVI levels were significantly elevated in the severe disease group relative to the mild-to-moderate disease group (28.9±1.08) than in the mild-to-moderate group (27.3±0.665) (P<0.001). Additionally, the E/e' ratio was elevated in the severe group (8.77±0.76)

Table 1. Comparison of demographic and echocardiographic data between the case and control groups

Variable	Seborrheic dermatitis patients (n=50)	Controls (n=30)	P value
Age (years)	27.8±6.55	29.0± .03	0.384
Sex (Male/female)	26/24	16/14	-
Disease severity (			
Mean SDASI score	31.0±5.5	N/A	-
SEDASI severity classification	30 (60%) severe 20 (40%) mild-moderate	N/A	-
Echocardiographic findings			
LAVI (mL/m²)	28.3±1.24	26.2±1.47	**<0.001**
E/e' ratio	8.52±0.77	7.16±0.70	**<0.001**
LVEDP (mmHg)	42.8±1.75	42.8±1.65	0.96
LVEF (%)	63.3±2.39	63.0±2.49	0.595

Data are shown as mean±standard deviation or n (%). SEDASI=Seborrheic Dermatitis Area and Severity Index, LAVI=Left Atrial Volume Index, LVEDP=Left Ventricular End-Diastolic Pressure, LVEF=Left Ventricular Ejection Fraction

compared to the mild-to-moderate group (8.15±0.657), with this difference also reaching statistical significance (P=0.004) (Table 2).

DISCUSSION

In our study, LAVI values and Patients diagnosed with seborrheic dermatitis exhibited significantly higher E/e' ratios when compared to the control group. Furthermore within the patient group, those with severe disease had significantly higher LAVI and E/e' values than those with mild-to-moderate disease. Given that elevated LAVI and E/e' ratios are associated with cardiovascular diseases and rhythm disorders such as atrial fibrillation, we recommend long-term follow-up for these patients to monitor potential cardiac involvement.

Seborrheic dermatitis is a common inflammatory skin disorder affecting infants, adolescents, and adults. It is characterized by symptoms such as scaling, erythema, and itching, and it predominantly affects areas with a high density of sebaceous glands, including the scalp, face, chest, back, armpits, and groin [1]. While the exact etiology remains unclear, seborrheic dermatitis is recognized as a chronic inflammatory condition. Several studies have explored its relationship with systemic inflammation and cardiovascular parameters.

For instance, a study investigating hematologic markers in 100 individuals found significantly ele-

Table 2. Comparison of LAVI and E/e' ratio in patients with severe and mild-moderate seborrheic dermatitis.

Variable	Severe SD (SEDASI Score ≥29) (n=30)	Mild-moderate SD (SEDASI Score <29) (n=20)	P value
LAVI (mL/m²)	28.9±1.08	27.3±0.665	**<0.001**
E/e' ratio	8.77±0.76	8.15±0.657	**0.004**

Data are shown as mean±standard deviation or n (%). LAVI=Left Atrial Volume Index, SEDASI=Seborrheic Dermatitis Area and Severity Index, SD=Seborrheic Dermatitis

vated C-reactive protein and platelet-to-lymphocyte ratio levels in seborrheic dermatitis patients compared to controls. Additionally, another study involving 80 participants demonstrated that epicardial fat thickness, a marker associated with cardiovascular risk A notable increase was observed in patients with seborrheic dermatitis [2, 5].

Assessing left atrial (LA) size using the LAVI is clinically important due to its strong prognostic value in various cardiovascular conditions. Moreover, incorporating LA function assessment alongside LAVI provides additional clinical and prognostic insights in different cardiovascular diseases [6].

LAVI is commonly accepted as a trustworthy indicator of diastolic dysfunction. Echocardiographic studies suggest that a normal LAVI is approximately 22±6 mL/m², while values exceeding 28 mL/m² are considered significant due to their sensitivity and specificity in predicting adverse cardiac events [7].

A retrospective study involving 1,160 patients identified elevated LAVI. This is commonly acknowledged as a standalone risk factor for heart-related illnesses, atrial fibrillation, and myocardial infarction [8]. Similarly, a systematic review of 11 cohort studies comprising 2,705 patients found that A significant decrease in major adverse cardiac events was observed in patients exhibiting lower LAVI levels, Furthermore, hospitalization rates were lower among individuals with low LAVI, A higher LAVI level was found to be an independent predictor of unfavorable outcomes in patients with acute coronary syndrom [9].

Additionally, research involving patients with psoriasis has demonstrated alterations in left atrial mechanics, which are believed to elevate the likelihood of developing atrial fibrillation [10-12]. Similar findings have been reported in individuals with hidradenitis suppurativa - another chronic inflammatory skin condition - where an increased incidence of cardiovascular disease has been observed [13-15].

In our study, the elevated LAVI values among patients with seborrheic dermatitis suggest a potential predisposition to cardiovascular complications, particularly atrial fibrillation, warranting long-term monitoring. However, these findings should be interpreted with caution, as subclinical cardiovascular risk factors such as early hypertension, mild obesity, or undiagnosed metabolic abnormalities, which were not systematically assessed in this study, may also contribute to increased LAVI values. Therefore, while routine echocardiographic screening cannot be universally recommended for all patients with seborrheic dermatitis, clinicians should maintain vigilance for possible subclinical cardiac alterations, particularly in those with additional cardiovascular risk factors Although the risk of major cardiovascular events is generally more pronounced when LAVI exceeds 32 mL/m², the threshold for increased cardiovascular risk is considered to begin at 28 mL/m². In our cohort, patient LAVI values were above this critical threshold, reinforcing the importance of close cardiac surveillance.

According to current Doppler echocardiography guidelines, we examined the early-to-late transmitral flow velocity ratio (E/A) and the early diastolic mitral annular velocity (E/e') are key parameters in assessing left ventricular diastolic function. Subclinical diastolic dysfunction may begin even in asymptomatic individuals when LAVI exceeds 28 and E/e' surpasses 8 [16].

Consistent with our findings, a study involving 40 patients with psoriasis reported an increased E/e' ratio and evidence of diastolic dysfunction [17]. Moreover, in patients treated with biological agents such as ustekinumab and ixekizumab, a reduction in E/e' was observed after one year of therapy, highlighting the potential reversibility of cardiac changes with effective anti-inflammatory treatment [18]. Likewise, another study focusing on myocardial infarction patients revealed that elevated E/e' ratios could predict in-hospital complications and were associated with adverse prognosis [19].

Limitations

This study has several limitations. Firstly, the relatively small sample size limits the generalizability of the findings. In addition, the single-center and cross-sectional design of the study makes it difficult to establish a causal relationship. Another limitation is that cardiac parameters were assessed solely by echocardiographic evaluation; the lack of advanced imaging modalities may have led to missing some subclinical findings. Finally, the absence of a detailed analysis of patients' comorbid systemic diseases might have overlooked their potential influence on cardiac parameters. Future studies with larger sample sizes, multicenter participation, and prospective design are needed to overcome these limitations and provide more robust evidence.

CONCLUSION

Our investigation revealed that both the left atrial volume index (LAVI) and the E/e' ratio were significantly higher in the patient group than in the control group. Although these values did not reach clearly pathological thresholds, the concurrent increase in both parameters underscores the potential cardiovascular risk in this population. The results emphasize the importance of extended surveillance for possible cardiovascular diseases and rhythm disturbances, such as atrial fibrillation, in patients with seborrheic dermatitis.

To the best of our knowledge, this study represents the first specific evaluation of LAVI in individuals with seborrheic dermatitis, adding novel insights to the existing literature. We believe our findings provide a foundation for future research, and we recommend that subsequent studies include larger cohorts and extended follow-up periods are necessary to fully understand the clinical importance of these cardiac changes. However, routine echocardiographic screening cannot yet be recommended for all patients with seborrheic dermatitis; instead, clinicians should consider closer cardiac evaluation in those with additional risk factors.

Ethics Approval and Consent to Participate

This study was approved by the Diyarbakır Gazi Yaşargil and Research Hospital Clinical Research Ethics Committee. (Date: 22.11.2024, Number: 267). The study was conducted in accordance with the principles of the Helsinki Declaration. Furthermore, it was conducted in accordance with the guidelines of good clinical practice. Informed consent was obtained from all individual participants included in the study.

Data Availability

The complete dataset supporting the results of this study can be found within the article itself. For any further inquiries, please reach out to the corresponding author

Authors' Contribution

Study Conception: MSÇ; Study Design: MSÇ; Supervision: MSÇ; Funding: MSÇ, AA; Materials: MSÇ, AA; Data Collection and/or Processing: MSÇ, AA; Statistical Analysis and/or Data Interpretation: AA; Literature Review: MSÇ; Manuscript Preparation: MSÇ; and Critical Review: MSÇ.

Conflict of interest

The author disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the author(s) and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Clark GW, Pope SM, Jaboori KA. Diagnosis and treatment of seborrheic dermatitis. Am Fam Physician. 2015;91(3):185-190.
2. Celik MS, Arpa A. Evaluation of epicardial fat tissue thickness in patients with seborrheic dermatitis. Arch Dermatol Res. 2025;317(1):364. doi: 10.1007/s00403-025-03909-5.
3. Imamoglu B, Hayta SB, Guner R, Akyol M, Ozcelik S. Metabolic syndrome may be an important comorbidity in patients with seborrheic dermatitis. Arch Med Sci Atheroscler Dis. 2016;1(1):e158-161. doi: 10.5114/amsad.2016.65075.
4. Linder D, Dreiher J, Zampetti A, Sampogna F, Cohen AD. Seborrheic dermatitis and hypertension in adults: a cross-sectional study. J Eur Acad Dermatol Venereol. 2014;28(11):1450-1455. doi: 10.1111/jdv.12310.
5. Tosun M, Yasak Güner R, Akyol M. Investigation of the relationship between inflammatory blood parameters and seborrheic dermatitis. J Cosmet Dermatol. 2022;21(10):5111-5115. doi: 10.1111/jocd.14984.
6. Mangia M, D'Andrea E, Cecchetto A, Beccari R, Mele D, Nistri S. Current and Clinically Relevant Echocardiographic Parameters to Analyze Left Atrial Function. J Cardiovasc Dev Dis. 2024;11(8):241. doi: 10.3390/jcdd11080241.

7. Hoit BD. Left atrial size and function: role in prognosis. J Am Coll Cardiol. 2014;63(6):493-505. doi: 10.1016/j.jacc.2013.10.055.
8. Tsang TS, Barnes ME, Gersh BJ, et al. Prediction of risk for first age-related cardiovascular events in an elderly population: the incremental value of echocardiography. J Am Coll Cardiol. 2003;42(7):1199-1205. doi: 10.1016/s0735-1097(03)00943-4.
9. Ahmeti A, Bytyçi FS, Bielecka-Dabrowa A, Bytyçi I, Henein MY. Prognostic value of left atrial volume index in acute coronary syndrome: A systematic review and meta-analysis. Clin Physiol Funct Imaging. 2021;41(2):128-135. doi: 10.1111/cpf.12689.
10. Atas H, Kepez A, Bozbay M, et al. Assessment of left atrial volume and function in patients with psoriasis by using real time three-dimensional echocardiography. Wien Klin Wochenschr. 2015;127(21-22):858-863. doi: 10.1007/s00508-015-0817-4.
11. Aksan G, Nar G, Soylu K, et al. Assessment of atrial electromechanical delay and left atrial mechanical functions in patients with psoriasis vulgaris. Echocardiography. 2015;32(4):615-622. doi: 10.1111/echo.12706.
12. Simsek H, Sahin M, Akyol A, et al. Increased risk of atrial and ventricular arrhythmia in long-lasting psoriasis patients. ScientificWorldJournal. 2013;2013:901215. doi: 10.1155/2013/901215.
13. Krajewski PK, Matusiak Ł, Ständer S, Thaçi D, Szepietowski JC, Zirpel H. Risk of cardiovascular disorders in hidradenitis suppurativa patients: a large-scale, propensity-matched global retrospective cohort study. Int J Dermatol. 2024;63(6):799-805. doi: 10.1111/ijd.17186.
13. Sánchez-Díaz M, Salvador-Rodríguez L, Cuenca-Barrales C, Arias-Santiago S, Molina-Leyva A. Exploring the role of systemic immune-inflammation index and neutrophil-lymphocyte ratio in cardiovascular risk stratification for patients with hidradenitis suppurativa: A cross-sectional study. J Dermatol. 2022;49(12):1238-1244. doi: 10.1111/1346-8138.16558.
15. Juhl CR, Miller IM, Jemec GB, Kanters JK, Ellervik C. Hidradenitis suppurativa and electrocardiographic changes: a cross-sectional population study. Br J Dermatol. 2018;178(1):222-228. doi: 10.1111/bjd.15778.
16. Mitter SS, Shah SJ, Thomas JD. A Test in Context: E/A and E/e' to Assess Diastolic Dysfunction and LV Filling Pressure. J Am Coll Cardiol. 2017;69(11):1451-1464. doi: 10.1016/j.jacc.2016.12.037.
17. Bülbül Şen B, Ekiz Ö, Rifaioğlu EN, et al. Assessment of subclinical left ventricular dysfunction in patients with psoriasis by speckle tracking echocardiography: A Speckle Tracking Study. Int J Dermatol. 2016;55(2):158-164. doi: 10.1111/ijd.12703.
18. Baykal Selçuk L, Şahin M, Aydemir B, Aksu Arıca D, Yaylı S. Effects of interleukin-12/23 inhibitors and interleukin-17 inhibitors on myocardial functions in patients with severe psoriasis. Turkderm-Turk Arch Dermatol Venereol. 2024;58(4):112-116. doi: 10.4274/turkderm.galenos.2024.20737.
19. Mikeladze B, Zhvania N, Nikolaishvili G. E/e' Ratio as a Predictor of In-Hospital Complications and Clinical Outcomes of Acute Myocardial Infarction. Cureus. 2024;16(8):e66795. doi: 10.7759/cureus.66795.

The European Research Journal 2025;11(6):1123-1131 Original Article

DOI: https://doi.org/10.18621/eurj.1679701 *Emergency Medicine*

Effect of negative pressure wound therapy applied in the emergency center on postoperative infection rates in patients with orthopedic open wounds

Semih Eriten[1], Vedat İzci[1], Serdar Menekşe[2]

[1]Department of Emergency Medicine, Sultanbeyli State Hospital, İstanbul, Türkiye; [2]Department of Orthopedics and Traumatology, Adana Seyhan State Hospital, Adana, Türkiye

ABSTRACT

Objectives: This study aimed to evaluate the effect of Negative Pressure Wound Therapy (NPWT) applied in the emergency department on postoperative infection rates in patients with orthopedic open injuries.

Methods: This prospective randomized controlled study included 80 patients who presented with orthopedic open injuries. Patients were divided into two groups: NPWT group (n=40) and control group receiving standard wound care (n=40). An average pressure of 125±25 mmHg was applied in the NPWT group. Demographic characteristics, wound characteristics, infection rates, length of hospital stay, and treatment outcomes were evaluated.

Results: The infection rate in the NPWT group (7.5%) was significantly lower compared to the control group (25.0%) (P=0.034). The NPWT group showed significantly lower numbers of surgical debridement (1.3±0.6 vs 2.1±0.8, P=0.008), shorter antibiotic use duration (10±3 vs 14±4 days, P=0.003), and shorter hospital stay (8±2 vs 14±3 days, P<0.001). Primary closure rate was significantly higher in the NPWT group (70%) compared to the control group (30%) (P=0.002).

Conclusions: NPWT applied in the emergency department significantly reduces postoperative infection rates, accelerates wound healing, and shortens hospital stay in orthopedic open injuries. These findings suggest that NPWT may have an important role in the early treatment protocol of open injuries.

Keywords: Negative pressure wound therapy, open fracture, infection, emergency department, orthopedic trauma, wound healing

Orthopedic open injuries constitute one of the most challenging areas of modern trauma surgery. In the management of these injuries, infection prevention strategies stand out as one of the most critical factors determining treatment success [1]. Especially in recent years, significant paradigm shifts have occurred in the treatment approaches of open injuries and new treatment modalities have been developed for infection control.

Negative Pressure Wound Therapy (NPWT) is accepted as a system that accelerates wound healing and reduces the risk of infection by applying controlled negative pressure to the wound surface [2]. The basic mechanism of action of NPWT is based on reducing

Received: April 19, 2025 **Accepted:** August 17, 2025 **Available Online:** August 19, 2025 **Published:** November 4, 2025

How to cite this article: Eriten S, Menekşe S. Effect of negative pressure wound therapy applied in the emergency center on postoperative infection rates in patients with orthopedic open wounds. Eur Res J. 2025;11(6):1123-1131. doi: 10.18621/eurj.1679701

Corresponding author: Semih Eriten, MD., Phone: +90 216 564 24 00, E-mail: semiheriten@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

edema in the wound bed, increasing microvascular blood flow, and accelerating wound healing.

It is reported in the literature that infection rates in open wounds vary between 5% and 50% depending on the severity and localization of the injury [3]. It has been shown that the use of NPWT in the management of open extremity wounds provides lower infection rates and shorter hospital stays compared to traditional dressing techniques [4].

In current approaches, it is emphasized that the management of soft tissue damage in open wounds is of critical importance for the success of treatment and that appropriate wound care strategies should be implemented with aggressive debridement in the early period [5]. It is known that infections developing after orthopedic trauma significantly increase patient morbidity and require a multidisciplinary approach [6]. The development of NPWT systems, especially with the introduction of instillation systems, allows effective application of antimicrobial solutions to the wound bed and reduces bacterial load more effectively [7]. Clinical studies have shown that NPWT is more effective than traditional dressing methods in contaminated and infected wounds and that successful results are achieved even in fracture cases with nonunion problems [8].

This study aimed to evaluate the effect of NPWT applied in the emergency department on postoperative infection rates in patients with orthopedic open injuries. Our study aims to reveal the effect of early application of NPWT on clinical outcomes and to evaluate the role of this treatment modality in preventing infection.

METHODS

Study Design and Patient Classification

This prospective randomized controlled study was conducted on patients with orthopedic open injuries who were admitted to our hospital emergency department between January 2024 and December 2024. The current Gustilo-Anderson classification, which also takes into account changes during surgery, was used to classify injuries [9].

Randomization Methodology

Patients were randomized into two groups using a computer-generated randomization sequence with a 1:1 allocation ratio. Block randomization with variable block sizes (4, 6, and 8) was employed to ensure balanced group allocation throughout the study period. Allocation concealment was maintained using sequentially numbered, opaque, sealed envelopes prepared by an independent statistician not involved in patient recruitment or treatment. The randomization sequence was generated using statistical software (SPSS version 25.0) with a predetermined seed number to ensure reproducibility.

Sample Size Calculation

Sample size was calculated based on the primary outcome of infection rate. Based on previous literature showing infection rates of approximately 25% with standard care in similar patient populations [3], we hypothesized that NPWT would reduce this rate to 5%. Using a two-sided significance level of 0.05 ($\alpha = 0.05$) and power of 80% ($\beta = 0.20$), the calculated minimum sample size was 36 patients per group. To account for potential dropouts and protocol violations, the sample size was increased by 10%, resulting in 40 patients per group (total n=80). Sample size calculation was performed using G*Power software version 3.1.9.2.

Negative Pressure Wound Treatment Protocol

A standardized wound treatment protocol was applied in the NPWT group. This protocol was performed with a pressure of -125 mmHg in line with the results of the WHIST study [10]. The effectiveness of the treatment was evaluated using standardized quality of life scales [11].

Detailed NPWT Protocol Specifications

NPWT was applied using a standardized protocol with the following specifications:

(a) Device: Vacuum-assisted closure system (V.A.C. therapy unit or equivalent)

(b) Sponge type: Black polyurethane foam for all cases, cut to fit wound dimensions with 1-2 cm overlap onto healthy skin

(c) Pressure setting: Continuous negative pressure of -125 mmHg (±25 mmHg tolerance)

(d) Dressing change frequency: Every 48-72 hours or earlier if clinically indicated (excessive drainage, loss of seal, signs of infection)

(e) Duration: NPWT was maintained until wound

bed was suitable for definitive closure or for a maximum of 14 days

(f) Monitoring: Daily assessment of vacuum seal integrity, drainage volume, and wound appearance.

Risk Assessment and Scoring

The DANGER (Diabetes, Antibiotic, Nature of trauma, Grade of fracture, Exposure of the fracture, and Relative risk of patient including use of tobacco, alcoholism, and psychiatric disorders) scoring system, which evaluates the degree of injury, contamination level, soft tissue damage, vascular status, and fracture pattern, was used to determine the risk of infection in all patients [12].

Antibiotic Prophylaxis

A customized antibiotic prophylaxis protocol was applied according to the type of injury in line with current guidelines [13].

Antibiotic prophylaxis was administered according to Gustilo-Anderson classification:

(1) Type I fractures: Cefazolin 2g IV administered within 3 hours of injury, followed by 1g IV every 8 hours for 24 hours

(2) Type II fractures: Cefazolin 2g IV every 8 hours for 48 hours

(3) Type III fractures: Cefazolin 2g IV + Gentamicin 5mg/kg IV every 8 hours for 72 hours

(4) Contaminated wounds: Addition of Clindamycin 600mg IV every 8 hours for anaerobic coverage

(5) Penicillin allergy: Vancomycin 15mg/kg IV every 12 hours + Gentamicin 5mg/kg IV daily.

Surgical Timing and Technique

The effect of surgical intervention timing on infection rates in open fractures of the hand region was evaluated in light of literature data [14]. A standardized surgical protocol was used in open fractures of the tibial plateau [15].

Wound Care and Irrigation

A standard protocol recommended in the literature was applied for wound irrigation [16]. Within the scope of this protocol, irrigation solution selection, pressure settings, irrigation amount and application technique were standardized.

Given the significant difference in admission time between groups (P=0.028), potential confounding effects were addressed through: (a) Multivariate logistic regression analysis including admission time as a covariate; (b) Stratified analysis by admission time (<6 hours vs ≥6 hours); and (c) Sensitivity analysis excluding patients with extreme admission times (>12 hours).

Statistical Analysis

Data analysis was performed using SPSS version 25.0 software. Following normality tests for continuous variables, Student's t-test or Mann-Whitney U test was used, and Chi-square or Fisher's exact test was used for categorical variables. Risk factors were evaluated with logistic regression analysis. Sample size calculation was performed with 80% power and 5% type I error rate [17].

RESULTS

Demographic and Clinical Features

A total of 80 patients were included in the study (NPWT group=40, control group=40). The mean age of the patients was 58.2±12.4 years in the NPWT group and 60.1±11.8 years in the control group (P=0.452). There were 22 (55%) male and 18 (45%) female patients in the NPWT group, and 28 (70%) male and 12 (30%) female patients in the control group (P=0.143).

When evaluated in terms of comorbidity, 15 (37.5%) patients in the NPWT group had diabetes mellitus and 10 (25%) patients had hypertension, while these rates were 12 (30%) and 14 (35%) in the control group, respectively (P=0.471 and P=0.312, respectively) (Table 1).

Injury Characteristics and Initial Assessment

In terms of injury localization, 28 patients (70%) in the NPWT group presented with lower extremity injuries, 12 (30%) patients presented with upper extremity injuries, while this distribution was 30 (75%) and 10 (25%) in the control group (P=0.602). The Gustilo Type III open fracture rate was determined as 45% (18 patients) in the NPWT group and 35% (14 patients) in the control group (P=0.362).

To address the significant difference in admission time between groups, multivariate logistic regression analysis was performed. After adjusting for admission time, age, Gustilo classification, and comorbidities,

Table 1. Demographic and clinical characteristics

Characteristics	NPWT group	Control group	P value
Age (years)	58.2±12.4	60.1±11.8	0.452
Male patients	22 (55%)	28 (70%)	0.143
Female patients	18 (45%)	12 (30%)	0.143
Diabetes mellitus	15 (37.5%)	12 (30%)	0.471
Hypertension	10 (25%)	14 (35%)	0.312

Data are shown as mean±standard deviation or n (%). NPWT=Negative Pressure Wound Therapy

NPWT remained independently associated with reduced infection risk (adjusted odds ratio [OR]: 0.28, 95% confidence interval [CI]: 0.09-0.87, P=0.027). Stratified analysis by admission time (<6 hours vs ≥6 hours) showed consistent benefits of NPWT in both subgroups (P=0.041 and P=0.039, respectively).

The mean time to presentation to the emergency department was 4.8±2.1 hours in the NPWT group, while it was 6.3 ± 3.5 hours in the control group (P=0.028) (Table 2).

Treatment Parameters and Application Results

The mean pressure value was applied as 125±25 mmHg in the NPWT group and the mean application time was 7±2 days. The frequency of dressing changes was determined as 2.5±0.5 days in the NPWT group and 1.2±0.3 days in the control group (P<0.001).

The number of surgical debridements was determined as 1.3±0.6 in the NPWT group and 2.1±0.8 in the control group (P=0.008). The mean duration of antibiotic use was 10 ± 3 days in the NPWT group and 14±4 days in the control group (P=0.003) (Fig. 1).

Wound Healing and Infection Parameters

Initial wound size was measured as 15.2±4.1 cm² in the NPWT group and 24.8±6.5 cm² in the control group (P=0.012). Granulation tissue formation was observed at 5.1±1.2 days in the NPWT group and 8.7±2.3 days in the control group (P=0.021).

CRP values were determined as 8.5±3.2 mg/L in the NPWT group and 18.9±5.4 mg/L in the control group (P<0.001). CRP normalization time was determined as 7±1.5 days in the NPWT group and 12±2.8 days in the control group (P=0.003) (Fig. 2).

Bacterial culture and antibiotic resistance data were not systematically collected in this study, which represents a limitation in the comprehensive evaluation of infection-related outcomes. This limits our ability to provide detailed microbiological analysis of the observed infections and their response to treatment.

Infection and Complication Rates

Total infection development was found in 3 patients (7.5%) in the NPWT group and 10 patients (25.0%) in the control group (P=0.034). The distribution of these infections was as follows: (a) Superficial surgical site infection: 2 (5.0%) patients in the NPWT group, 7 (17.5%) patients in the control group (P=0.077); (b) Deep surgical site infection: 1 (2.5%) patients in the NPWT group, 3 (7.5%) patients in the control group (P=0.615); and (c) Osteomyelitis: 1 (2.5%) patients in the NPWT group, 4 (10.0%) pa-

Table 2. Injury and treatment characteristics

Characteristics	NPWT group	Control group	P value
Lower extremity injury	28 (70%)	30 (75%)	0.602
Upper extremity injury	12 (30%)	10 (25%)	0.602
Gustilo Type III open fracture	18 (45%)	14 (35%)	0.362
Admission time (hours)	4.8±2.1	6.3±3.5	0.028

Data are shown as mean±standard deviation or n (%). NPWT=Negative Pressure Wound Therapy

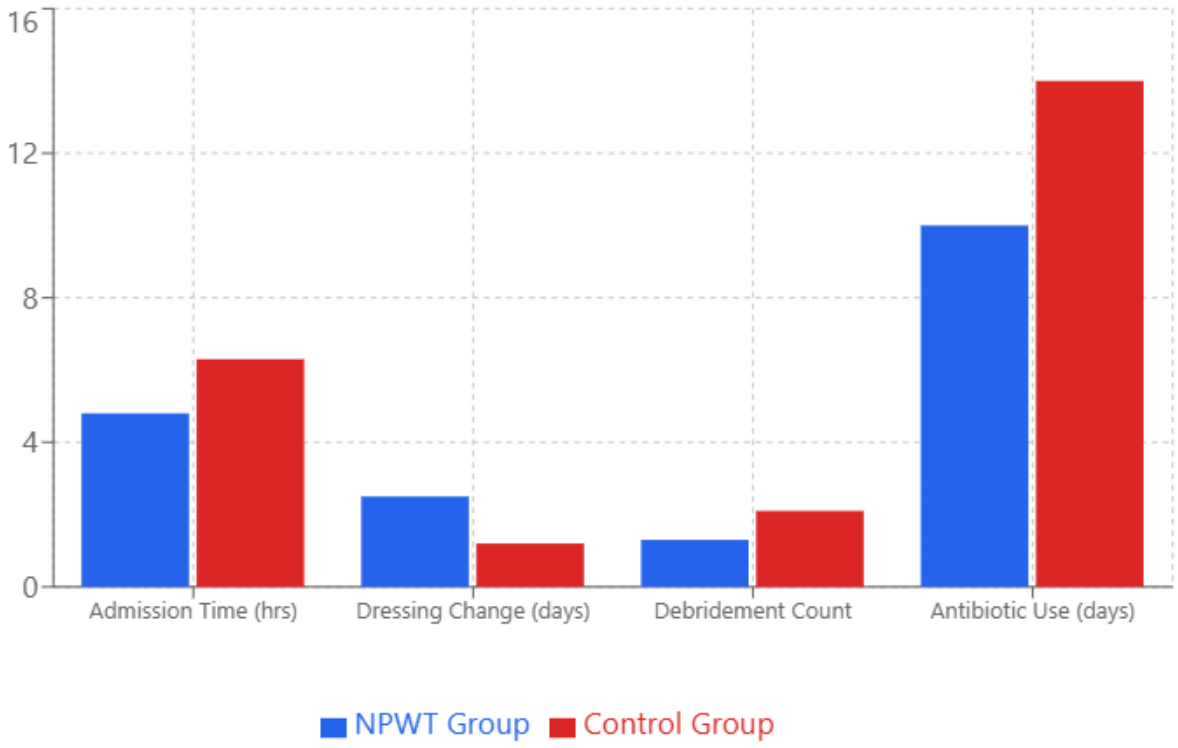

Fig. 1. Clinical and treatment parameters. NPWT=Negative Pressure Wound Therapy.

tients in the control group (P=0.359).

Other complications was the following: (a) Wound dehiscence: 2 (5.0%) patients in the NPWT group, 5 (12.5%) patients in the control group (P=0.432), (b) Seroma: 3 (7.5%) patients in the NPWT group, 8 (20.0%) patients in the control group (P=0.193); (c) Hematoma: 1 (2.5%) patients in the NPWT group, 3 (7.5%) patients in the control group (P=0.615); and (d) Amputation requirement: 0 patients in the NPWT group, 2 (5.0%) in the control group (P=0.494).

Treatment Results and Healing Times

Hospitalization time was determined as 8±2 days in the NPWT group and 14±3 days in the control group (P<0.001). Wound closure time was determined as 21±4 days in theNPWT group and 35±6 days in the control group (P<0.001).

The reduction in wound size was measured as 65±12% in the NPWT group and 35±10% in the control group (P<0.001). The bacterial load reduction rate was determined as 85±9% in the NBYT group and 45±15% in the control group (P<0.001).

Primary closure as the final treatment method could be applied in 28 (70%) patients in the NPWT group and 12 (30%) patients in the control group (P=0.002). Skin graft or flap requirement occurred in 8 (20%) patients in the NBYT group and 22 (55%) patients in the control group. Pain scores (VAS) were evaluated as 2.1±0.8 in the NPWT group and 5.7±1.2 in the control group (P<0.001).

DISCUSSION

This study aimed to evaluate the effect of NPWT applied in the emergency department on postoperative infection rates in orthopedic open injuries. In our study, the infection rate was 7.5% in the NPWT group, while this rate was 25% in the control group. In the study conducted by Kim *et al*. [18] on Gustilo type III open tibia fractures with a mean follow-up of 22.7 months, 15% infection rate, 85% fracture union, mean union time of 31 weeks and 80% independent mobilization rate were reported in patients who underwent NPWT.

The meta-analysis results show that NPWT significantly reduces the overall infection rate (9.8% vs. 13.3%, P=0.005, 95% CI: 0.260-0.683) and acute wound infection rate (7.5% vs. 16.8%, P=0.009) compared to conventional dressings. In addition, no significant difference was found between the groups in

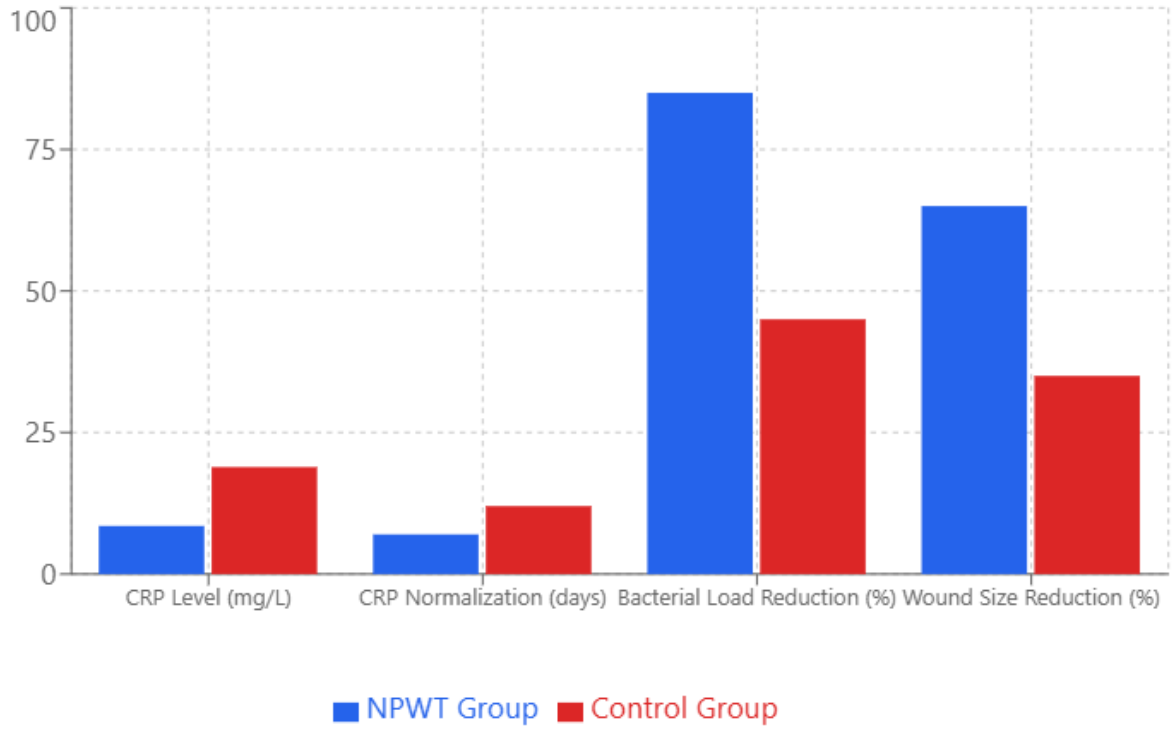

Fig. 2. Healing and infection parameters. NPWT=Negative Pressure Wound Therapy.

terms of functional outcomes in DRI scores at 3 and 6 months and EQ-5D quality of life scale (P=0.99) [19].

Systematic reviews show that NPWTi-d systems provide a significant increase in complete wound closure rates compared to conventional NPWT (OR=2.006, 95% CI: 1.315-3.058, P=0.023). It has been reported that clinical infection clearance time decreased from 25.9 days to 6.0 days and total treatment cost decreased (\$799 vs \$2217) despite higher daily costs (\$194.80 vs \$106.08) [20].

Although a formal cost-effectiveness analysis was not performed in this study, the economic implications of NPWT merit discussion. Literature reports indicate that while NPWT has higher daily costs ($150-250 per day depending on the system), the overall treatment costs may be reduced through several mechanisms [20, 26]. In our study, patients in the NPWT group demonstrated significantly shorter hospital stays (8±2 vs 14±3 days, P<0.001), reduced need for surgical debridement (1.3±0.6 vs 2.1±0.8, P=0.008), and shorter antibiotic treatment duration (10±3 vs 14±4 days, P=0.003).

These factors collectively suggest potential cost savings that could offset the higher daily costs of NPWT. The reduced infection rate (7.5% vs 25.0%) is particularly significant from an economic perspective, as orthopedic infection treatment can cost $50,000-100,000 per case due to prolonged hospitalization, multiple surgeries, and long-term antibiotic therapy [27]. Future studies should include formal cost-effectiveness analyses to provide definitive economic evidence for NPWT implementation in emergency department protocols.

It has been reported in the literature that every 6-hour delay increases the risk of infection by 0.17%, and this increase increases to 0.23% in Gustilo-Anderson type III injuries. The mean debridement time in infected cases was 23.8±32.4 hours, while it was 18.8±27.4 hours in non-infected cases (P<0.001) [21]. In our study, the number of surgical debridements in the NPWT group (1.3±0.6 vs. 2.1±0.8, P=0.008) was found to be significantly lower.

In soft tissue management, definitive closure is recommended within the first 7 days and according to the Gustilo classification, infection rates are reported as 0-2% in Type I, 2-10% in Type II and 10-50% in Type III. The risk of infection in Type III injuries can be up to 50% [22]. In our study, primary closure rate (70% vs. 30%, P=0.002) and granulation tissue development rate (5.1±1.2 vs. 8.7±2.3 days, P=0.021) were

found to be significantly higher in the NPWT group.

In perioperative management, single dose (Co-amoxiclav 1.2g IV or Cefuroxime 1.5g IV), 24 hours in Type II and 72 hours in Type III antibiotic prophylaxis is recommended. It has been reported that Clindamycin 600mg IV can be used in patients with penicillin allergy and the first dose should be administered within 3 hours [23]. In our study, the duration of antibiotic use (10±3 vs. 14±4 days,P=0.003) and CRP normalization time (7±1.5 vs. 12±2.8 days, P=0.003) were found to be significantly shorter in the NPWT group.

According to the OPEN study data, 69.0% of patients are started on antibiotics in the emergency department, 22.0% in the prehospital period, and the rate of NPWT use in complex injuries reaches 67.7%. 90.3% of surgical procedures are performed between 08:00 and 20:00, and the rate of consultant surgeon availability is reported as 89.2% [24].

At the molecular level, NPWT has been shown to shift macrophage polarization to the M2 phenotype by increasing MBD2 gene expression, decreasing inflammatory cytokines such as IL-6 and G-CSF, and increasing IL-12 levels. This immunological modulation plays an important role in controlling bacterial colonization and accelerating wound healing [25].

Strengths and Limitations

Strengths of our study are prospective randomized controlled design, standardized treatment protocol, detailed data collection and analysis, and long follow-up period.

The original contributions of our study to the literature are as follows: (a) Demonstration of the effectiveness of early NPWT application in infection control in the emergency department; (b) Significant increase in the rate of bacterial load reduction (85±9% vs. 45±15%); (c) Significant reduction in CRP normalization time (7±1.5 vs. 12±2.8 days); (d) Decrease in the need for debridement (1.3±0.6 vs. 2.1±0.8); and (e) Shortening of the duration of antibiotic use (10±3 vs. 14±4 days).

Limitations of our study are single center experience, relatively small sample size, possible differences in NPWT application technique, and no cost analysis Methodological limitations are the following: (a) Single-center experience limits generalizability of results to different healthcare settings; (b) Relatively small sample size (n=80); (c) Significant difference in admission time between groups (P=0.028), although this was addressed through multivariate analysis; and (d) Injury severity was not completely homogeneous between groups (Type III injury NPWT group 45%, control group 35%).

Technical limitations are possible differences in standardization of NPWT application technique: (1) NPWT pressure values were not optimized on a patient-by-patient basis; (2) NPWT change times varied depending on clinical indication, and (3) No comparison was made between different NPWT systems.

Data collection limitations are lack of standardized quality of life assessment scales, absence of long-term functional outcome evaluation, limited follow-up period in the evaluation of long-term complications, and lack of systematic bacterial culture and antibiotic resistance data collection.

Economic limitations are also the following a) Lack of formal cost-effectiveness analysis; (b) Absence of detailed calculation of NPWT costs; (c) No evaluation of indirect costs such as rehabilitation and lost productivity; and (d) Limited assessment of healthcare resource utilization beyond hospital stay. Other limitations also included potential influence of surgical team experience and learning curve effects, lack of standardization of variables in emergency room conditions, limited analysis of the interaction between comorbidities and treatment outcomes, and absence of patient-reported outcome measures and satisfaction scores.

Recommendations for Future Research

In future studies, the following should be considered: (1) Conducting multicenter studies with larger sample groups; (2) Comparison of different NPWT protocols in emergency department conditions; (3) Detailed cost-effectiveness analyses; and (4) Evaluation of long-term functional outcomes.

In light of these findings, the following recommendations can be made in clinical practice:

(1) In orthopedic open injuries, NPWT should be started as soon as possible after the initial evaluation in the emergency department

(2) Especially in high-risk patients (Gustilo Type III injuries), NPWT should be the primary choice for infection control

(3) NPWT application should be evaluated as part of the standard protocol to increase the chance of primary closure

(4) For control of bacterial load and inflammatory response, NPWT should be preferred in the early period

CONCLUSION

This prospective randomized controlled trial comprehensively demonstrated the effect of negative pressure wound therapy applied in the emergency department on clinical outcomes in orthopedic open injuries. Unlike the literature, our study evaluated the effects of NPWT initiated early in the emergency department and obtained significant results. The most striking finding of our study is that the infection rate in the NPWT group was significantly lower than in the control group (7.5% vs. 25%). This rate is lower than the general infection rates reported in the literature (9.8%-13.3%), emphasizing the importance of early application. In addition, the significant increase in primary closure rates (70% vs. 30%) demonstrates the positive effect of NPWT on soft tissue healing. In conclusion, this study demonstrates the effectiveness of NPWT initiated in the emergency department in postoperative infection control and suggests a new treatment algorithm in the management of orthopedic open injuries. The results obtained show that early initiation of NPWT in the emergency department is an important strategy that increases treatment success and reduces complications.

Ethics Approval and Consent to Participate

This study was approved by the Adana City Training and Reserach Hospital Scientific Research Ethics Committee (Decision No: 2025-12/447; date: 10.04.2025). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Written informed consent was obtained from all individual participants included in the study.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: SE, SM; Study Design: SE, SM, Vİ; Supervision: SE, SM, Vİ; Funding: SE, SM, Vİ; Materials: SE, SM, Vİ; Data Collection and/or Processing: SE, SM, Vİ; Statistical Analysis and/or Data Interpretation: SE, SM, Vİ; Literature Review: SE, SM, Vİ; Manuscript Preparation: SE; and Critical Review: SE, SM, Vİ.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The author would like to thank the Emergency Department staff and IT staff of Sultanbeyli State Hospital for their support during the data collection process.

Generative Artificial Intelligence Statement

The author(s) declare that artificial intelligence tools were used in accordance with academic ethical standards during the preparation of this manuscript. Google translator was used for language control and also ChatGPT was used in Figs. 1 and 2. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Rupp M, Popp D, Alt V. Prevention of infection in open fractures: Where are the pendulums now? Injury. 2020;51 Suppl 2:S57-S63. doi: 10.1016/j.injury.2019.10.074.

2. Zaver V, Kankanalu P. Negative Pressure Wound Therapy. 2023 Sep 4. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan.
3. Wang B, Xiao X, Zhang J, Han W, Hersi SA, Tang X. Epidemiology and microbiology of fracture-related infection: a multicenter study in Northeast China. J Orthop Surg Res. 2021;16(1):490. doi: 10.1186/s13018-021-02629-6.
4. Van Rysselberghe NL, Gonzalez CA, Calderon C, Mansour A, Oquendo YA, Gardner MJ. Negative Pressure Wound Therapy for Extremity Open Wound Management: A Review of the Literature. J Orthop Trauma. 2022;36(Suppl 4):S6-S11. doi: 10.1097/BOT.0000000000002430.
5. Hollins AW, Mithani SK. Advances in Soft Tissue Injuries Associated with Open Fractures. Hand Clin. 2023;39(4):605-616. doi: 10.1016/j.hcl.2023.05.008.
6. Dvorak JE, Lasinski AM, Romeo NM, Hirschfeld A, Claridge JA. Fracture related infection and sepsis in orthopedic trauma: A review. Surgery. 2024;176(2):535-540. doi: 10.1016/j.surg.2024.04.031.
7. Wu L, Wen B, Xu Z, Lin K. Research progress on negative pressure wound therapy with instillation in the treatment of orthopaedic wounds. Int Wound J. 2022;19(6):1449-1455. doi: 10.1111/iwj.13741.
8. Patro BP, Khuntia S, Sahu NK, Das G, Patra SK. Negative Pressure Wound Therapy Assisted Closure: An Effective Mode of Management for Infected and Contaminated Wound With Non-Union Fracture Femur. Cureus. 2020;12(7):e9037. doi: 10.7759/cureus.9037.
9. Axelrod D, Comeau-Gauthier M, Prada C, et al. Change in Gustilo-Anderson classification at time of surgery does not increase risk for surgical site infection in patients with open fractures: A secondary analysis of a multicenter, prospective randomized controlled trial. OTA Int. 2022;6(1):e231. doi: 10.1097/OI9.0000000000000231.
10. Costa ML, Achten J, Knight R, et al; WHIST Trial Collaborators. Effect of Incisional Negative Pressure Wound Therapy vs Standard Wound Dressing on Deep Surgical Site Infection After Surgery for Lower Limb Fractures Associated With Major Trauma: The WHIST Randomized Clinical Trial. JAMA. 2020;323(6):519-526. doi: 10.1001/jama.2020.0059.
11. Atwan Y, Sprague S, Slobogean GP, et al; FLOW Investigators. Does negative pressure wound therapy reduce the odds of infection and improve health-related quality of life in patients with open fractures? Bone Jt Open. 2022;3(3):189-195. doi: 10.1302/2633-1462.33.BJO-2021-0199.R1.
12. Messori M, Touloupakis G, Gilli A, et al. The risk of infection in open distal tibial fracture: the DANGER score. Eur J Orthop Surg Traumatol. 2023;33(7):2965-2970. doi: 10.1007/s00590-023-03517-x.
13. Garner MR, Sethuraman SA, Schade MA, Boateng H. Antibiotic Prophylaxis in Open Fractures: Evidence, Evolving Issues, and Recommendations. J Am Acad Orthop Surg. 2020;28(8):309-315. doi: 10.5435/JAAOS-D-18-00193.
14. Tait MA, Bracey JW, Roulette P, Lewis DR. Infection Rates in Open Hand Fractures: Can Surgical Treatment Be Delayed? J Surg Orthop Adv. 2024;33(4):222-224.
15. Reátiga Aguilar J, Gonzalez Edery E, Guzmán Badrán J, et al. Open tibial plateau fractures: Infection rate and functional outcomes. Injury. 2023;54 Suppl 6:110720. doi: 10.1016/j.injury.2023.04.007.
16. Gardezi M, Roque D, Barber D, et al. Wound Irrigation in Orthopedic Open Fractures: A Review. Surg Infect (Larchmt). 2021;22(3):245-252. doi: 10.1089/sur.2020.075.
17. Dwivedi AK. How to write statistical analysis section in medical research. J Investig Med. 2022;70(8):1759-1770. doi: 10.1136/jim-2022-002479.
18. Kim Y, Kong CG, Park HY, Lee KS, Sur YJ. Clinical and radiographic outcomes of negative pressure wound therapy combined with polymethylmethacrylate sealant for wound management of Gustilo type III open tibia fractures. Acta Orthop Traumatol Turc. 2023;57(6):366-371. doi: 10.5152/j.aott.2023.22104.
19. Qian H, Lei T, Hu Y. Negative pressure wound therapy versus gauze dressings in managing open fracture wound of lower limbs: A meta-analysis of randomized controlled trials. Foot Ankle Surg. 2022;28(7):1120-1128. doi: 10.1016/j.fas.2022.03.012.
20. De Pellegrin L, Feltri P, Filardo G, et al. Effects of negative pressure wound therapy with instillation and dwell time (NPWTi-d) versus NPWT or standard of care in orthoplastic surgery: A systematic review and meta-analysis. Int Wound J. 2023;20(6):2402-2413. doi: 10.1111/iwj.14072.
21. Cortez A, Urva M, Subramanian A, et al. Delays in Debridement of Open Femoral and Tibial Fractures Increase Risk of Infection. J Bone Joint Surg Am. 2023;105(20):1622-1629. doi: 10.2106/JBJS.23.00074.
22. von Rüden C, Wunder J, Schirdewahn C, Augat P, Hackl S. Initial treatment of severe soft-tissue injuries in closed and open fractures to prevent fracture-related infection. Injury. 2024;55 Suppl 6:111935. doi: 10.1016/j.injury.2024.111935.
23. Loh B, Lim JA, Seah M, Khan W. Perioperative management of open fractures in the lower limb. J Perioper Pract. 2022;32(5):100-107. doi: 10.1177/17504589211012150.
24. Winstanley RJH, Hadfield JN, Walker R, et al; Open-Fracture Patient Evaluation Nationwide (OPEN) collaborators. The Open-Fracture Patient Evaluation Nationwide (OPEN) study: the management of open fracture care in the UK. Bone Joint J. 2022;104-B(9):1073-1080. doi: 10.1302/0301-620X.104B9.BJJ-2022-0202.R1.
25. Jia B, Xue R, Li J, Guo J, Liu J. Novel insights into vancomycin-loaded calcium sulfate and negative pressure wound therapy in preventing infections in open fractures. J Orthop Surg Res. 2024;19(1):517. doi: 10.1186/s13018-024-04931-5.

The European Research Journal 2025;11(6):1132-1139 Original Article

DOI: https://doi.org/10.18621/eurj.1661556 *Anatomy Education*

Satisfaction of medical students with the artistic anatomy lecture: A questionnaire study

Zekiye Karaca Bozdağ, Buse Naz Çandır Gürses

Department of Anatomy, Faculty of Medicine, Istanbul Yeni Yüzyıl University, İstanbul, Türkiye

ABSTRACT

Objectives: The aim of this study is to evaluate the effects of an artistic anatomy lecture on medical students and to identify the role of the lecture in education.

Methods: A 20-question online satisfaction survey was administered to 32 out of 45 medical students who attended the artistic anatomy lecture at our university in the fall semester of 2023-2024. In the survey, questions were asked to reveal the students' performance regarding the relationship between art and anatomy. SPSS Statistics 22 software (IBM SPSS, Turkey) was used for statistical analysis of the data obtained, and $P<0.05$ was considered statistically significant.

Results: Ninety-three point eight percent of students stated that the artistic anatomy lecture contributed to their understanding of the relationship between art and anatomy. Survey results indicated that 87.5% of participants reported the lecture improved their observational skills, while 81.3% said it enhanced their clinical observation skills. In addition, 90.6% of the participants stated that the lecture helped them to understand the importance of art in medical education.

Conclusions: The results show that the artistic anatomy lecture supports the development of important skills such as visual memory, analytical thinking and observational skills in medical students. In line with the literature, arts-based educational approaches enable students to gain an interdisciplinary perspective and develop the ability to pay attention to detail in clinical practice. We suggest that our findings may be useful for integrating artistic anatomy lectures into the curriculum of medical education, and may provide guidance to anatomists in this regard.

Keywords: Artistic anatomy, medical education, anatomy education, clinical observation, visual memory

Medical education is an interdisciplinary process that aims to enable students to learn in detail the structure, function and pathological conditions of the human body [1]. Anatomy is one of the cornerstones of this educational process and plays a critical role in the understanding and application of anatomical structures in clinical practice. Traditional anatomy teaching is usually based on cadaver dissection and plastic models. In recent years, however, more creative and multidisciplinary approaches have been incorporated into medical education. One such approach is artistic anatomy lectures [2, 3]. Seen as a meeting point between art and medicine, these lectures aim to deepen learning by approaching anatomical knowledge from an artistic perspective [4].

Artistic anatomy is a discipline that has developed

Received: March 22, 2025 **Accepted:** April 30, 2025 **Available Online:** June 9, 2025 **Published:** November 4, 2025

How to cite this article: Karaca Bozdağ Z, Çandır Gürses BN. Satisfaction of medical students with the artistic anatomy lecture: A questionnaire study. Eur Res J. 2025;11(6):1132-1139. doi: 10.18621/eurj.1661556

Corresponding author: Zekiye Karaca Bozdağ, PhD, Asst. Prof. Phone: +90 212 444 50 01, E-mail: zekiye.karaca@yeniyuzyil.edu.tr

Available Online at https://dergipark.org.tr/en/pub/eurj

throughout history, with artists such as Leonardo da Vinci and Michelangelo having an in-depth knowledge of the body's structure. Particularly during the Renaissance, anatomy became one of the basic disciplines that nourished art, and knowledge of anatomy became very important in terms of drawing the correct human figure [5].

Art-based anatomy education strengthens students' visual memory and facilitates the learning process by providing an understanding of body structures from a different perspective. Integrating art and anatomy into medical education enables students to acquire not only theoretical knowledge, but also skills such as observation, attention to detail and analytical thinking [6]. Research has shown that arts-based approaches encourage creativity in medical students, improving their problem-solving and visual analysis skills, which are extremely useful when assessing patients in the clinic [7, 8].

Artistic anatomy lectures allow students not only to perceive the human body visually, but also to evaluate the anatomical structure of the body from an artistic point of view. This can be particularly useful for students with visual learning memories. Studies show that medical students learn more effectively using visual learning methods and are able to apply this knowledge more quickly in clinical practice [7, 9].

The aim of this study was to investigate the satisfaction of medical students with the Artistic Anatomy lecture, to evaluate the contribution of the lecture to the students, and to reveal how the connection of the lecture to clinical practice is perceived from the students' perspective. We believe that the results can highlight the impact of Artistic Anatomy on medical education and provide guidance for its integration into medical education.

METHODS

This study was carried out with 45 students of the Faculty of Medicine who took the lecture "Artistic Anatomy", which was opened as an elective lecture in the fall semester 2023-2024 of our university. At the end of the semester, the satisfaction survey prepared with the Google survey application was applied online. The survey form consisted of a total of 20 questions. There were 19 questions on a 5-point Likert scale and 1 open question.

Ethics committee approval for the study was obtained from our university's Science and Health Sciences Research Ethics Committee (IRB: 2024/12-1399).

Statistical Analysis

The conformity of the variables to normal distribution was checked with the Shapiro Wilk Test. The categorical variables were compared between the groups by using the Pearson Chi-Square, Fisher's Exact Chi-Square, Fisher Freeman Halton Test, and McNemar Test. According to the normality test results, the Mann-Whitney U Test was used for the comparisons between binary groups, and the Wilcoxon Test was used for the comparisons between dependent samples. Categorical variables were reported as n (%). Fisher Freeman Halton test was used for comparing categorical variables. The Statistics Statistical Analysis SPSS 22 (SPSS IBM) program was used, and $P<0.05$ was considered significant.

RESULTS

This lecture was attended by 45 students and 32 students responded to the questionnaire. Among the students who participated in the study, 90.7% stated that they chose the Art Anatomy lecture voluntarily. Before the lecture, 56.3% of the students were interested in visual arts, whereas 18.7% were not. While 37.5% of the students were aware of the relationship between art and anatomy before the lecture, 25% were not aware of the relationship and 37.5% were undecided. On the other hand, the students who learned about the relationship between art and anatomy after the artistic anatomy lecture represent 90.6% of the students who participated in the study. A total of 93.8% of the students said that the art anatomy lecture showed that anatomy and art are intertwined. However, the students who agreed with the statement that this lecture would provide a deeper understanding of anatomy for medical students represented 75% of the students who participated in the study. The statement "Observation, imagination and creativity are very important when drawing anatomy" was accepted by 93.8% of the students, while 6.2% were undecided. In total, 84.4% of the participating students agreed with the statement

Table 1. Relationship between voluntary lecture choice and the importance of anatomy knowledge

		"I choose the lecture voluntarily"						
		0	**1**	**2**	**3**	**4**	**Total**	**P value**
"Anatomy knowledge is essential for good and accurate human works'.	**0**	0	0	0	0	0	0	**<0.001[a]**
	1	0	0	0	0	1 (100%)	1 (100%)	
	2	0	0	0	0	1 (100%)	1 (100%)	
	3	0	1 (14.3%)	1 (14.3%)	5 (71.4%)	0	7 (100%)	
	4	0	1 (4.3%)	0	1 (4.3%)	21 (91.3%)	23 (100%)	

[a]Fisher Freeman Halton Test

5-point Likert scale (0: Strongly Disagree; 1: Disagree 2: Undecided 3: Agree 4: Strongly Agree)

"The ability to analyse art plays an important role in improving the clinical observation skills of medical students." According to the survey, 87.5% of students stated that this lecture improved their observation skills, and 81.3% said that it improved their clinical observation skills. After taking the Artistic Anatomy lecture, 90.6% of the students stated that they understood the importance of art in medical education. And 93.8% of the students agreed with the statement "Knowledge of anatomy is essential for good and accurate human works". The anatomical structures that give aesthetics and shape to the external structure of the body, the contours formed by these structures under the skin, and the organs that give perspective to the superficial structure of the body have contributed to my understanding of superficial and topographical anatomy. Among the participants, 87.5% agreed with this statement, 9.4% were undecided, and 3.1 % disagreed. While 81.2% of the students who participated in the study agreed that the subjects taught in the Artistic Anatomy lecture were useful for their professional development, 9.4% stated that they did not find it useful. It was found that 87.5% of the students thought that it was a great privilege to be able to draw anatomical structures correctly while training to become a doctor. It was found that 87.5% of the students agreed with the statement 'depicting the human body through art is a wonderful experience', 3.1% disagreed and 9.4% were undecided. A proportion of 87.5 % of the students stated that their interest in both art and the anatomy lecture increased thanks to the enjoyable artistic anatomy sessions. When the students were asked to rate their satisfaction with the Art Anatomy lecture in general, taking into account all the conditions highlighted in the study, it was found that 93.8% of the students who took this lecture and participated in the study were satisfied and 75% of them were very satisfied.

In response to the open-ended question, 'What do you think is the most important thing you have learnt during this lecture?', the most common answers were:

Table 2. Relationship between lecture impact and art analysis in enhancing clinical observation skills

		"This lecture improved my observation skills"						
		0	**1**	**2**	**3**	**4**	**Total**	**P value**
"Being able to analyse art plays an important role in improving medical students"	**0**	0	0	0	0	0	0	**<0.001[a]**
	1	0	0	0	0	0	0	
	2	2 (40%)	0	1 (20%)	2 (40%)	0	5 (100%)	
	3	0	0	0	6 (75%)	2 (25%)	8 (100%)	
	4	0	0	1 (5.3%)	1 (5.3%)	17 (89.5%)	19 (100%)	

[a]Fisher Freeman Halton Test

5-point Likert scale (0: Strongly Disagree; 1: Disagree 2: Undecided 3: Agree 4: Strongly Agree)

Table 3. Association between lecture impact and art analysis in improving clinical observation skills

		"This lecture improved my clinical observation skills"						
		0	**1**	**2**	**3**	**4**	**Total**	**P value**
"Being able to analyse art plays an important role in improving the clinical observation skills of medical students"	**0**	0	0	0	0	0	0	**0.002[a]**
	1	0	0	0	0	0	0	
	2	2 (40%)	1 (20%)	1 (20%)	1 (20%)	0	5 (100%)	
	3	0	0	0	4 (50%)	4 (50%)	8 (100%)	
	4	0	0	2 (10.5%)	3 (15.8%)	14 (73.7%)	19 (100%)	

[a]Fisher Freeman Halton Test
5-point Likert scale (0: Strongly Disagree; 1: Disagree 2: Undecided 3: Agree 4: Strongly Agree

'I have learnt that anatomy is not just a lesson, it is in every aspect of our lives. Without art, there would be great deficiencies in the most important part of medical science today. The pioneer in the development of anatomy is its intertwining with art'. The flow of the lecture was excellent. We realised once again that anatomy and art are intertwined. It was very important to understand and grasp anatomy through drawings. It was a fun and educational lesson.' To be a good doctor, it is imperative to know the science of anatomy, and artistic anatomy has multiplied this importance. With this lecture I saw the wonderful combination of anatomy and art and it increased my interest in art.

It was determined that there was a statistical difference between the answers given to the question 'I chose the lecture voluntarily' and the answers given to the question 'anatomy knowledge is essential for good and accurate human works' (P<0.001). The rate of those who agreed with the question 'knowledge of anatomy is essential for human works with good and accurate measurements' was higher in the group of those who agreed with the question 'I chose the anatomy lecture voluntarily'. The rate of those who strongly agree with the question 'knowledge of anatomy is essential for good and accurate human works' is higher in the group of those who strongly agree with the question 'I chose the anatomy lecture voluntarily' (Table 1). When the answers given to the question 'I chose the lecture voluntarily' were compared with the answers given to the other questions in the questionnaire, it was determined that there was no difference between the answers (P>0.05).

It was found that there was a difference between the responses to the question 'This lecture improved

Table 4. Relationship between drawing anatomical structures and understanding superficial and topographic anatomy

		"It is a great privilege to be able to draw anatomical structures correctly while studying to become a doctor"						
		0	**1**	**2**	**3**	**4**	**Total**	**P value**
"The anatomical structures that give aesthetic and shape to the external structure of the body, the contours formed by these structures under the skin and the organs that give a perspective view to the superficial structure of the body contribute to my understanding of superficial and topographic anatomy"	**0**	0	0	0	0	0	0	**0.016[a]**
	1	1(100%)	0	0	0	0	1 (100%)	
	2	0	0	1(33.3%)	1(33.3%)	1 (33.3%)	3 (100%)	
	3	0	0	2(22.2%)	1(11.1%)	6 (66.6%)	9 (100%)	
	4	0	0	0	3(15.8%)	16 (84.2%)	19 (100%)	

[a]Fisher Freeman Halton Test
5-point Likert scale (0: Strongly Disagree; 1: Disagree 2: Undecided 3: Agree 4: Strongly Agree)

Table 5. Relationship between drawing anatomical structures and professional development

		"I think it is a great privilege to draw anatomical structures correctly during my training to become a doctor"						
		0	**1**	**2**	**3**	**4**	**Total**	**Pvalue**
"I think the subjects are useful for my development in terms of my profession"	**0**	0	0	0	0	1 (100%)	1 (100%)	**0.001[a]**
	1	1 (50%)	0	0	0	1 (50%)	2 (100%)	
	2	0	0	3 (100%)	0	0	3 (100%)	
	3	0	0	0	2 (50%)	2 (50%)	47 (100%)	
	4	0	0	0	3 (13.6%)	19 (86.4%)	22 (100%)	

[a]Fisher Freeman Halton Test

5-point Likert scale (0: Strongly Disagree; 1: Disagree 2: Undecided 3: Agree 4: Strongly Agree)

my observation skills' and the responses to the question 'Being able to analyse art plays an important role in improving medical students' clinical observation skills' (P<0.001). The proportion of those who agreed with the statement 'Being able to analyse art plays an important role in improving the clinical observation skills of medical students' was higher in the group of those who agreed with the statement 'This lecture improved my observation skills'. The proportion of those who strongly agreed with the statement 'Being able to analyse art plays an important role in developing medical students' clinical observation skills' was higher in the group who agreed with the statement 'This lecture improved my observation skills'. The proportion of students who strongly agreed with the statement 'This lecture has improved my observational skills' was higher in the group. In summary, students who believe that art analysis plays an important role in improving medical students' clinical observation skills also believe that the lecture has improved their observational skills (Table 2).

It was determined that there was a difference between the answers given to the question 'This lecture improved my clinical observation skills.' and the answers given to the question 'Being able to analyse art plays an important role in improving the clinical observation skills of medical students' (P=0.002). The rate of those who answered 'I am undecided' to the question 'Being able to analyse art plays an important role in improving the clinical observation skills of medical students' is higher in the group of those who strongly disagree with the question 'This lecture improved my clinical observation skills' (Table 3).

There was a statistically significant relationship between the question 'It is a great privilege to be able to draw anatomical structures correctly while studying

Table 6. Statistical relationship between representing the human body through art and professional development

		"It is a great experience to represent the human body through art"						
		0	**1**	**2**	**3**	**4**	**Total**	**P value**
"I think the subjects are useful for my development in terms of my profession"	**0**	0	0	0	0	1 (100%)	1 (100%)	**0.004[a]**
	1	1 (50%)	0	1 (50%)	0	0	2 (100%)	
	2	0	0	1 (33.3%)	2 (66.7%)	0	3 (100%)	
	3	0	0	0	2 (50%)	2 (50%)	4 (100%)	
	4	0	0	1 (4.5%)	3 (13.6%)	18 (81.8%)	22 (100%)	

[a]Fisher Freeman Halton Test

5-point Likert scale (0: Strongly Disagree; 1: Disagree 2: Undecided 3: Agree 4: Strongly Agree)

to become a doctor.' and the question 'The anatomical structures that give aesthetic and shape to the external structure of the body, the contours formed by these structures under the skin and the organs that give a perspective view to the superficial structure of the body contribute to my understanding of superficial and topographic anatomy' (P=0.016). Students who think that the body's external and superficial structures contribute to understanding anatomy also think it is a privilege to draw them correctly during medical training (Table 4).

There was a difference in responses to two questions: 'I think it is a great privilege to draw anatomical structures correctly during my training to become a doctor' and 'I think the subjects are useful for my development in terms of my profession' (P=0.001). Students who think the subjects are useful for their profession consider it a privilege to draw anatomical structures correctly while studying to become a doctor (Table 5).

The answers to the questions 'It is a great experience to represent the human body through art' and 'I think the subjects are useful for my development in terms of my profession' differed (P=0.04). Students who think that the subjects are useful for their development in terms of their profession think that it is a great experience to represent the human body through art (Table 6).

DISCUSSION

Artistic anatomy is an educational approach in which the anatomical structures of the human body are addressed from both a scientific and artistic perspective. These lectures are usually located at the intersection of art and medicine disciplines and aim to provide students with a deeper understanding of anatomical knowledge through artistic methods [10].

While artists learn anatomy to accurately represent body structures, medical students can use this approach as a method that enriches their learning process. Artistic anatomy lectures use drawing, sculpture, and other visual art techniques to understand the shape, volume, and movements of the body [11, 12]. In our study, it was observed that the artistic anatomy lecture made significant contributions in different areas such as visual memory, observational skills, and interdisciplinary thinking of medical students.

Visual Learning and Memory

Our study showed that the Artistic Anatomy lecture improved students' visual memory and analytical skills. In particular, 87.5% of students reported that their observation skills had improved as a result of the lecture. The importance of visual memory in medical education is widely discussed in the literature. For instance, previous research has emphasized the significance of visual memory, particularly in grasping topics that rely on visual elements such as surface anatomy. The findings indicate that when students engage in a learning process centered on visual memory during anatomy lectures, they retain the information more effectively [9]. Similarly, Tyler and Likova found that arts-based teaching strengthens visual memory and helps students to better remember what they have learned. The results of our study are consistent with these findings in the literature, confirming that art is an effective way to visualise knowledge in the medical education process. Therefore, the fact that students in our study reported that they strengthened their visual memory by taking this lecture supports previous research that art provides memory retention in the educational process [13].

Contribution to Observation and Analytical Thinking

Another important finding of our study was that the artistic anatomy lecture reinforced students' observation and analytical thinking skills. Students reported that the lecture increased their clinical observation skills 81.3% and helped them understand the connection between art and anatomy 90.6%. The effect of art-based education on observation skills has long been discussed in the literature. Mehta and Agius [14] revealed that observation through art improves students' clinical skills and enables them to pay more attention to details. They reported that art strengthens analytical thinking skills in students and allows them to achieve more successful results, especially in areas that require attention, such as patient assessment. In this context, the fact that the students in our study stated that their observational skills improved coincides with the findings in the literature that art-based anatomy education can be beneficial in clinical practice [14]. Another study found that the arts improved students' skills such as attention to detail and analytical thinking [6]. The

proportion of students in our study who reported improved observational skills is consistent with previous research that arts-based education improves not only visual memory, but also attention to detail and analytical thinking skills.

Interdisciplinary Learning and Integration of Art and Medicine

In our study, 90.6% of the students stated that the artistic anatomy lecture enabled them to make the link between art and anatomy and provided an interdisciplinary perspective. The contribution of interdisciplinary education to medical education is widely discussed in the literature. Emara stated that interdisciplinary education provides medical students with information integrity and contributes to the acquisition of different perspectives in clinical practice. This helps students to develop not only anatomical structures, but also aesthetic elements and analytical skills [15]. In this context, the results of our study support the interdisciplinary contribution of arts-based education to medical students. Bardes *et al*. [16] also found that art-based anatomy lectures not only improve students' anatomical knowledge, but also enhance their aesthetic perspective and support analytical thinking. In line with these findings, we see in our study that students' awareness of art has increased and this awareness, combined with medical knowledge, provides more comprehensive learning.

Convenience in Clinical Practice and the Role of Art

The usefulness and contribution of art anatomy lectures to clinical practice was also observed in our study. Of the students who participated in the study, 81.3% reported that the art anatomy lecture contributed to their clinical observation skills. This finding supports studies in the literature that address the contribution of art-based education to clinical assessment and observation skills. Weiss and Casazza [17] found that art-based education helps medical students to be better equipped in their transition to clinical practice and improves their observational skills. In this context, the proportion of students in our study who reported that the art-based anatomy lecture contributed to their clinical observation skills is consistent with the literature highlighting the clinical benefits of integrating art into medical education.

The results of our study demonstrate that art-based anatomy education enhances medical students' development in areas such as visual memory, observation skills, interdisciplinary thinking, and clinical practice. These findings, which are also supported by the literature, suggest that integrating art into medical education provides students with a more holistic learning experience. Previous research has indicated that art positively influences the development of visual memory and observation skills, aligning with the results of our study [9, 13]. Additionally, studies have shown that art contributes to clinical observation and analytical thinking skills, supporting the observed improvement in observation skills in our findings [6, 14]. Research has also highlighted that interdisciplinary education fosters information integration and helps students gain diverse perspectives; in this regard, the increased awareness of art observed in our study underscores the benefits of interdisciplinary approaches [15, 16]. Finally, literature on arts-based education suggests that it enhances clinical observation skills in practice, which is consistent with the outcomes of our study [17].

Limitations

This study was conducted at a single center with a small, volunteer sample (n = 32). Consequently, selection bias may have been introduced, and because the data rely solely on self-report, objective performance measures and long-term follow-up are lacking.

CONCLUSION

Evaluating the results obtained in comparison with the existing studies in the literature, we can say that art-based anatomy lectures have positive effects on medical education. We believe that the initial results of our study can guide anatomy educators on how much artistic anatomy can be included in medical education.

Ethics Approval and Consent to Participate

This study was approved by the İstanbul Yeni Yüzyıl University Science and Non-Medical Interventional Health Sciences Research Ethics Committee (Decision no.: 2024/12-1399, date: 03.12.2024).

Data Availability

All data generated or analyzed during this study are

included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: ZKB, BNÇG; Study Design: ZKB, BNÇG; Supervision: N/A; Funding: N/A; Materials: ZKB, BNÇG; Data Collection and/or Processing: ZKB, BNÇG; Statistical Analysis and/or Data Interpretation: ZKB, BNÇG; Literature Review: ZKB; Manuscript Preparation: ZKB and Critical Review: ZKB, BNÇG.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Louw G, Eizenberg N, Carmichael SW. The place of anatomy in medical education: AMEE Guide no 41. Med Teach. 2009;31(5):373-386. doi: 10.1080/01421590902825149.
2. Elbe H, Gümüş E, Koçak EN, Çetinavcı D, Yücel A. Art-Based Special Study Module: Improving Visual Thinking and Communication Skills of Medical Students. World Med Educ. 2023;22(67):15-28. doi: 10.25282/ted.1249701.
3. Turney BW. Anatomy in a modern medical curriculum. Ann R Coll Surg Engl. 2007;89(2):104-107. doi: 10.1308/003588407X168244.
4. Naghshineh S, Hafler JP, Miller AR, et al. Formal art observation training improves medical students' visual diagnostic skills. J Gen Intern Med. 2008;23(7):991-997. doi: 10.1007/s11606-008-0667-0.
5. Dalia Y, Milam EC, Rieder EA. Art in Medical Education: A Review. J Grad Med Educ. 2020;12(6):686-695. doi: 10.4300/JGME-D-20-00093.1.
6. Shapiro J, Rucker L, Beck J. Training the clinical eye and mind: using the arts to develop medical students' observational and pattern recognition skills. Med Educ. 2006;40(3):263-268. doi: 10.1111/j.1365-2929.2006.02389.x.
7. Kibibi Wairimu H. Innovative Approaches to Integrating Arts in Medical Education. Eurasian Exp J Sci Appl Res. 2024;5(3):66-69.
8. Pories SE, Piawah S, Abel GA, Mullangi S, Doyle J, Katz JT. What is the role of the arts in medical education and patient care? A survey-based qualitative study. J Med Humanit. 2018;39(4):431-445. doi: 10.1007/s10912-018-9530-4.
9. Aspanani A, Sadeqhi H, Omid A. The relationship between visual memory and spatial intelligence with students' academic achievement in anatomy. BMC Med Educ. 2023;23(1):336. doi: 10.1186/s12909-023-04327-9.
10. Bell LT, Evans DJ. Art, anatomy, and medicine: Is there a place for art in medical education? Anat Sci Educ. 2014;7(5):370-378. doi: 10.1002/ase.1435.
11. Na Y, Clary DW, Rose-Reneau ZB, et al. Spatial Visualization of Human Anatomy through Art Using Technical Drawing Exercises. Anat Sci Educ. 2022;15(3):587-598. doi: 10.1002/ase.2080.
12. Borrelli M, Leung B, Morgan M, Saxena S, Hunter A. Should drawing be incorporated into the teaching of anatomy? J Contemp Med Educ. 2018;6(2):34-38. doi: 10.5455/jcme.20180411105347.
13. Tyler CW, Likova LT. The role of the visual arts in enhancing the learning process. Front Hum Neurosci. 2012;6:8. doi: 10.3389/fnhum.2012.00008.
14. Mehta A, Agius S. The Use of Art Observation Interventions to Improve Medical Students' Diagnostic Skills: A Scoping Review. Perspect Med Educ. 2023;12(1):169-178. doi: 10.5334/pme.20.
15. Emara S. Doctors without borders-integrating visual art in the medical school curriculum. J Vis Commun Med. 2024;47(2):70-74. doi: 10.1080/17453054.2024.2431804.
16. Bardes CL, Gillers D, Herman AE. Learning to look: developing clinical observational skills at an art museum. Med Educ. 2001;35(12):1157-1161. doi: 10.1046/j.1365-2923.2001.01088.x.
17. Weiss V, Casazza K. Using the art of medical illustration to enhance anatomy education. Int J Arts Humanit Soc Sci. 2021;2(9):1-5.

The European Research Journal 2025;11(6):1140-1146 Original Article

DOI: https://doi.org/10.18621/eurj.1634522 *Emergency Medicine*

Assessment of end-tidal carbon dioxide levels in patients presenting to the emergency department with gastrointestinal bleeding

Emine Kılınç[1], Rezan Karaali[2], Ahmet Kayalı[3], Ecem Ermete Güler[4], Hüseyin Acar[3], Fatih Esad Topal[3]

[1]Deparment of Emergency Medicine, İzmir Buca Seyfi Demirsoy Training and Research Hospital, İzmir, Türkiye; [2]Deparment of Emergency Medicine, İzmir Demokrasi University Faculty of Medicine, İzmir, Türkiye; [3]Deparment of Emergency Medicine, İzmir Katip Çelebi University Faculty of Medicine, İzmir, Türkiye; [4]Deparment of Emergency Medicine, İzmir Atatürk Training and Research Hospital, İzmir, Türkiye

ABSTRACT

Objectives: Gastrointestinal bleeding is a common condition in emergency departments and can be fatal if diagnosis and treatment are delayed. In this study, we aimed to explore the relationship between end-tidal carbon dioxide ($ETCO_2$) levels and Glasgow Blatchford Score (GBS) and AIMS65 scores, as well as its impact on assessing morbidity and mortality in patients presenting to the emergency department with gastrointestinal bleeding.

Methods: The research involved 103 eligible patients diagnosed with gastrointestinal bleeding. $ETCO_2$ measurements were taken on admission and data on hospitalization, GBS/AIMS65 scores, endoscopically detected active bleeding and 30-day mortality were recorded. Statistical analysis was performed on the collected data.

Results: When $ETCO_2$ values obtained from the patients were compared according to hospitalization status, GBS score, AIMS65 score, presence of endoscopically detected active bleeding and mortality status; $ETCO_2$ levels were significantly lower in patients with active bleeding, those who died, patients with AIMS65 scores ≥ 2, and those with GBS scores ≥ 12 ($P<0.05$).

Conclusions: This study demonstrates that ETCO2 levels are significantly lower in patients with gastrointestinal bleeding, especially in those with active bleeding, high mortality risk, and elevated GBS or AIMS65 scores. $ETCO_2$ may serve as a rapid and practical marker for assessing hypovolemia and clinical status in emergency settings.

Keywords: Gastrointestinal bleeding, end-tidal capnography, Glasgow Blatchford score, AIMS65 score

Gastrointestinal (GI) bleeding is a common condition in emergency departments and can lead to fatal outcomes if diagnosis and treatment are delayed. Upper GI bleeding represents 5% of emergency department admissions [1]. Mortality rates range from 2% to 15% [2]. In most cases, GI bleeding may resolve spontaneously without the need for endoscopic intervention, blood transfusion, or surgery. Nevertheless, timely intervention is crucial for patients with life-threatening bleeding [3]. For this purpose,

Received: February 6, 2025 **Accepted:** March 11, 2025 **Available Online:** June 5, 2025 **Published:** November 4, 2025

How to cite this article: Kılınç E, Karaali R, Kayalı A, Ermete Güler E, Acar H, Topal FE. Assessment of end-tidal carbon dioxide levels in patients presenting to the emergency department with gastrointestinal bleeding. Eur Res J. 2025;11(6):1140-1146. doi: 10.18621/eurj.1634522

Corresponding author:Ecem Ermete Güler, MD., Phone: +90 232 243 43 43, E-mail: ecemermete@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

commonly used and validated risk stratification tools, such as the Glasgow Blatchford Score (GBS) and the AIMS65 score, are utilized [4, 5].

Capnography is a noninvasive method for measuring the partial pressure of carbon dioxide (CO_2) throughout the respiratory cycle. It offers valuable insights into ventilation (the effectiveness of carbon dioxide removal), perfusion (vascular transport of CO_2), and metabolism (CO_2 production from cellular processes). The principle behind end-tidal capnography is to detect carbon dioxide levels in the exhaled breath. Variations in the end-tidal CO_2 ($ETCO_2$) waveform provide physicians with crucial information in various scenarios, such as evaluating disease severity, assessing cardiac arrest (quality of chest compressions, return of spontaneous circulation, correct endotracheal tube placement, prognosis), guiding procedural sedation, and predicting critical illness [6-13].

In this study, we sought to explore the relationship between $ETCO_2$ levels and GBS and AIMS65 scores, and assess their effectiveness in evaluating morbidity and mortality in patients presenting to the emergency department with GI bleeding.

METHODS

This study was conducted as a prospective observational study at İzmir Atatürk Training and Research Hospital. The study commenced following the approval of the ethics committee (dated 12.05.2020, approval number 702) and was conducted from 1st June 2020 to 31st December 2020. Patients who were admitted to our emergency department, a tertiary education and research hospital, diagnosed with GI bleeding, and met the inclusion criteria were included in the study.

Inclusion criteria were patients presenting to the emergency department with signs of GI bleeding (blood from the mouth, blood from the rectum, black-colored stools, vomiting like coffee grounds, fainting), those who consented to participate in the study, and patients aged 18 years and older. Exclusion criteria were patients under 18 years of age, pregnant women, trauma patients, individuals with respiratory conditions that elevate CO_2 levels (such as asthma, COPD, pneumonia, etc.), patients transferred from other hospitals, those with heart failure, and patients who declined participation in the study.

Upon admission to the emergency department, ETCO2 measurement was performed with a UTECH VS2000® capnography device under 2 L/min oxygen support for approximately 5 minutes using a Mediplus A202Mx capnomask® mask and recorded on the prepared case form. Demographic data (age, gender), blood pressure, and pulse rate, comorbid diseases (liver failure, heart failure), history of syncope, and the presence of melena on rectal examination were recorded on the patient's form. Endoscopy /colonoscopy results (with/without active bleeding), international normalized ratio (INR), albumin levels, and blood urea nitrogen (BUN) levels, and hemoglobin (Hb) values of laboratory parameters were recorded via the automation system. GBS and AIMS65 scores were calculated using the obtained

Table 1. Glasgow Blatchford score (GBS)

Admission risk marker	Score component value
Blood urea (mmol/L)	
6.5-8.0	2
8.0-10.0	3
10.0-25.0	4
>25.0	6
Haemoglobin (g/dL) for men	
12.0-12.9	1
10.0-11.9	3
<10.0	6
Haemoglobin (g/dL) for women	
10.0-11.9	1
<10.0	6
Systolic blood pressure (mm Hg)	
100-109	1
90-99	2
<90	3
Other markers	
Pulse ≥100/min	1
Melaena	1
Syncope	2
Hepatic disease	2
Cardiac failure	2

data (Tables 1 and 2). Patients were grouped as GBS≥12 and GBS<12 according to the GBS score and AIMS≥2 and AIMS<2 according to the AIMS65 score [14, 15]. Hospitalization status (discharged, admission to the ward, admission to the intensive care unit) was recorded. Within 30 days, any death due to GI bleeding was followed up, and 30- day mortality was recorded.

Sample Size

The sample size was calculated using G Power 3.1.9.6. Since there was no other study similar to ours, the sample size was determined based on the effect size. With the effect size established at 0.7, the sample size was calculated as 45 for each group, totaling 90 total.

Statistical Analysis

The data collected were analyzed using the SPSS 23.0 software package. For demographic variables, categorical data such as gender were expressed as frequencies and percentages, while other numerical variables were presented as means and standard deviations. The normality of continuous variables in independent groups was assessed using histogram curves and the ShapiroWilk test. For comparisons between two independent groups, the Student's t-test was used for normally distributed data, while the Mann-Whitney U test was applied for non-normally distributed data. One-way ANOVA was used to compare more than two groups, as the data were normally distributed. A P-value of <0.05 was considered statistically significant in all tests.

RESULTS

During the study period, a total of 402 patients presented to the emergency department with symptoms of gastrointestinal bleeding. Of these, 191 were diagnosed with GI bleeding, and 103 patients who met the inclusion criteria were included in the study. The mean age of the patients was 67.89±16.55 years, with 51.5% of the participants being male.

Syncope occurred in 4.9% of the patients. Additionally, 18.4% had an AIMS65 score ≥2, and 77.7% had a Glasgow-Blatchford score (GBS) ≥12. 32.1% of our patients were hospitalized in the intensive care unit (ICU), while 42.7% were admitted to the general ward. The mortality rate was 18.4% (Table 3). When comparing the $ETCO_2$ values obtained from patients based on hospitalization status, GBS score, AIMS65 score, presence of actively bleeding detected endoscopically, and mortality status, it was found that $ETCO_2$ values were statistically significantly lower in patients with active bleeding, deceased patients, patients with AIMS65 ≥2, and patients with GBS >12 (P<0.05) (Table 4).

Table 2. AIMS65 Score

Variable	Score
Albumin level <3.0 mg/dL	1
INR >1.5	1
Altered mental status	1
Systolic blood pressure ≤90 mm Hg	1
Age >65 years	1

INR=international normalized rate

Table 3. Demographic characteristics of patients and distribution of variables

Variable		Data
Age (years)		67.89±16.55
Gender	Male	53 (51.5)
	Female	50 (48.5)
Syncope	Yes	5 (4.9)
	No	98 (95.1)
Melena	Yes	59 (57.3)
	No	44 (42.7)
Active bleeding	Yes	52 (49.5)
	No	51 (50.5)
AIMS65	<2	91 (81.6)
	≥2	12 (18.4)
GBS	<12	23 (22.3)
	≥12	80 (77.7)
Hospitalization	ICU	33 (32.1)
	Ward	44 (42.7)
	Discharged	26 (25.2)
Mortality	Yes	19 (18.4)
	No	84 (81.6)

Data are shown as mean±standard deviation or n (%).
GBS=Glasgow Blatchford score, ICU=Intensive care unit

Table 4. Relationship of ETCO2 value with AIMS65, GBS and patient outcomes

Variable		ETCO2	P value
Active bleeding	Yes	23.34±5.40	**0.031***
	No	26.01±6.80	
AIMS65	AIMS<2	25.08±6.11	**0.026***
	AIMS≥2	21.48±6.91	
GBS	GBS<12	28.44±5.95	**0.001***
	GBS≥12	23.58±5.97	
Hospitalization	ICU	22±23±6.02	0.481***
	Ward	24.91±5.99	
	Discharged	27.32±6.12	
Mortality	Exitus	20 (15.70-23.40)	**0.002****
	Survivor	26.10 (20.70-29.58)	

Data are shown as mean±standard deviation or median (IQR:25-75). ETCO2=end-tidal carbon dioxide, GBS=Glasgow Blatchford score, ICU=Intensive care unit.

*Student T test, **Mann whitney U test, ***One way ANOVA

DISCUSSION

Gastrointestinal bleeding is an important clinical picture because it is frequently encountered as a reason for presentation to the emergency department, associated mortality, high cost of diagnosis and treatment, and may require hospitalization and subsequent intensive care follow-up. Predicting the severity of the disease and mortality is important in the approach to the emergency unit [3]. Previous studies have established $ETCO_2$ as an effective predictor of critical illness and mortality. Ladde *et al*. [16] assessed the utility of $ETCO_2$ in emergency department triage, finding mean levels of 33 mmHg (range 32-34) in survivors and 22 mmHg (range 18-26) in deceased patients [16]. They also highlighted $ETCO_2$ as a valuable indicator in patients requiring intensive care unit admission. In contrast, our study specifically focused on patients presenting to the emergency department with gastrointestinal bleeding. Among studies evaluating $ETCO_2$ in hemorrhagic conditions, Guzman *et al*. [17] demonstrated its effectiveness in detecting supply dependency, noting rapid increases in $ETCO_2$ correlating with changes in oxygen consumption during resuscitation in experimentally induced hemorrhagic shock in dogs. Similarly, Okamoto *et al*. [18] investigated $ETCO_2$'s relationship with circulation and respiration in canine models, showing $ETCO_2$ elevation alongside increases in blood flow, cardiac output, and hemoglobin concentration, influencing CO_2 elimination rates. Belenky *et al*. [19] reported that $ETCO_2$ levels below 20mm Hg predict hemorrhagic shock severity, decreasing further with shock progression in animal models. These experimental studies underscore the impact of blood volume reduction on $ETCO_2$ levels via CO_2 transfer dynamics influenced by blood flow, hemoglobin, and cardiac output. In human studies, Bulger *et al*. [20] noted lower $ETCO_2$ levels (26.5 mmHg vs. 32.5 mmHg) in trauma patients with hemorrhagic shock compared to those without, emphasizing its potential as an indicator for hemorrhagic shock. Wilson *et al*. [21] identified a threshold $ETCO_2$ value of 35 mmHg for predicting hemorrhagic shock and massive transfusion need in trauma patients. In the study by O'Connor *et al*. [22], were evaluated 56 patients admitted to the emergency department with GI bleeding. The mean $ETCO_2$ value was measured as 30.52±5.30 mmHg in the group with bleeding and 37.43±4.11 mmHg in the group without bleeding and $ETCO_2$ value was found to be significantly lower in patients with bleeding [22]. Similarly, our measurements in GI bleeding patients showed $ETCO_2$ levels of 26.01±6.85 mmHg in those without active bleeding versus 23.34±5.40 mmHg in those with active bleeding, con-

firming $ETCO_2$'s predictive value in identifying bleeding events. $ETCO_2$ measurement reflects CO_2 production in tissues transported to the lungs via blood; decreased blood volume in GI bleeding impairs oxygen delivery to tissues, reducing CO_2 production and excretion via respiration, thereby lowering $ETCO_2$ levels. Additionally, ventilation perfusion mismatch due to impaired lung perfusion further contributes to $ETCO_2$ reduction. Thus, $ETCO_2$ serves as an early indicator of hypovolemia in GI bleeding patients, detecting tissue-level hemorrhage effects. These findings underscore $ETCO_2$'s utility as a marker for assessing hypovolemia severity in GI bleeding.

The GBS is a scoring system used to assess the severity of bleeding in patients with gastrointestinal bleeding and to predict which patients can be safely discharged from the emergency department and which require emergency endoscopy. Gonçalves *et al*. [23]. reported that GBS can predict the need for blood transfusion in patients with GI bleeding. Chaudhary *et al*. [14], in a study evaluating endoscopy time in patients with GI bleeding, reported that patients with GBS ≥12 should be urgently taken to endoscopy and even severe bleeding should be suspected in these patients. In our study, we found that the $ETCO_2$ value obtained from patients with GBS ≥12 (23.58±5.97 mmHg) was statistically significantly lower than the $ETCO_2$ value obtained from patients with GBS <12 (28.44±5.95 mmHg). These findings indicate that low levels of $ETCO_2$ in patients with GBS ≥12 are a marker of severe bleeding. $ETCO_2$ level is a value that can be obtained quickly and noninvasively. However, GBS is a challenging condition to evaluate in the emergency department due to the time required to obtain laboratory results for the necessary parameters, the difficulty in distinguishing melena, especially in patients using iron medication, and the inability of patients and anxious relatives to fully describe the syncope condition. Therefore, we believe that in patients with GI bleeding, $ETCO_2$ levels provide more practical and rapid results compared to the GBS, allowing us to assess the presence and severity of bleeding in the emergency department.

Another scoring system considered in our study is the AIMS65 score [5]. A statistically significant difference in $ETCO_2$ values was observed between the groups of patients classified based on the AIMS65 score. When evaluated from this perspective, we see that the $ETCO_2$ value of patients with an AIMS65 score ≥2 is lower than that of patients with an AIMS65 <2. This score is a system used to predict the risk of mortality in patients with gastrointestinal bleeding. An AIMS65 score of ≥2 is associated with a high mortality risk [5, 15]. However, specific laboratory tests such as albumin, which are not routinely examined in emergency departments, are required for the calculation of the AIMS65 score. The low $ETCO_2$ value in patients with an AIMS65 score ≥2, which is considered to be at high risk for mortality, indicates that $ETCO_2$ can also be used to predict the risk of mortality. Furthermore, when assessed in relation to mortality, we found a statistically significant difference between the $ETCO_2$ values of deceased patients (20 mmHg) and surviving patients (26.10 mmHg). We can conclude that $ETCO_2$ is a good predictor of survival. Supporting our thesis, Campion *et al*. [24] reported that trauma patients in hemorrhagic shock with a pre-hospital measured $ETCO_2$ value of 18 mmHg resulted in patient mortality, while those with a value of 33 mmHg survived.

In our study, 51.5% of the 103 patients were male. Research has demonstrated that gastrointestinal bleeding is more prevalent in men than in women [25, 26]. The mean age of the patients in our study was 67.89±16.55 years, which is consistent with the average age reported in the literature [22, 25]. The patient population in our study aligns with findings in existing studies.

Limitations

This study was conducted as a single-center prospective observational study with a limited sample size, which may restrict the generalizability of the findings. Moreover, $ETCO_2$ measurements were obtained only at the time of admission, and changes over the course of follow-up were not evaluated. $ETCO_2$ levels may be influenced by various factors, particularly patient-related conditions such as altered mental status, anxiety in confined spaces, and poor compliance with the capnography device, all of which may affect the accuracy of the measurements. Furthermore, the 30-day mortality follow-up relied solely on hospital records, and deaths occurring after discharge may not have been captured.

CONCLUSION

This study shows that $ETCO_2$ levels in patients admitted to the emergency department for gastrointestinal bleeding are significantly lower in cases of active bleeding, increased mortality risk, GBS ≥12, and AIMS65 ≥2. These findings suggest that $ETCO_2$ could serve as a valuable marker for detecting hypovolemia and evaluating the patient's clinical condition. Additionally, compared to GBS and AIMS65 scores, $ETCO_2$ appears to be a faster, more practical, and applicable value in the emergency department. This result suggests that $ETCO_2$ may offer greater advantages in predicting the presence of bleeding, its severity, and the need for transfusion, compared to the GBS and AIMS65 scores. Our study has demonstrated the potential advantages of $ETCO_2$ in the evaluation of gastrointestinal tract bleeding, which is critical for rapid and effective intervention in the emergency department setting. Future studies involving larger patient populations and diverse clinical scenarios will provide a deeper understanding of how these findings can be applied in clinical practice.

Ethics Approval and Consent to Participate

This study was approved by the İzmir Kâtip Çelebi University Non-Interventional Clinical Research Ethics Committee (Decision no. 702, date: 12.05.2020). All procedures performed during data collection, review of patient records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Informed consent was waived because of the retrospective nature of the study and the analysis used anonymous clinical data.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: EK, RK, AK, FET; Study Design: EK, AK, EEG; Supervision: RK, AK, HA, FET; Funding: EK, AK, RK, FET; Materials: EK, EEG, AK, HA; Data Collection and/or Processing: EK, RK, EEG, HA; Statistical Analysis and/or Data Interpretation: EK, RK, AK, FET; Literature Review: EK, AK, EEG, HA; Manuscript Preparation: EK, RK, EEG and Critical Review: RK, AK, EEG, FET.

Conflict of interest

he authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgements

This article is derived from the specialization thesis. The authors are grateful to everyone who contributed to this study and provided moral support.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Hearnshaw SA, Logan RF, Lowe D, Travis SP, Murphy MF, Palmer KR. Acute upper gastrointestinal bleeding in the UK: patient characteristics, diagnoses and outcomes in the 2007 UK audit. Gut. 2011;60(10):1327-1335. doi: 10.1136/gut.2010.228437.
2. van Leerdam ME, Vreeburg EM, Rauws EA, et al. Acute upper GI bleeding: did anything change? Time trend analysis of incidence and outcome of acute upper GI bleeding between 1993/1994 and 2000. Am J Gastroenterol. 2003;98(7):1494-1499. doi: 10.1111/j.1572-0241.2003.07517.x.
3. Sugawa C, Steffes CP, Nakamura R, et al. Upper GI bleeding in an urban hospital. Etiology, recurrence, and prognosis. Ann Surg. 1990;212(4):521-526; discussion 526-527. doi:

10.1097/00000658-199010000-00014.
4. Blatchford O, Murray WR, Blatchford M. A risk score to predict need for treatment for uppergastrointestinal haemorrhage. Lancet. 2000;356(9238):1318-1321. doi: 10.1016/S0140-6736(00)02816-6.
5. Saltzman JR, Tabak YP, Hyett BH, Sun X, Travis AC, Johannes RS. A simple risk score accurately predicts in-hospital mortality, length of stay, and cost in acute upper GI bleeding. Gastrointest Endosc. 2011;74(6):1215-1224. doi: 10.1016/j.gie.2011.06.024.
6. Whitaker DK. Time for capnography - everywhere. Anaesthesia. 2011;66(7):544-549. doi: 10.1111/j.1365-2044.2011.06793.x.
7. Kodali BS. Capnography outside the operating rooms. Anesthesiology. 2013;118(1):192-201. doi: 10.1097/ALN.0b013e318278c8b6.
8. Nassar BS, Schmidt GA. Capnography During Critical Illness. Chest. 2016;149(2):576-585. doi: 10.1378/chest.15-1369.
9. Thompson JE, Jaffe MB. Capnographic waveforms in the mechanically ventilated patient. Respir Care. 2005;50(1):100-8; discussion 108-109.
10. Blanch L, Romero PV, Lucangelo U. Volumetric capnography in the mechanically ventilated patient. Minerva Anestesiol. 2006;72(6):577-585.
11. Zwerneman K. End-tidal carbon dioxide monitoring: a VITAL sign worth watching. Crit Care Nurs Clin North Am. 2006;18(2):217-25, xi. doi: 10.1016/j.ccell.2006.02.002.
12. Manifold CA, Davids N, Villers LC, Wampler DA. Capnography for the nonintubated patient in the emergency setting. J Emerg Med. 2013;45(4):626-32. doi: 10.1016/j.jemermed.2013.05.012.
13. Ward KR, Yealy DM. End-tidal carbon dioxide monitoring in emergency medicine, Part 1: Basic principles. Acad Emerg Med. 1998;5(6):628-36. doi: 10.1111/j.1553-2712.1998.tb02473.x.
14. Chaudhary S, Stanley AJ. Optimal timing of endoscopy in patients with acute upper gastrointestinal bleeding. Best Pract Res Clin Gastroenterol. 2019;42-43:101618. doi: 10.1016/j.bpg.2019.05.005.
15. Stanley AJ, Laine L, Dalton HR, et al; International Gastrointestinal Bleeding Consortium. Comparison of risk scoring systems for patients presenting with upper gastrointestinal bleeding: international multicentre prospective study. BMJ. 2017;356:i6432. doi: 10.1136/bmj.i6432.
16. Ladde JG, Miller S, Chin K, et al. End-tidal carbon dioxide measured at emergency department triage outperforms standard triage vital signs in predicting in-hospital mortality and intensive care unit admission. Acad Emerg Med. 2023;30(8):832-841. doi: 10.1111/acem.14703.
17. Guzman JA, Lacoma FJ, Najar A, Kruse JA. End-tidal partial pressure of carbon dioxide as a noninvasive indicator of systemic oxygen supply dependency during hemorrhagic shock and resuscitation. Shock. 1997;8(6):427-31.
18. Okamoto H, Hoka S, Kawasaki T, Okuyama T, Takahashi S. Changes in end-tidal carbon dioxide tension following sodium bicarbonate administration: correlation with cardiac output and haemoglobin concentration. Acta Anaesthesiol Scand. 1995;39(1):79-84. doi: 10.1111/j.1399-6576.1995.tb05596.x.
19. Belenkiy SM, Berry JS, Batchinsky AI, et al. The noninvasive carbon dioxide gradient (NICO2G) during hemorrhagic shock. Shock. 2014;42(1):38-43. doi: 10.1097/SHK.0000000000000177.
20. Bulger N, Harrington B, Krieger J, et al. Prehospital end-tidal carbon dioxide predicts hemorrhagic shock upon emergency department arrival. J Trauma Acute Care Surg. 2021;91(3):457-464. doi: 10.1097/TA.0000000000003312.
21. Wilson BR, Bruno J, Duckwitz M, Akers N, Jeanmonod D, Jeanmonod R. Prehospital end-tidal CO2 as an early marker for transfusion requirement in trauma patients. Am J Emerg Med. 2021;45:254-257. doi: 10.1016/j.ajem.2020.08.056.
22. O'Connor R, Kazemi A, Venkat A, Asaly M, Aoun E. Sa1468 Differentiation of Patients With Acute Gastrointestinal Bleeding From Those With Occult Bleeding by End-Tidal Carbon Dioxide Levels: a Pilot Study. Gastrointest Endosc. 2015;81(5):AB227. doi 10.1016/j.gie.2015.03.1305.
23. Cúrdia Gonçalves T, Barbosa M, Xavier S, et al. Optimizing the Risk Assessment in Upper Gastrointestinal Bleeding: Comparison of 5 Scores Predicting 7 Outcomes. GE Port J Gastroenterol. 2018;25(6):299-307. doi: 10.1159/000486802.
24. Campion EM, Cralley A, Robinson C, et al. Prehospital end-tidal carbon dioxide predicts massive transfusion and death following trauma. J Trauma Acute Care Surg. 2020;89(4):703-707. doi: 10.1097/TA.0000000000002846.
25. Diamantopoulou G, Konstantakis C, Kottorou A, et al. Acute Lower Gastrointestinal Bleeding: Characteristics and Clinical Outcome of Patients Treated With an Intensive Protocol. Gastroenterology Res. 2017;10(6):352-358. doi: 10.14740/gr914w.
26. Lau JYW, Yu Y, Tang RSY, et al. Timing of Endoscopy for Acute Upper Gastrointestinal Bleeding. N Engl J Med. 2020;382(14):1299-1308. doi: 10.1056/NEJMoa1912484.

The European Research Journal 2025:11(6):1147-1156 Original Article

DOI: https://doi.org/10.18621/eurj.1751730 *Pathology*

Investigation of the protective role of fisetin against doxorubicin-induced liver injury in rats

Ömür Gülsüm Deniz

Department of Histology and Embryology, Faculty of Medicine, Bolu Abant İzzet Baysal University, Bolu, Türkiye

ABSTRACT

Objectives: The aim of the present research was to histopathologically investigate the potentially advantageous effects of the flavonoid fisetin on liver damage induced by the chemotherapeutic drug doxorubicin (DOX) in rats.

Methods: Thirty-five Wistar albino female rats were randomized in five as Control, dimethyl sulfoxide (DMSO, solvent), fisetin (50 mg/kg/day; 7 days i.p.), DOX (single dose, 10 mg/kg; i.p.) and DOX+fisetin (10 mg/kg DOX+50 mg/kg/day fisetin for 7 days). Livers were harvested, fixed in 10% formalin, and processed for histopathology by hematoxylin-eosin staining. Central to these analyses of damage parameters for inflammation, sinusoidal dilatation, and hepatocyte injury were examined histopathologically.

Results: The DOX group had severe hepatocyte degeneration, inflammation, and sinusoidal dilatation. In contrast, the DOX+fisetin group expressed mild sinusoidal dilatation and insignificant inflammatory change. In this context, statistical significance was found between the DOX group and the DOX+fisetin group in terms of hepatocyte degeneration and inflammation ($P<0.01$), and sinusoidal dilatation ($P=0.038$). However, no significant differences were observed in between the Control, fisetin, and DMSO groups ($P=1.000$).

Conclusions: Fisetin conferred substantial histological protection from DOX-induced liver injury in rats. The amelioration may have resulted from the fisetin's antioxidant, anti-inflammatory, and perhaps membrane-stabilizing effects, thus proving its potential as a hepatoprotective agent.

Keywords: Doxorubicin, fisetin, histopathology, liver, rat

Doxorubicin (DOX), one of the first two anthracyclines to be recovered from Streptomyces peucetius, was first discovered during the 1960s [1]. It is a typical chemotherapeutic antibiotic that is effective in the therapy of various cancers, for example, breast cancer, leukemia, bronchogenic carcinoma, stomach carcinoma, sarcomas, thyroid carcinomas, and hematologic malignancies [2]. Although DOX has powerful antitumor activities, it is significant toxicity in non-malignant cells and therefore restricting its clinical utility [3]. The drug is metabolized by hepatic microsomal enzymes and cytoplasmic reductases, leading to the accumulation of toxic and immunogenic intermediates that contribute to liver injury [4]. Emerging research demonstrates that DOX-caused toxicity is due to several mechanisms that are largely diverse according to the cell type. In tumor cells, DOX mainly triggers DNA damage that leads to the suppres-

Received: July 26, 2025 **Accepted:** August 29, 2025 **Available Online:** October 8, 2025 **Published:** November 4, 2025

How to cite this article: Deniz ÖG. Investigation of the protective role of fisetin against doxorubicin-induced liver injury in rats. Eur Res J. 2025;11(6):1147-1156. doi: 10.18621/eurj.1751730

Corresponding author: Ömür Gülsüm Deniz, PhD., Assist. Prof., Phone: +90 374 254 10 00 ext. 6519, E-mail: omur.deniz@ibu.edu.tr, omur.denizomu@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

sion of proliferation [5], in non-malignant cells, it promotes oxidative stress and disrupts mitochondrial function through the accumulation of reactive oxygen species (ROS) [6]. Additionally, DOX has been shown to exacerbate oxidative stress by suppressing specific transcription factors [7]. This disruption of redox homeostasis by DOX has been documented in several organs, including the heart [8], kidneys [9], and liver [10, 11].

The liver bears the major role in detoxification through its ability to convert lipophiles into more aqueous compounds that can be easily removed from the body via urine [12]. Drug-induced hepatotoxicity or drug-induced liver injury (DILI) is an acute or chronic reaction to a natural or synthetic chemical [13]. Histopathological evaluation of liver tissue from animal models treated with DOX has revealed significant structural alterations. Studies have documented various hepatic abnormalities, including bile duct hyperplasia, sinusoidal dilation, and central vein congestion [14]. Other findings include hepatocyte vacuolation, nuclear condensation, degeneration of hepatic cords [15], cellular edema, focal necrosis, and disrupted hepatic trabecular organization [16]. Further observations encompass biliary duct proliferation, parenchymal necrosis [17], increased intercellular spaces, mitochondrial vacuolization and swelling [18], as well as lymphocyte infiltration [19]. These different findings reflect the extensive hepatic damage caused by DOX, with both degenerative and inflammatory liver histoarchitectural changes.

Fisetin (3,3′,4′,7-tetrahydroxyflavone) is a bioactive flavonoid found in numerous fruits and vegetables, such as strawberries, apples, grapes, onions, and cucumbers [20]. Fisetin demonstrates a wide range of pharmacological properties, including potent antioxidant [21], anti-inflammatory, anti-mitotic, and anticancer effects. Naturally found flavonoid fisetin has the potential to scavenge free radicals, control the signal pathways of cell proliferation [22], as well as suppress inflammatory mediators. Remarkably, the anticancer activity of fisetin was noticed to result through various mechanisms involving cell cycle arrest, apoptosis induction [23], as well as inhibition of the signal transduction pathway involving metastasis the spread of cancer cells from the place where they first formed to another part of the body [24, 25]. Fisetin demonstrates a favorable safety profile based on preclinical toxicity assessments. Acute toxicity studies employing animal models reported no adverse effects, with no observed changes in body weight or clinical signs of toxicity such as gastrointestinal disturbances (e.g., diarrhea), neurological symptoms (e.g., restlessness or coma), respiratory dysfunction, as well as motor abnormality. These findings suggest a wide therapeutic window for fisetin administration [26]. Histopathological studies revealed no detrimental effects in the heart, lungs, kidneys, liver, stomach, intestines, spleen, or reproductive organs [27-30].
DOX, although highly effective in fighting cancer, is a major cause of liver damage that is produced because of oxidative stress and inflammation. While a few protective drug studies have undertaken work on finding ways to restore tissue damage following the use of drug therapy. Despite rising interest in the therapeutic prospects of natural flavonoids, relatively little investigation has specifically addressed the histological effect of fisetin when used in the context of DOX elicited liver injury. Most current studies concentrate on the biochemical and antioxidant activities of the fisetin, with insufficient attention to its ability to preserve tissue arrangement at the tissue level. Also, DOX is widely known to generate intense hepatotoxicity through mechanisms involving oxidative stress, inflammation, and mitochondrial injury. However, the histological impact of potential protective compounds like fisetin remains a largely untapped area of research. Bridging this gap holds significant promise for understanding the structural and functional recovery of hepatic tissue following chemotherapeutic therapy.

While there is a growing recognition of natural flavonoids as powerful therapeutic agents, the research on the role of fisetin as the main agent in the restoration of liver tissue structure after the injury caused by DOX is poorly defined. Different scientific papers emphasize fisetin's antioxidant and anti-inflammatory effects, but few address its capacity to preserve histological integrity and reserve liver tissue architecture intactness after DOX-induced stress. The study aims to fill that gap by investigating the potential of fisetin to maintain the integrity of liver tissue structure in rats that have been given DOX. Thus, the current study aims to investigate the histologic protective activity of fisetin against DOX-induced liver injury in rats. In this context, it was concentrated on histopathological changes to reveal more about the role of fisetin

in mitigating liver damage caused by chemotherapy, which is an important part of the process of enhancing the patient's condition.

METHODS

Experimental Design

All experimental procedures carried out in our study were approved by the Bolu Abant İzzet Baysal University (BAİBÜ) Animal Experiments Local Ethics Committee with the decision numbered 2025/24 and dated 16/04/2025. Following the ethics committee approval, 35 adults female *Wistar albino* rats weighing 200-250 grams were obtained from the BAİBÜ Experimental Animals Application and Research Center. The animals were maintained in an environment with a stable temperature (24±2°C) and humidity (55±15%) at the BAIBU Experimental Animal Application and Research Center until the commencement of the investigation and throughout its duration. The rats were granted unrestricted access to regular pellet food and tap water prior to the experiment. In our study, 35 female rats were used. The subjects were randomly divided into 5 groups, and the experimental procedure was started. The groups and the procedures to be applied are given below in order:

-Control group: 1 ml/kg saline was given intraperitoneally (i.p.) for 7 days.

-Fisetin group: 50 mg/kg fisetin dissolved in 0.1% Dimethyl sulfoxide (DMSO) [31] was injected i.p. for 7 days [32].

-DOX group: 10 mg/kg single dose of DOX [33] was given as i.p., and rats were sacrificed after 7 days.

-DOX+Fisetin group: 10 mg/kg was given as a single dose of DOX i.p. and 50 mg/kg fisetin was injected as i.p. for 7 days.

-DMSO (solvent) group): 1 ml/kg 0.1% DMSO was administered i.p. for 7 days.

Histopathological Evaluation

At the end of the experimental period, the rats were anesthetized by injecting a 5/1 mixture of ketamine (50 mg/kg)/xylazine (10 mg/kg) i.p. and their liver tissues were removed. The tissues were fixed in 10% buffered neutral formalin solution for two weeks for histopathological evaluation and then embedded in paraffin blocks with routine histological follow-up (consisting of alcohol, xylene and paraffin series). From the obtained paraffin blocks, 4 µm thick sections were taken using a systematic random sampling method for histological procedures using a Thermo Shandon Finesse ME microtome device (Thermo Fisher Scientific, Cheshire, UK) and stained with Hematoxylin-eosin staining. Histopathological evaluation was performed using a Carl Zeiss Axiocam ERc5 model digital camera-attached microscope from the sections, and 5 different areas representing the relevant area were determined for each section taken from each animal by systematic random sampling in the relevant liver tissue, and parameters such as inflammation, hepatocyte damage, and sinusoidal dilatation were taken into consideration and evaluated under a light microscope. These morphological parameters were scored semi quantitatively between 0 and 3. If there was no change, the score was 0, if there was mild damage, the score was 1, if there was moderate damage, the score was 2, and if there was severe damage, the score was 3 [34]. All sampled were coded for the sake of all analyses would be done blindly.

Statistical Analysis

SPSS version 21.0 analysis program was used in the statistical analysis of the numerical data belonging to the groups obtained from our study. The conformity of the data to the normal distribution assumption was evaluated with the Shapiro-Wilk test. In comparing continuous variables specified by measurement, data conforming to normal distribution were evaluated using One-Way ANOVA and Bonferroni tests as post-hoc tests. In the statistical evaluations, the difference was accepted to be statistically significant when $P<0.05$.

RESULTS

In the study, histopathological analysis was performed on sections stained with hematoxylin-eosin. When the groups were evaluated, it was seen that there was no histological difference between the control, fisetin and DMSO groups. In the relevant groups, the liver tissue contained euchromatic (Normal, uncondensed chromatin pattern), centrally located, and polygonal hepatocytes with regular cell borders around the central vein located in the center of the lobule, which had a

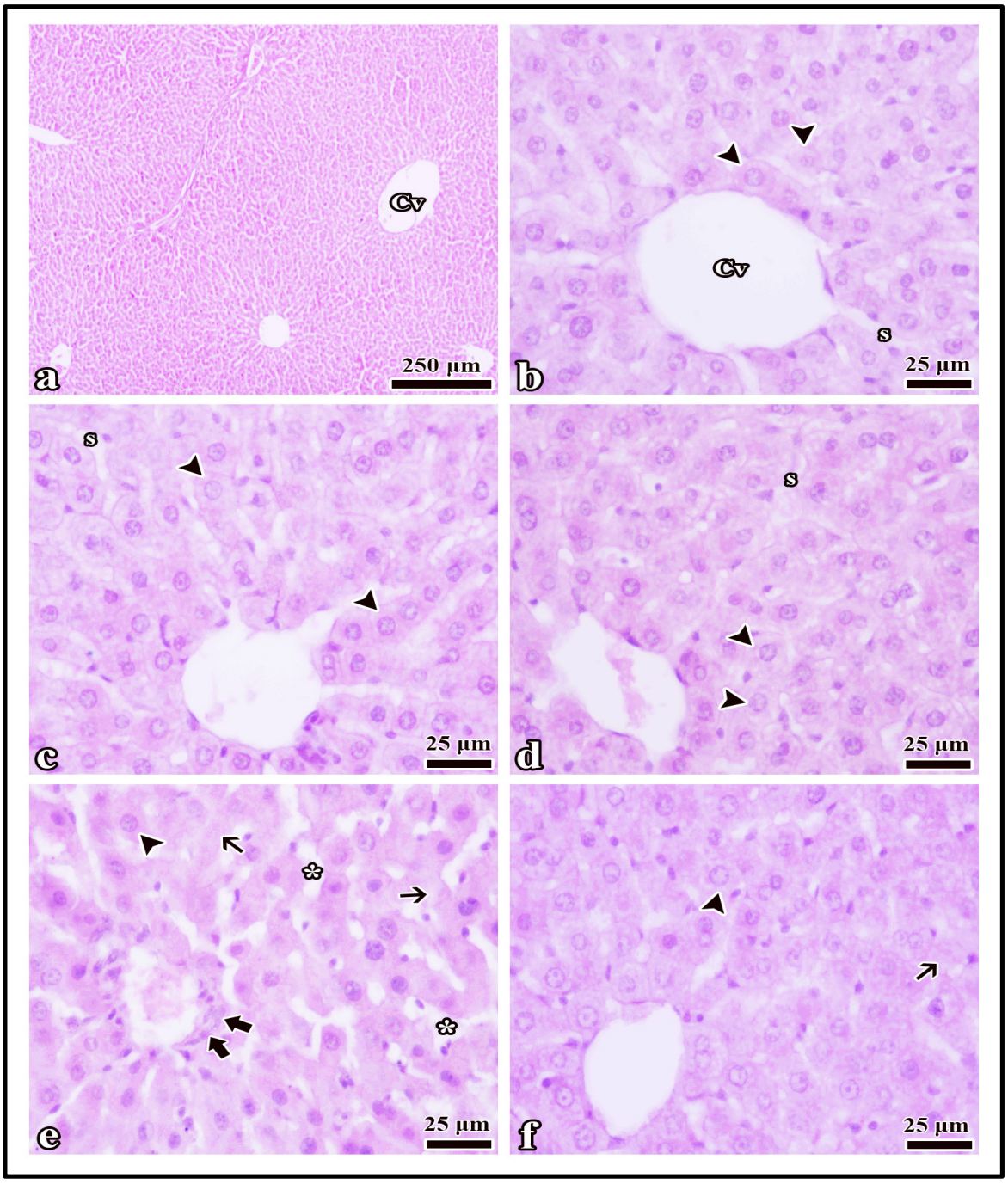

Fig. 1. The micrograph shows histological images of all groups. (a) Panaromic view of the liver. (b) Control group, (c) Fisetin group, (d) DMSO group, (e) DOX group, (f) DOX+fisetin group. Arrowhead; healthy hepatocyte, thin arrow; damaged hepatocyte, thick arrow; inflammation, (*); dilated sinusoid, Cv; central vein, s; sinusoid. (b), (c), and (d) were in normal morphology. (f) indicated that fisetin may minimize the detrimental effects of DOX (e). Hematoxylin-eosin staining ×40.

small amount of stroma showing normal morphology cell cords extending radially from there and sinusoids (Vascular channels between hepatocyte cords) between these cords. The portal area, which includes the hepatic artery, portal vein, and bile duct, called the portal triad at the corners of the hepatic lobule, appeared healthy. In the DOX group, sinusoidal dilatation, inflammation, and degeneration-related shape changes in hepatocytes in different parts of the parenchyma were remarkable. In addition, a morpho-

logical appearance similar to the control group was observed in the DOX+fisetin group, except for mild sinusoidal dilatation. No infiltration or congestion findings were observed in the relevant group (Fig. 1). When all these findings were evaluated in light of this; no statistically significant difference was observed between the control, fisetin and DMSO groups (P=1.000), while a statistical significance was found between these groups and the DOX and DOX+Fisetin groups (P<0.01). In addition, the significance between DOX and DOX+Fisetin groups was found to be at P<0.01(Hepatocyte damage, inflammation) (Figs. 2a and c, respectively) and P=0.038 (sinusoidal dilatation) (Fig. 2b) levels.

DISCUSSION

Liver is the main organ that detoxes the detoxification of chemicals, excretion of waste in bile and performing various vital functions like synthesis, storage and redistribution of carbohydrates, lipids and proteins [35]. Like many chemotherapeutic agents, DOX is metabolized in the liver and after that it is excreted in bile. In case the liver is not functioning properly, the elimination of DOX becomes slower and delayed, and it starts to accumulate in the body which in turn leads to adverse effects on the other tissues by increased systemic toxicity. DOX is extensively metabolized by the liver to its major metabolite, doxorubicinol, along with

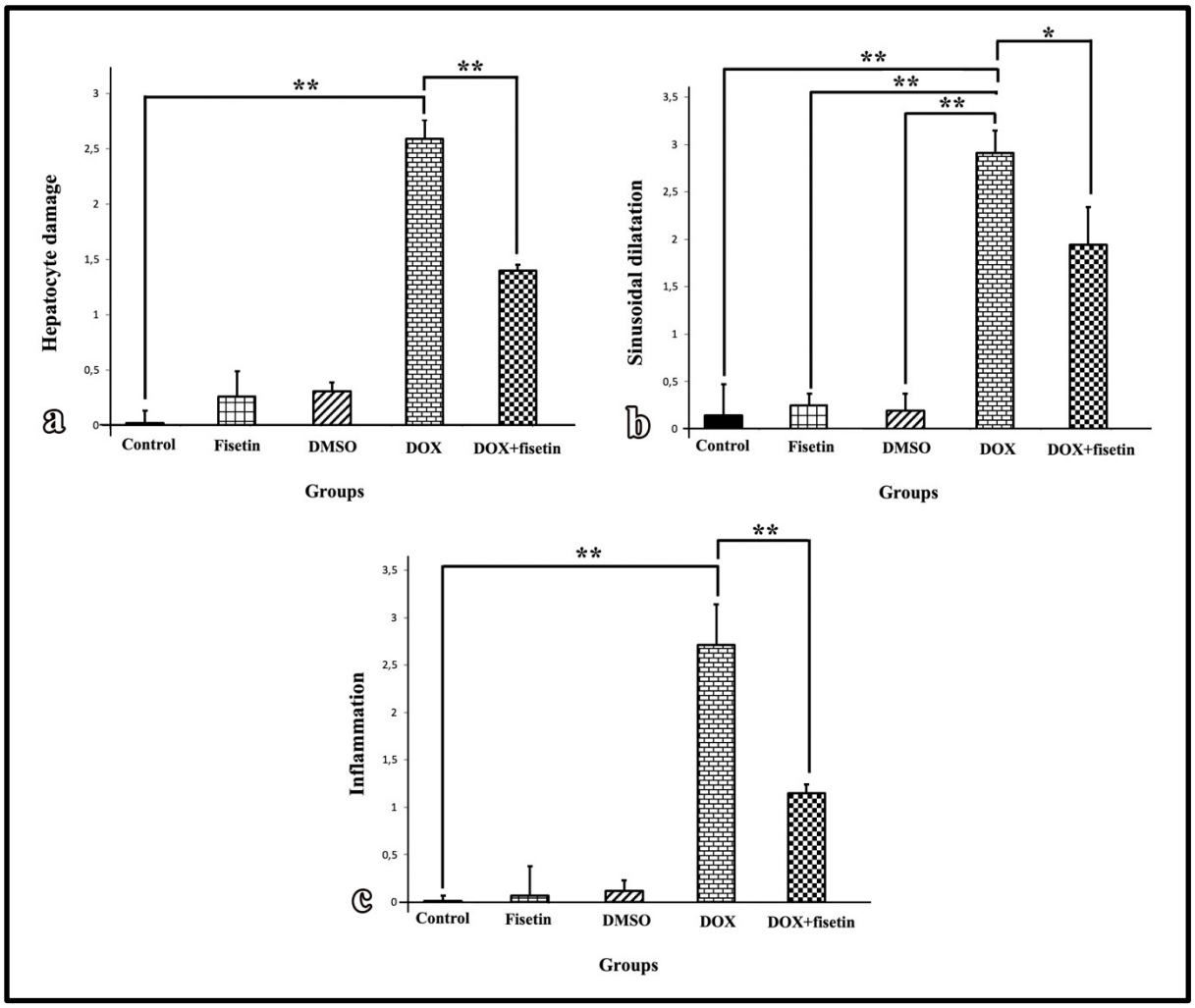

Fig. 2. Graph indicates the hepatocyte damage (a), sinusoidal dilatation (b) and inflammation (c) parameters belonging to all groups. Statistically significant differences at P<0.01 are shown by (**), while significant differences at P<0.05 are demonstrated by (*). The data obtained in the relevant graph show that DOX caused adverse effects on the liver, while fisetin minimized these effects.

several cytotoxic aglycone metabolites [36, 37]. For these drugs to be metabolized, they must first pass through the hepatic sinusoidal endothelium to enter the space of Disse and gain access to the hepatocytes [38]. Hence, understanding the mechanisms of hepatic drug clearance and how DOX induces liver toxicity liver toxicity is critical for developing new therapeutic drugs that can reduce these toxic effects. This underscores the urgent need for therapeutic agents that can mitigate such liver damage without compromising DOX's antitumor efficacy.

Multiple studies have documented the histological consequences of DOX-induced liver toxicity. They indicate that DOX cause expanded perisinusoidal spaces, edematous, containing material of damaged hepatocyte cytoplasm, and connective tissue proliferation and apoptosis [10, 11, 39]. Similarly, a study on male rats by El-Sayyad *et al*. [40], the most notable histological deficiencies in rats treated with doxorubicin at a dose of 1 mg/kg body weight on alternate days for 20 days included dissolution of hepatic cords appearing as hollow vacuoles organized by lines of necrotic hepatocytes. Focal inflammatory cells or necrotic tissues were observed in hepatic tissues. A higher tendency for liver fibrosis, characterized by multiple foci of focal cellular granulomatous lesions, was also a feature of DOX-treated rats. Our current findings corroborate these results in which DOX-induced hepatic toxicity was evidenced in the current histopathological investigation by the DOX group exhibited a liver that was severely damaged along with the signs of sinusoidal dilatation, hepatocyte arrangement that was disrupted, cellular degeneration, and inflammation. Another histopathology analysis of liver showed disintegration of histological architecture, there was excessive degeneration and dissolution of hepatic cords, focal inflammation and necrotic areas and increased apoptosis, vacuolations of hepatocytes with round to oval and elliptical shaped nuclei and sinusoidal dilatation, hemorrhage and congested portal vein and mononuclear cells infiltration in rats given intraperitoneally 10 mg/kg DOX [41]. Mansouri *et al*. [17] showed that the animals receiving DOX revealed mild to moderate pathological changes in the liver tissue such as degeneration of hepatocyte cords, dilatation of sinusoids, vacuolation of hepatocytes and condensation of nuclei which confirmed by our data.

Furthermore, DOX administration altered hepatic biochemical composition, resulted in a significant reduction in the levels of glycogen and nucleic acids and a significant rise in the quantity of lipids and proteins in the liver [42]. Other models, such as DOX-induced nonalcoholic steatohepatitis, revealed similar patterns of periportal necrosis and fatty degeneration, reinforcing the hepatotoxicity of DOX [43].

In the present study, histopathological evaluation of the hematoxylin-eosin-stained liver sections observed that there were no statistically significant differences among the control, fisetin, and DMSO groups ($P>0.05$). Related groups were all free of structural liver anomalies and they had normal liver histology that comprised of euchromatic hepatocytes with a central nucleus, sinusoidal structures without any damage, and healthy portal triads. Significant differences were observed between these groups and both the DOX and DOX+fisetin groups ($P<0.01$). Fisetin demonstrates notable hepatoprotective properties by preserving liver tissue structure and supporting normal hepatic function. Its antioxidant and anti-inflammatory actions contribute to reducing cellular stress and preventing liver damage under various pathological conditions.

Fisetin has been shown to exert multiple beneficial effects on liver metabolism and structure. One study reported that fisetin disrupts hepatic mitochondrial energy metabolism by inhibiting oxidative phosphorylation, reducing the mitochondrial NADH/NAD^+ redox ratio, suppressing ketogenesis, and enhancing oleate oxidation indicating a shift toward a more oxidized mitochondrial state and a predominantly prooxidant effect on liver mitochondria [44]. Additional studies have demonstrated its protective role in alcoholic liver disease by accelerating ethanol clearance and reducing oxidative stress [45]. Furthermore, dietary fisetin has been found to prevent hepatic steatosis through the modulation of lipid metabolism genes and miR-378 expression in mice [46]. A review showed that fisetin's had protective mechanisms in organ systems treated with chemotherapy, including modulation of oxidative stress and anti-inflammatory effects. While liver-specific data is limited, it validates the fisetin's ability to reduce DOX-induced toxicity [47]. These findings support the potential of fisetin as a multifunctional compound with both regulatory and protective roles in liver health and consistent with our findings that

show that fisetin alone group rats exhibited no structural liver anomalies, maintaining normal hepatic architecture with euchromatic hepatocytes featuring central nuclei, intact sinusoidal structures, and healthy portal triads, thereby reinforcing fisetin's hepatoprotective potential under toxic conditions.

Our study found that coadministration of fisetin with DOX significantly improved liver histology compared to the DOX-only group of ($P<0.05$). Liver architecture improved significantly in the DOX+fisetin group. This preserved lobular structure, sinusoidal dilatation and inflammation were reduced. Liver architecture improved significantly compared to the control group, suggesting significant hepatoprotection with fisetin. This hepatoprotective efficacy is likely due to fisetin's well-documented antioxidant and membrane-stabilizing properties that counteract DOX-induced oxidative injury. In methotrexate (MTX)-induced hepatotoxicity, fisetin administration effectively reversed all these changes, which points toward its hepatoprotective potential. Histopathology also indicated that fisetin diminished MTX-induced hepatocyte necrosis, fibrosis, and inflammatory infiltration. By inhibiting oxidative stress, inflammation, and apoptotic pathways, fisetin can emerge as a promising drug that leads to counteract MTX-induced hepatotoxicity [31]. Previous studies warrant these findings, providing evidence that administration of fisetin significantly reduces hepatic inflammation, preserves hepatocyte morphology, and stabilizes cell membranes in chemical-induced liver injury models [32, 36, 48]. These findings establish that fisetin enhances MTX-induced liver injury in rats through multimodal modulation of biochemical and histological damages markers [31]. Similarly, in thioacetamide-induced hepatic fibrosis, fisetin reduced the progress of histologic hepatic fibroplasia and diminished hepatic expression of α-smooth muscle actin (α-SMA) and cyclin D1 [49]. These consistent findings across various hepatotoxicity models reinforce fisetin's protective role in maintaining hepatic tissue integrity.

Fisetin administration exhibited remarkable hepatoprotective advantages by considerably lowering immunohistochemical and biochemical changes well-documented. DOX-induced oxidative stress in the liver triggers molecular events that contribute to hepatotoxicity. Hepatic enzymes metabolize DOX, generating ROS and reactive metabolites [50]. These harmful byproducts disrupt critical cellular processes, including DNA replication, transcription, and protein synthesis. The resulting DNA damage and impaired gene expression compromise hepatocyte regeneration, reducing the liver's ability to repair itself and maintain normal function [8]. Fisetin scavenges ROS, inhibits lipid peroxidation, and modulates cellular signaling pathways involved in inflammation and apoptosis [36]. Khairani and Ilyas's [50] review highlights the protective role of flavonoids, including that of fisetin, inhibiting DOX-induced oxidative liver injury. Flavonoids possess hepatoprotection through inhibiting oxidative stress through free radical formation and lipid peroxidation prevention. Flavonoids also safeguard mitochondrial integrity and reduce hepatocyte apoptosis. According to such mechanisms, flavonoids are candidate therapeutic drugs that can inhibit DOX-induced liver injury. Fisetin has been shown in Li *et al*. [51] study alleviate liver damage, reduce cell apoptosis, and minimize oxidative stress caused by hepatic ischemia-reperfusion injury, primarily through activation of the Nrf2/HO-1 signaling pathway. Fisetin protects against acetaminophen (APAP)-induced liver damage in mice and attenuates APAP-induced cytotoxicity in L-02 cells by promoting autophagy and inhibiting inflammasome activation [52]. A study investigating the protective effects of fisetin against subacute and chronic co-exposure to arsenic and fluoride in mice revealed that oxidative stress and arsenic burden in essential organs, including the liver, kidneys, and heart, were dose-dependently reduced by treatment with fisetin. These findings led to the conclusion that antioxidant activity and arsenic burden reduction by fisetin are responsible for its protective effect [53].

Strengths and Limitations

The present study has several strengths, including the use of a randomized controlled animal model, blinded histopathological evaluation, and comprehensive semi-quantitative scoring, which enhance the validity of our findings. Furthermore, the originality of investigating fisetin in DOX-induced hepatotoxicity highlights its translational potential. Nevertheless, some limitations should be acknowledged. It primarily focuses on histopathological evaluation. For future research, we recommend integrating corresponding bio-

chemical or molecular markers (e.g., serum liver enzymes, oxidative stress biomarkers, and cytokine levels). Additionally, the duration of fisetin treatment, dose-response relationships, and long-term outcomes were not assessed.

CONCLUSION

Our findings show that fisetin preserved hepatic lobular architecture, reduced sinusoidal dilatation, and prevented inflammation and hepatocyte degeneration caused by DOX administration. These protective actions appear to be exerted via the potent antioxidant, anti-inflammatory, and membrane stabilizing activities of fisetin. These findings indicate the potential of fisetin as a promising adjunctive therapeutic agent in the treatment of chemotherapy-induced hepatotoxicity. Nevertheless, more research remains to be conducted to identify its molecular targets and its long-term safety profile in human models.

In conclusion, this study suggests that fisetin is a promising natural compound with the ability to preserve liver histoarchitecture and reduce DOX-induced hepatic injury. Given its favorable safety profile and multi-targeted mechanisms of action, fisetin may offer a novel adjunctive therapeutic strategy for preventing chemotherapy-associated liver toxicity. These promising results support further exploration of fisetin as a potential adjunctive therapeutic agent to minimize chemotherapy associated with liver injury.

Ethics Approval and Consent to Participate

This study was approved by the Bolu Abant İzzet Baysal University animal experiments local ethics committee under decision no 2025/24 and date 16/04/2025. All experimental procedures were conducted in compliance with the U.K. Animals (Scientific Procedures) Act, 1986, the EU Directive 2010/63/EU for animal experimentation, and the National Institutes of Health guidelines for the care and use of laboratory animals (NIH Publications No. 8023, revised 1978).

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: ÖGD; Study Design: ÖGD; Supervision: ÖGD; Funding: ÖGD; Materials: ÖGD; Data Collection and/or Processing: ÖGD; Statistical Analysis and/or Data Interpretation: ÖGD; Literature Review: ÖGD; Manuscript Preparation: ÖGD; and Critical Review: ÖGD.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Carvalho C, Santos RX, Cardoso S, et al. Doxorubicin: the good, the gad and the ugly effect. Curr Med Chem. 2009;16(25):3267-3285. doi: 10.2174/092986709788803312.
2. Thorn CF, Oshiro C, Marsh S, et al. Doxorubicin pathways: Pharmacodynamics and adverse effects. Pharmacogenet Genomics. 2011;21(7):440-446. doi: 10.1097/FPC.0b013e32833ffb56.
3. Li S, Yuan S, Zhao Q, Wang B, Wang X, Li K. Quercetin enhances

chemotherapeutic effect of doxorubicin against human breast cancer cells while reducing toxic side effects of it. Biomed Pharmacother. 2018;100:441-447. doi: 10.1016/j.biopha.2018.02.055.
4. Licata S, Saponiero A, Mordente A, Minotti G. Doxorubicin metabolism and toxicity in human myocardium: Role of cytoplasmic deglycosidation and carbonyl reduction. Chem Res Toxicol. 2000;13(5):414-420. doi: 10.1021/tx000013q.
5. Edwardson D, Narendrula R, Chewchuk S, Mispel-Beyer K, Mapletoft J, Parissenti A. Role of drug metabolism in the cytotoxicity and clinical efficacy of anthracyclines. Curr Drug Metab. 2015;16(6):412-426. doi: 10.2174/1389200216888150915112039.
6. Asensio-López MC, Soler F, Pascual-Figal D, Fernández-Belda F, Lax A. Doxorubicin-induced oxidative stress: The protective effect of nicorandil on HL-1 cardiomyocytes. PLoS One. 2017;12(2):e0172803. doi: 10.1371/journal.pone.0172803.
7. Jakobs P, Serbulea V, Leitinger N, Eckers A, Haendeler J. Nuclear factor (Erythroid-Derived 2)-like 2 and thioredoxin-1 in atherosclerosis and ischemia/reperfusion injury in the heart. Antioxid Redox Signal. 2017;26(12):630-644. doi: 10.1089/ars.2016.6795.
8. Cappetta D, De Angelis A, Sapio L, et al. Oxidative stress and cellular response to doxorubicin: A common factor in the complex milieu of anthracycline cardiotoxicity. Oxid Med Cell Longev. 2017;2017:1521020. doi: 10.1155/2017/1521020.
9. Elsherbiny NM, El-Sherbiny M. Thymoquinone attenuates doxorubicin-induced nephrotoxicity in rats: Role of Nrf2 and NOX4. Chem Biol Interact. 2014;223:102-108. doi: 10.1016/j.cbi.2014.09.015.
10. Barakat BM, Ahmed HI, Bahr HI, Elbahaie AM. Protective effect of boswellic acids against doxorubicin-induced hepatotoxicity: Impact on Nrf2/HO-1 defense pathway. Oxid Med Cell Longev. 2018;2018:8296451. doi: 10.1155/2018/8296451.
11. Alasmari AF, Alharbi M, Alqahtani F, et al. Diosmin alleviates doxorubicin-induced liver injury via modulation of oxidative stress-mediated hepatic inflammation and apoptosis via NfkB and MAPK pathway: A preclinical study. Antioxidants (Basel). 2021;10(12):1998. doi: 10.3390/antiox10121998.
12. Grant DM. Detoxification pathways in the liver. J Inherit Metab Dis. 1991;14(4):421-430. doi: 10.1007/BF01797915.
13. Fisher K, Vuppalanchi R, Saxena R. Drug-induced liver injury. Arch Pathol Lab Med. 2015;139(7):876-887. doi: 10.5858/arpa.2014-0214-RA.
14. Shivakumar P, Usha Rani M, Gopala Reddy A, Anjaneyulu Y. A study on the toxic effects of doxorubicin on the histology of certain organs. Toxicol Int. 2012;19(3):241-244. doi: 10.4103/0971-6580.103656.
15. Mansouri E, Jangaran A, Ashtari A. Protective effect of pravastatin on doxorubicin-induced hepatotoxicity. Bratisl Lek Listy. 2017;118(5):273-277. doi: 10.4149/BLL_2017_054.
16. Martinel Lamas DJ, Nicoud MB, Sterle HA, et al. Selective cytoprotective effect of histamine on doxorubicin-induced hepatic and cardiac toxicity in animal models. Cell Death Discov. 2015;1:15059. doi: 10.1038/cddiscovery.2015.59.
17. Mete R, Oran M, Topcu B, et al. Protective effects of onion (Allium cepa) extract against doxorubicin-induced hepatotoxicity in rats. Toxicol Ind Health. 2016;32(3):551-557. doi: 10.1177/0748233713504807.
18. Kalender Y, Yel M, Kalender S. Doxorubicin hepatotoxicity and hepatic free radical metabolism in rats: The effects of vitamin E and catechin. Toxicology. 2005;209(1):39-45. doi: 10.1016/j.tox.2004.12.003.
19. Yao PA, Wu R, Zhang YL, et al. Alleviation of doxorubicin-induced hepatic toxicity with fermented Cordyceps sinensis via regulating hepatic energy metabolism in rats. Pharmacogn Mag. 2018;14(56):283-289. doi: 10.4103/pm.pm_354_17.
20. Li D, Liu X, Pi W, et al. Fisetin attenuates doxorubicin-induced cardiomyopathy in vivo and in vitro by inhibiting ferroptosis through SIRT1/Nrf2 signaling pathway activation. Front Pharmacol. 2022;12:808480. doi: 10.3389/fphar.2021.808480.
21. Khan N, Syed DN, Ahmad N, Mukhtar H. Fisetin: A dietary antioxidant for health promotion. Antioxid Redox Signal. 2013;19(2):151-162. doi: 10.1089/ars.2012.4901.
22. Afzal M, Alarifi A, Karami AM, et al. Antiproliferative mechanisms of a polyphenolic combination of kaempferol and fisetin in triple-negative breast cancer cells. Int J Mol Sci. 2023;24(7):6393. doi: 10.3390/ijms24076393.
23. Sundarraj K, Raghunath A, Panneerselvam L, Perumal E. Fisetin, a phytopolyphenol, targets apoptotic and necroptotic cell death in HepG2 cells. Biofactors. 2020;46(1):118-135. doi: 10.1002/biof.1577.
24. Ma T, Kandhare AD, Mukherjee-Kandhare AA, Bodhankar SL. Fisetin, a plant flavonoid ameliorates doxorubicin-induced cardiotoxicity in experimental rats: the decisive role of caspase-3, COX-II, cTn-I, iNOs and TNF-α. Mol Biol Rep. 2019;46(1):105-118. doi: 10.1007/s11033-018-4450-y.
25. Lin KH, Ramesh S, Agarwal S, et al. Fisetin attenuates doxorubicin-induced cardiotoxicity by inhibiting the insulin-like growth factor II receptor apoptotic pathway through estrogen receptor-α/-β activation. Phytother Res. 2023;37(9):3964-3981. doi: 10.1002/ptr.7855.
26. Prasath GS, Subramanian SP. Modulatory effects of fisetin, a bioflavonoid, on hyperglycemia by attenuating the key enzymes of carbohydrate metabolism in hepatic and renal tissues in streptozotocin-induced diabetic rats. Eur J Pharmacol. 2011;668(3):492-496. doi: 10.1016/j.ejphar.2011.07.021.
27. Naeimi AF, Alizadeh M. Antioxidant properties of the flavonoid fisetin: An updated review of in vivo and in vitro studies. Trends Food Sci Technol. 2017;70:34-44. doi: 10.1016/j.tifs.2017.10.003.
28. Althunibat OY, Al Hroob AM, Abukhalil MH, et al. Fisetin ameliorates oxidative stress, inflammation and apoptosis in diabetic cardiomyopathy. Life Sci. 2019;221:83-92. doi: 10.1016/j.lfs.2019.02.017.
29. Li D, Liu X, Pi W, et al. Fisetin attenuates doxorubicin-induced cardiomyopathy in vivo and in vitro by inhibiting ferroptosis through SIRT1/Nrf2 signaling pathway activation. Front Pharmacol. 2022;12:808480. doi: 10.3389/fphar.2021.808480.
30. El Shaer DF, Abd El Halim HI. The possible ameliorating role of fisetin on hepatic changes induced by fluoxetine in adult male albino rats: histological, immunohistochemical, and biochemical study. J Microsc Ultrastruct. 2023;11(3):161-171. doi: 10.4103/jmau.jmau_84_22.
31. Kayali A, Bora ES, Acar H, Yilmaz G, Erbaş O. Fisetin ameliorates methotrexate induced liver fibrosis. Eur Rev Med Pharmacol

Sci. 2024;28(8):3112-3119. doi: 10.26355/eurrev_202404_36027.
32. Hawi AJ, Al-Shawi NN. Evaluation of the effect of fisetin against cyclophosphamide-induced myelosuppression and oxidative stress in male albino rats. Iraqi J Pharm Sci. 2023;32(2):120-127. doi: 10.31351/vol32iss2pp120-127.
33. Cecen E, Dost T, Culhaci N, Karul A, Ergur B, Birincioglu M. Protective effects of silymarin against doxorubicin-induced toxicity. Asian Pac J Cancer Prev. 2011;12(10):2697-2704.
34. Elma B, Mammadov R, Süleyman H, et al. The effect of rutin on experimentally induced acute heart contusion in rats: Biochemical and histopathological evaluation. Ulus Travma Acil Cerrahi Derg. 2022;28(8):1073-1081. doi: 10.14744/tjtes.2021.97760.
35. Bechmann LP, Hannivoort RA, Gerken G, Hotamisligil GS, Trauner M, Canbay A. The interaction of hepatic lipid and glucose metabolism in liver diseases. J Hepatol. 2012;56(4):952-964. doi: 10.1016/j.jhep.2011.08.025.
36. Ballet F, Vrignaud P, Robert J, Rey C, Poupon R. Hepatic extraction, metabolism and biliary excretion of doxorubicin in the isolated prefused rat liver. Cancer Chemother Pharmacol. 1987;19(3):240-245. doi: 10.1007/BF00252979.
37. Bachur NR, Fox BM. Loss of fluorescence by anthracycline antibiotics: effects of xanthine oxidase and identification of the nonfluorescent metabolites. Cancer Res. 1987;47(4):1036-1039.
38. Le Couteur DG, Fraser R, Hilmer S, Rivory LP, McLean AJ. The hepatic sinusoid in aging and cirrhosis: Effects on hepatic substrate disposition and drug clearance. Clin Pharmacokinet. 2005;44(2):187-200. doi: 10.2165/00003088-200544020-00004.
39. Song S, Chu L, Liang H, et al. Protective effects of dioscin against doxorubicin-induced hepatotoxicity via regulation of SIRT1/FoxO1/NF-κB signal. Front Pharmacol. 2019;10:1030. doi: 10.3389/fphar.2019.01030.
40. El-Sayyad HI, Ismail MF, Shalaby FM, et al. Histopathological effects of cisplatin, doxorubicin and 5-flurouracil (5-FU) on the liver of male albino rats. Int J Biol Sci. 2009;5(5):466-473. doi: 10.7150/ijbs.5.466.
41. Chaudhary D, Khatiwada S, Sah SK, Tamang MK, Bhattacharya S, Jha CB. Effect of doxorubicin on histomorphology of liver of wistar albino rats. J Pharm Pharmacol. 2016;4(4):186–190. doi: 10.17265/2328-2150/2016.04.005.
42. Orak N, Çakmak Arslan G, Kaya ST. An Evaluation of damages caused by doxorubicin in liver tissue and potential protective effect of propolis on these damages. Konuralp Med J. 2022;14(1):104-113. doi: 10.18521/ktd.1049919.
43. Maslova GS, Skrypnyk IM, Yeroshenko GA. Morphological features of doxorubicin-induced liver damage associated with nonalcoholic steatohepatitis. World Med Biol. 2020;16(73):189-194. doi: 10.26724/2079-8334-2020-3-73-189-194.
44. Constantin RP, Constantin J, Pagadigorria CLS, et al. Prooxidant activity of fisetin: Effects on energy metabolism in the rat liver. J Biochem Mol Toxicol. 2011;25(2):117-126. doi: 10.1002/jbt.20367.
45. Sun Q, Zhang W, Zhong W, Sun X, Zhou Z. Dietary fisetin supplementation protects against alcohol-induced liver injury in mice. Alcohol Clin Exp Res. 2016;40(10):2076-2084. doi: 10.1111/acer.13172.
46. Jeon T-Il, Park JW, Ahn J, Jung CH, Ha TY. Fisetin protects against hepatosteatosis in mice by inhibiting miR-378. Mol Nutr Food Res. 2013;57(11):1931-1937. doi: 10.1002/mnfr.201300071.
47. Rauf A, Abu-Izneid T, Imran M, et al. Therapeutic potential and molecular mechanisms of the multitargeted flavonoid fisetin. Curr Top Med Chem. 2023;23(21):2075-2096. doi: 10.2174/1568026623666230710162217.
48. Singh M, Kadhim MM, Jalil AT, et al. A systematic review of the protective effects of silymarin/silibinin against doxorubicin-induced cardiotoxicity. Cancer Cell Int. 2023;23(1):88. doi: 10.1186/s12935-023-02936-4.
49. El-Fadaly AA, Afifi NA, El-Eraky W, et al. Fisetin alleviates thioacetamide-induced hepatic fibrosis in rats by inhibiting Wnt/β-catenin signaling pathway. Immunopharmacol Immunotoxicol. 2022;44(3):355-366. doi: 10.1080/08923973.2022.2047198.
50. Khairani D, Ilyas S. Flavonoids for hepatoprotection: Potential natural approaches to mitigate doxorubicin-induced liver toxicity: A mini review. J Med Pharm Chem Res. 2025;7(7):1464-1485. doi: 10.48309/JMPCR.2025.471206.1367.
51. Li Z, Wang Y, Zhang Y, et al. Protective effects of fisetin on hepatic ischemia-reperfusion injury through alleviation of apoptosis and oxidative stress. Arch Med Res. 2021;52(2):163-173. doi: 10.1016/j.arcmed.2020.10.009.
52. Zhang J, Zhao L, Hu C, et al. Fisetin prevents acetaminophen-induced liver injury by promoting autophagy. Front Pharmacol. 2020;11:162. doi: 10.3389/fphar.2020.00162.
53. Nayak S, Rakshit D, Yadav V, et al. Protective effect of fisetin on subacute and chronic arsenic and fluoride co-exposure induced hepatic, renal and cardiac toxicities in mice. J Biochem Mol Toxicol. 2025;39(2):e70113. doi: 10.1002/jbt.70113.

The European Research Journal 2025;11(6):1157-1166 Original Article

DOI: https://doi.org/10.18621/eurj.1777758 *Chest Diseases*

Investigation of the potential predictive value of flow-volume curve changes in patients with obstructive sleep apnea syndrome

Fatma Arzu Akkuş[1], Şebnem Yosunkaya[2]

[1]Department of Immunology and Allergy, Konya City Hospital, Konya, Türkiye; [2]Department of Internal Medicine, Division of Pulmonary Diseases, Necmettin Erbakan University, Faculty of Medicine, Konya, Türkiye

ABSTRACT

Objectives: The relationship between pulmonary function tests (PFT) performed while awake and sleep-related breathing disorders is not yet well defined. This study aimed to examine the potential of flow-volume curve abnormalities in assessing both the presence and severity of obstructive sleep apnea syndrome.

Methods: This retrospective study reviewed medical records of patients evaluated for suspected obstructive sleep apnea syndrome (OSAS) between May 2011 and April 2013. Among 141 patients with available PFT, 99 were diagnosed with OSAS, and 42 were normal based on polysomnography results. Patients were classified into OSAS and control groups, and pulmonary function parameters, including flow-volume curve features, were analyzed.

Results: Of the OSAS patients, 80 (80.8%) were male, while 22 (52.4%) males were present in the control group, which consisted of patients not diagnosed with obstructive sleep apnea syndrome. Among pulmonary function test parameters, the ratio of maximal mid-expiratory flow to forced vital capacity (FEF25-75/FVC) was lower in the obstructive sleep apnea syndrome group than in the control group ($P<0.05$). Patients with the sawtooth sign showed significantly higher apnea-hypopnea index, apnea-hypopnea index during rapid eye movement sleep, and total apnea scores ($P<0.05$). The presence of sawtooth signs in the flow-volume loop was useful for identifying more severe cases of obstructive sleep apnea syndrome.

Conclusions: The complex pathophysiology of obstructive sleep apnea syndrome complicates the identification of patients who need polysomnography. Flow-volume curve abnormalities, indicating airway instability in the upper respiratory tract, are common in these patients and may serve as early indicators of obstructive sleep apnea syndrome.

Keywords: Obstructive sleep apnea, sleep apnea, polysomnography, pulmonary function test, flow-volume curve, spirometry

Obstructive sleep apnea syndrome (OSAS), a type of sleep-related breathing disorder (SBD), is characterized by recurrent episodes of complete (apnea) or partial (hypopnea) obstruction of the upper respiratory tract (URT), resulting in sleep fragmentation, increased daytime sleepiness, and typically a reduction in blood oxygen saturation [1]. URT patency maintains the balance between the collapsing ef-

Received: September 4, 2025 **Accepted:** October 8, 2025 **Available Online:** October 14, 2025 **Published:** November 4, 2025

How to cite this article: Akkuş FA, Yosunkaya Ş. Investigation of the potential predictive value of flow-volume curve changes in patients with obstructive sleep apnea syndrome. Eur Res J. 2025;11(6):1157-1166. doi: 10.18621/eurj.1777758

Corresponding author: Fatma Arzu Akkuş, MD., Phone: +90 332 310 50 00, E-mail: arzusmr@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

fect of negative intraluminal pressure during inspiration and the activity of dilator muscles in the URT. This balance is highly complex, as it is influenced by several anatomical, mechanical, neuromuscular, and central factors. In OSAS patients, URTs differ significantly from those of healthy individuals, with more than 75% experiencing airway collapse in the retropalatal region. The obstruction is typically not localized to a single level [2].

Factors that reduce the width of the URT or facilitate obstruction increase the susceptibility to OSAS. Risk factors for OSAS include obesity, male gender, and increasing age [3]. Epidemiological studies have reported an OSAS incidence of 24% in men and 9% in women aged 30-60 years [4]. Male gender is an independent risk factor, as men tend to accumulate more fat in the abdomen and neck. Obesity is the most critical factor, with weight gain directly correlating to increased sleep apnea severity. Fat accumulation in the tongue narrows the URT, raising the risk of collapse [3]. Furthermore, SBD have been reported in several congenital and inherited diseases involving structural alterations in the URT and affecting the respiratory control center [4].

The gold-standard diagnostic method for SBD is polysomnography (PSG) [3]. This test, which is performed under the supervision of a trained technician, is both time-consuming and costly. PSG requires the patient to be hospitalized overnight. According to PSG results, stages of sleep should be examined in detail. Therefore, reliable screening tests are needed to identify patients who require PSG, particularly in cases with a high clinical suspicion index.

Spirometry, a non-invasive physiological test, is commonly used as a general respiratory screening tool [5]. Previous studies have reported flow-volume curve abnormalities that suggest airway instability in the URT of OSAS patients [6]. However, the predictive value of spirometric parameters for OSAS diagnosis remains unclear.

This study aimed to investigate whether flow-volume curve abnormalities and pulmonary function test (PFT) parameters differ between OSAS patients and controls, as well as their association with PSG results and clinical significance. In particular, this study provides a focused evaluation of sawtooth sign prevalence and the FEF50/FIF50 ratio - parameters that are not routinely assessed in OSAS patients - to explore their potential utility in clinical screening. In addition to conventional spirometric indices, previous studies have emphasized the potential value of specific parameters such as the FEF25–75/FVC ratio and FEF50/FIF50 ratio, and saw tooth sign in reflecting upper airway instability in OSAS patients [7-10]. Based on this literature, we hypothesized that these parameters, although not routinely assessed in OSAS evaluation, might show measurable differences between OSAS patients and controls and could correlate with PSG-derived severity markers.

METHODS

Study Design

This retrospective cross-sectional study was approved by the Ethics Committee (approval no: 2012-28) in accordance with the principles of the Declaration of Helsinki. The medical records of 196 patients who underwent PSG due to suspected OSAS between May 2011 and April 2013 were reviewed. Patients with a previous diagnosis of OSAS, obstructive lung disease, active smoking status, or incomplete medical records were excluded from the study. All patients had also undergone an ear, nose, and throat (ENT) examination as part of their clinical evaluation. Patients with significant upper airway abnormalities were excluded based on this assessment. Among the remaining patients, 141 individuals with available and complete PFT results were included in the final analysis. These inclusion and exclusion criteria were applied to minimize potential confounding factors affecting spirometry results and to ensure data integrit.

The diagnosis of OSAS was made by a pulmonologist specializing in sleep disorders based on the combination of clinical symptoms and PSG results, following AASM guidelines. Patients with an AHI >5 accompanied by relevant symptoms were diagnosed with OSAS, while those with an AHI ≤5 and no significant symptoms were classified as controls.

Patients' demographics, clinical histories, and laboratory results were obtained from the patient files and electronic medical records. Anthropometric measurements, including body mass index (BMI), neck circumference, and waist circumference, were recorded. BMI was calculated as weight (in kilograms) divided by height (in meters squared). Neck circumference

was measured at the widest part of the neck at the level of the cricoid cartilage.

Polysomnography

Our sleep laboratory used a digital polysomnographic system (VIASYS Healthcare GmbH, Leibnizstraβe 7, 97204 Hoechberg, Germany). Electroencephalography (EEG), electrooculography (EOG), and submental electromyography (EMG) recordings were obtained to evaluate sleep. Respiration was monitored using an oro-nasal flowmeter and a thermistor placed in the nose. Thoracic and abdominal movements were recorded with a thoracoabdominal motion sensor. Oxyhemoglobin saturation and heart rate were monitored using a pulse oximeter. Leg motions were recorded by an EMG sensor placed on the anterior tibial muscle of one leg. Scoring of sleep stages was performed manually according to the AASM scoring criteria [11, 12].

Obstructive apnea was defined as a reduction in airflow by ≥90% for a duration of ≥10 seconds with ongoing respiratory effort. Hypopnea, on the other hand, was defined as a reduction in airflow by ≥30% for a duration of ≥10 seconds, accompanied by an oxy-

Table 1. Demographic, clinical, and polysomnography features of the OSAS patients and the control group

Parameter	OSAS (n=99)	Control (n=42)	P value*
Demographic and clinical features			
Age (years)	48.41±10.37	44.33±12.66	**0.048**
Height (cm)	169.41±9.83	167.38±10.59	0.274
Weight (kg)	93.89±16.48	83.14±20.01	**0.001**
BMI (kg/m^2)	33.64±9.88	29.93±8.84	**0.038**
Neck circumference (cm)	44.02±13.34	38.78±3.17	**0.013**
Waist circumference (cm)	106.54±12.76	97.04±9.10	**<0.001**
Systolic blood pressure (mm Hg)	133.23±19.42	121.43±10.94	**<0.001**
Diastolic blood pressure (mm Hg)	73.74±12.74	64.29±11.07	**<0.001**
Hemoglobin (g/dl)	14.64±1.58	13.90±1.72	**0.014**
Hematocrit (%)	44.43±4.56	42.19±3.77	**0.006**
PSG results			
AHI total	36.60±24.97	2.24±1.55	**<0.001**
AHI side	26.20±25.29	2.83±5.94	**<0.001**
AHI supine	50.06±30.59	3.49±3.51	**<0.001**
Total apnea	26.40±23.38	0.81±0.97	**<0.001**
Total hypopnea	10.55±8.75	1.42±1.14	**<0.001**
Minimum SaO_2 (%)	73.74±12.10	87.85±3.30	**<0.001**
Mean SaO_2 (%)	88.91±5.08	93.42±1.98	**<0.001**
Duration of <90% SaO_2	40.91±35.69	5.68±13.02	**<0.001**
Duration of REM (%)	10.18±6.91	5.65±6.28	**<0.001**
Sleep efficiency	78.15±13.63	73.58±15.90	0.085

Data are shown as mean±standard deviation. OSAS=Obstructive sleep apnea syndrome, PSG=Polysomnography, AHI=Apnea hypopnea index, Mean SaO_2=Mean oxygen saturation, Min SaO_2=Minimum oxygen saturation, REM=Rapid eye movement.

*Mann-Whitney U test

gen desaturation of ≥3% or arousal.

The Apnea-Hypopnea Index (AHI) value was calculated by dividing the total number of apnea and hypopnea episodes by the total sleep duration (in hours). When the AHI was between 5 and 15, it was classified as mild OSAS, 16-30 as moderate OSAS, and >30 as severe OSAS [4]. The diagnosis of OSAS was established by considering both symptoms and polysomnographic results together.

From the full-night standard polysomnography (PSG) reports of the patients, the following parameters were recorded: total AHI, AHI during the rapid eye movement (REM) sleep stage (AHI REM), AHI in the supine position (AHI supine), total apneas, minimum oxygen saturation (min SaO_2), the lowest value of oxygen saturation during sleep, mean oxygen saturation (mean SaO_2) recorded night long, the mean of oxygen saturation during sleep, and the ratio of the duration with oxygen saturation below 90% to the duration of full-night sleep (ratio of duration of <90% SaO_2) were recorded. Additionally, for all patients, sleep efficiency (SE), defined as the ratio of total sleep duration to the total time spent in bed, and the ratio of the REM sleep stage duration to the total sleep duration (ratio of duration of REM) were also recorded.

Pulmonary Function Test

PFT was performed in the sitting position using a computerized system (Vmax 22, Sensormedics, USA). The patients exhaled into the mouthpiece at normal tidal volume for spirometric maneuvers, followed by deep inhalation and exhalation. The exhalation duration was at least 6 seconds without interruption. Pulmonary function tests were conducted in accordance with the American Thoracic Society/European Respiratory Society (ATS/ERS) guidelines [13]. A technician instructed each patient to perform at least three maneuvers, and the best values were recorded. PFT data were recorded in liters and as a percentage of the expected values based on age, gender, weight, and height.

From the PFT results of the patients, the following parameters were recorded: forced vital capacity (FVC), forced expiratory volume in the 1st second (FEV1), peak expiratory flow (PEF), FEV1/FVC ratio (Tiffeneau ratio), expiratory flow rates (FEF), maximal mid-expiratory flow rate (FEF %25-75), and maximal expiratory flow rates measured at 25% (FEF 25%), 50% (FEF 50%), and 75% (FEF 75%) of FVC exhaled. From the inspiratory curve, the following parameters were estimated and recorded: forced inspiratory volume at 1 second (FIV1), peak inspiratory flow (PIF), forced inspiratory flow at 50% (FIF 50%), and FEF50/FIF50 ratio. The FEF25-75/FVC ratio was derived by dividing the FEF25-75 value by FVC.

Maximal inspiratory and expiratory flow-volume curves were examined for the presence of a sawtooth sign. According to the definition by Sanders and colleagues, the sawtooth sign was determined by the presence of at least three consecutive, periodic fluctuations in the inspiratory and/or expiratory flow-volume curve [14].

Statistical Analysis

Continuous variables were expressed as mean±standard deviation (SD), whereas categorical variables were expressed as numbers and percentages. Normally distributed variables were compared using the Independent Samples t-test, while non-normally distributed variables were compared using the Mann-Whitney U test. Pearson's chi-squared test and Fisher's exact test were used to compare categorical variables. The relationships between PSG parameters and PFT values were analyzed using Spearman correlation analysis. Statistical analyses were performed using SPSS version 16.0 (SPSS Inc., Chicago, IL, USA). A P-value < 0.05 was considered statistically significant.

RESULTS

Demographic, Clinical, and Polysomnography Results

A total of 141 patients who underwent PSG due to suspected OSAS and whose PFT results were obtained were included in the study. Of these patients, 99 were diagnosed with OSAS, while 42 were considered normal. Among the OSAS patients, 80 (80.8%) were male, whereas 22 (52.4%) of the control group were male (P=0.001). OSAS patients had significantly higher age, weight, neck circumference, waist circumference, and BMI compared to the control group (Table 1). There were no significant differences between the groups in terms of the ratio of ex-smokers, smokers, and smoking-naive individuals.

PSG findings were within normal limits in the control group, while the OSAS group showed signif-

Table 2. PFT parameters of the OSAS patients and the control group

PFT parameters	OSAS (n=99)	Control (n=42)	P value
FEV1 (L)	3.40±1.10	3.41±1.10	0.946
FVC (L)	4.19±1.34	4.17±1.47	0.940
FIV1 (L)	3.38±1.19	3.22±1.02	0.456
FEF50 (L/sec)	4.33±1.74	4.40±1.25	0.834
FIF50 (L/sec)	4.41±1.73	3.96±1.70	0.155
PEF (L/sec)	8.16±6.48	7.14±2.60	0.324
PIF (L/sec)	4.73±1.74	4.31±1.67	0.188
FEV1/FVC (%)	81.10±7.90	83.14±5.44	0.129
FEF50/FIF50 ratio	1.15±0.92	1.21±0.37	0.721
FEF25-75/FVC ratio	0.87±0.30	2.55±1.62	**<0.001**

Data are shown as mean±standard deviation. OSAS=Obstructive sleep apnea syndrome, PFT=Pulmonary function test, FEV1=Forced expiratory volume in the 1st second, FVC=Forced vital capacity, FIV1=Forced inspiratory volume in the 1st second, FEF50=Maximal expiratory flow rates 50%, FIF50=Maximal inspiratory flow rates 50%, PEF=Peak expiratory flow, PIF=Peak inspiratory flow, FEF25-75=Maximal mid-expiratory flow rate

*Mann-Whitney U test

icant abnormalities. Group differences in PSG parameters are summarized in Table 1.

Pulmonary Function Test Results

In Table 2, PFT parameters were compared between the two groups. The FEF25-75/FVC ratio was significantly lower in the OSAS patients compared to the control group (P<0.001). Sawtooth signs were observed in 4 OSAS patients and 2 controls; however, this difference was not statistically significant due to the small sample size.

Correlation Between PSG Parameters and Spirometry Results

In OSAS patients, correlation analysis revealed that total hypopnea count was significantly negatively associated with FEV1, FVC, FIV1, FEF50, FIF50, and PEF (P<0.05). Total apnea count showed a significant positive correlation with PEF (P<0.05). Minimum SaO_2 levels correlated positively with FEV1, FVC, FIV1, and FEF50, whereas mean SaO_2 was only positively correlated with FEF50 (P<0.05). Additionally, the duration of time with oxygen saturation below 90% was negatively correlated with FEV1, FVC, FIV1, and FEF50 values. These results are summarized in Table 3.

Comparison of OSAS Patients with and without Sawtooth Sign

In OSAS patients with a sawtooth sign, AHI total, AHI REM, AHI side, and total apnea were higher compared to those without a sawtooth sign. Mean SaO_2, however, was significantly lower in OSAS patients with a sawtooth sign than in those without (Table 4).

Comparison of OSAS Patients with and without FEF50/FIF50 > 1

Forty-six (46.5%) of the OSAS patients had an FEF50/FIF50 > 1, whereas 30 out of 42 individuals (71.4%) in the control group had this value. A comparison between OSAS patients with and without an FEF50/FIF50 > 1 is summarized in Table 5.

DISCUSSION

In this study, no significant differences were observed in PFT parameters between OSAS patients and the control group, except for the FEF25-75/FVC ratio, which was significantly lower in OSAS patients (0.87±0.30) compared to controls (2.55±1.62). AHI, AHI REM, AHI side, and total apnea values of the

Table 3. Correlation between PSG parameters and PFT values in patients with OSAS

PSG results	PFT values									
	FEV1 (L)	FVC (L)	FIV1 (L)	FEF50 (L/sec)	FIF50 (L/sec)	PEF (L/sec)	PIF (L/sec)	FEV1/FVC (%)	FEF50/FIF50 ratio	FEF25-75/FVC ratio
	r	r	r	r	r	r	r	r	r	r
AHI total	-0.028	-0.019	-0.028	0.009	-0.069	0.117	-0.024	0.035	0.041	0.017
AHI side	-0.129	-0.127	-0.129	-0.075	-0.104	-0.003	-0.075	0.027	0.007	0.008
AHI supine	0.077	0.098	0.077	0.094	-0.078	0.168	-0.040	0.044	0.110	0.044
Total apnea	0.052	0.054	0.052	0.077	-0.022	0.208*	0.004	0.031	0.038	0.017
Total hypopnea	-0.295*	-0.269*	-0.295*	-0.237*	-0.210*	-0.244*	-0.177	0.009	0.005	0.017
Min SaO_2 (%)	0.307*	0.291*	0.307*	0.236*	0.170	0.121	0.173	0.021	0.066	0.032
Mean SaO_2 (%)	0.188	0.149	0.188	0.225*	0.092	0.035	0.103	0.117	0.108	0.139
Duration of <90% SaO_2	-0.214*	-0.183	-0.214*	-0.203*	-0.100	-0.079	-0.104	-0.098	-0.074	-0.112
Duration of REM (%)	0.130	0.120	0.130	0.161	0.066	0.165	0.065	0.112	0.090	0.105
Sleep efficiency	0.046	0.040	0.046	0.073	0.012	0.075	-0.008	0.059	0.003	0.058

OSAS=Obstructive sleep apnea syndrome, PFT=Pulmonary function test, PSG=Polysomnography, FEV1=Forced expiratory volume in the 1st second, FVC=Forced vital capacity, FEV1: Forced expiratory volume in the 1st second, FEF50=Maximal expiratory flow rates 50%, FIF50=Maximal inspiratory flow rates 50%, PEF=Peak expiratory flow, FEF25-75=Maximal mid-expiratory flow rate, PIF=Peak inspiratory flow, AHI=Apnea hypopnea index, REM=Rapid eye movement, Mean SaO_2=Mean oxygen saturation, Min SaO_2=Minimum oxygen saturation
*r=Spearman correlation coefficient; P<0.05

OSAS patients with sawtooth signs were determined to be significantly higher than those without sawtooth signs. The presence of sawtooth signs in the OSAS patients was found to be beneficial in indicating severe OSAS. Furthermore, correlation analysis also demonstrated significant associations between specific PFT parameters and hypopnea and desaturation indices in OSAS patients.

Alterations in the URT seen in OSAS may be reflected in PFTs performed while awake. A previous study showed that the quantitative flow-volume curve standard has high sensitivity and specificity in detecting upper airway obstruction [15]. Notably, a FEF50/FIF50 ratio >1 and the sawtooth pattern observed in the flow-volume curve were both strongly associated with OSAS, indicating URT obstruction and tissue vibration [6, 16]. In a study of 138 OSAS patients, 26.1% had a FEF50/FIF50 ratio >1. No significant differences in age, BMI, AHI, desaturation, or standard PFT parameters were found between patients with and without a FEF50/FIF50 ratio >1 [7]. In our study, while the FEF50/FIF50 ratio did not differ significantly between OSAS patients and controls, a noteworthy distinction emerged within the OSAS group itself. OSAS patients with a ratio >1 exhibited significantly lower inspiratory flow values (PIF and FIV1) compared to those with a ratio ≤1. Additionally, patients with a lower FEF50/FIF50 ratio ≤1 had significantly reduced FEF25-75/FVC and FEV1/FVC ratios, suggesting both small airway dysfunction and potentially greater upper airway resistance. These findings indicate that while the FEF50/FIF50 ratio may offer insights into upper airway resistance, its utility as a predictor of OSAS severity remains limited.

The FEF25-75/FVC ratio has been suggested to reflect small airway function and possibly disproportionate development between the pulmonary parenchyma and airways, which may be associated with increased upper airway resistance [17]. Previous studies have linked a low FEF25-75/FVC ratio to airway hypersensitivity and to a higher prevalence of OSAS in patients with unstable asthma or obesity [8, 18]. While some reports have found no direct correlation between this ratio and PSG parameters [19], our findings demonstrated a significant decrease in the FEF25-75/FVC ratio among OSAS patients compared to controls. This reduction may reflect elevated upper airway collapsibility or resistance, potentially related

Table 4. Demographic features, PFT parameters, and PSG results of the OSAS patients with and without the sawtooth sign

Variables	Those with the sawtooth sign (n=4)	Those without the sawtooth sign (n=95)	P value*
Age (years)	49.50±4.45	48.37±10.55	0.832
BMI (kg/m^2)	32.77±5.15	33.61±10.04	0.859
FEV1 (L)	2.95±0.98	3.42±1.11	0.378
FVC (L)	3.45±1.08	4.22±1.35	0.263
FEV1/FVC ratio	84.00±6.61	80.97±7.95	0.457
PEF (L/sec)	6.16±1.18	8.23±6.60	0.531
FEF50 (L/sec)	3.85±1.60	4.36±1.75	0.557
FIF50 (L/sec)	3.05±1.31	4.40±1.73	0.110
FEF50/FIF50 ratio	1.44±0.97	1.19±0.92	0.524
FEF25-75/FVC ratio	0.97±0.27	0.86±0.30	0.492
AHI total	64.27±29.84	35.43±24.24	**0.023**
AHI REM	78.47±30.37	32.66±25.09	**0.001**
AHI supine	69.65±39.67	49.24±30.14	0.193
AHI side	60.95±37.83	24.74±23.82	**0.004**
Total apnea	56.37±27.73	25.11±22.46	**0.008**
Mean SaO_2 (%)	84.00±8.40	89.12±4.85	**0.048**
Min SaO_2 (%)	64.75±15.84	74.12±11.87	0.130

OSAS=Obstructive sleep apnea syndrome, PFT=Pulmonary function test, PSG=Polysomnography, BMI=Body mass index, FEV1=Forced expiratory volume in the 1st second, FVC=Forced vital capacity, PEF=Peak expiratory flow, FEF50=Maximal expiratory flow rates 50%, FIF50=Maximal inspiratory flow rates 50%, FEF25-75=Maximal mid-expiratory flow rate, AHI=Apnea hypopnea index, REM=Rapid eye movement, Mean SaO^2=Mean oxygen saturation, Min SaO_2=Minimum oxygen saturation

*Mann-Whitney U test

to increased BMI and reduced pharyngeal muscle tone, both of which are commonly observed in OSAS patients. These findings support the potential utility of this ratio as a supportive indicator of upper airway instability in the context of OSAS.

In this study, correlation analysis revealed significant relationships between total hypopnea and oxygen saturation parameters and spirometric indices that reflect inspiratory and expiratory flow limitations in OSAS patients. Total hypopnea count showed negative correlations with FEV1, FVC, FIV1, FEF50, FIF50, and PEF, suggesting that increased hypopnea burden may be associated with notable impairments in respiratory function. While minimum and mean SaO_2 values showed positive correlations with PFT parameters, the duration of <90% SaO_2 was negatively correlated. These findings support the notion that individuals with better respiratory function may experience less desaturation in the context of OSAS. However, the lack of a significant correlation between AHI and spirometric parameters suggests that PFT may not directly indicate the presence of OSAS but may instead reflect URT instability. Similar findings have been reported in previous studies, where no significant correlation was found between AHI severity and spirometric values such as FEV1 or FVC, despite the presence of pulmonary function abnormalities in OSAS patients [20].

In a previous study, among patients undergoing spirometry for any indication, those with the sawtooth sign were found to have a higher likelihood of being diagnosed with OSA compared to those without the

Table 5. Comparison of demographic features, PFT parameters, and PSG results of the OSAS patients with and without a FEF50/FIF50 >1

Variables	FEF50/FIF50 >1 (n=46)	FEF50/FIF50 ≤1 (n=53)	P value
Age (years)	48.15±10.66	48.65±10.21	0.810
BMI (kg/m^2)	33.03±6.62	34.19±12.14	0.564
FEV1 (L)	3.52±1.07	3.29±1.13	0.292
FVC (L)	4.13±1.32	4.25±1.37	0.640
FEV1/FVC ratio	85.78±4.97	76.86±7.69	**<0.001**
PEF (L/sec)	7.57±2.56	8.70±8.62	0.391
PIF (L/sec)	4.08±1.54	5.31±1.71	**<0.001**
FIV1 (L)	3.09±1.16	3.64±1.17	**0.021**
FEF25-75/FVC ratio	1.06±0.24	0.70±0.24	**<0.001**
AHI total	37.88±25.62	35.43±24.57	0.628
AHI REM	36.38±28.81	32.85±24.86	0.519
AHI supine	53.55±31.35	46.91±29.84	0.283
AHI side	26.27±25.18	26.14±25.62	0.981
Total apnea	27.02±23.43	25.83±23.55	0.804
Mean SaO_2 (%)	89.00±5.41	88.84±4.80	0.881
Min SaO_2 (%)	73.51±12.98	73.96±11.36	0.854

OSAS=Obstructive sleep apnea syndrome, PFT=Pulmonary function test, PSG=Polysomnography, BMI=Body mass index, FEV1=Forced expiratory volume in the 1st second, FVC=Forced vital capacity, PEF=Peak expiratory flow, PIF=Peak inspiratory flow, FIV1=Forced inspiratory volume in the 1st second, FEF25-75=Maximal mid-expiratory flow rate, AHI=Apnea hypopnea index, MeanSaO_2=Mean oxygen saturation, MinSaO_2=Minimum oxygen saturation

*Mann-Whitney U test

sign [10]. Several studies have since explored this association in more detail. In a study examining flow-volume curves of 401 patients who presented with snoring and were investigated for OSAS, apnea was found to be more severe in patients with a sawtooth sign during both inspiration and expiration. Sensitivity and specificity of the sawtooth sign in identifying OSAS have been reported to vary widely, ranging from 29% to 85% and 56% to 95%, respectively [10]. In another study, the FEF50/FIF50 ratio >1 and/or sawtooth sign was observed in 40% of OSAS patients, while these features were absent in the control group. OSAS patients with a sawtooth sign had significantly larger height, weight, and neck circumference, and worse PSG results. However, there were no differences in spirometric parameters between OSAS patients with and without the sawtooth sign [19]. In a study investigating the diagnostic value of the flow-volume curve for OSAS, sawtooth signs were significantly more prevalent in OSAS patients compared to other groups ($P < 0.01$). OSAS patients with sawtooth signs had a higher apnea index and more significant reductions in SaO_2 [21]. Another study found that 12.3% of 138 OSAS patients exhibited sawtooth signs, but no significant differences were observed in age, BMI, AHI, desaturation index, or spirometric parameters between those with and without sawtooth signs [7]. In this study, sawtooth signs were present in 4 OSAS patients and 2 controls. While no significant differences were found between these groups regarding demographic or PFT parameters, OSAS patients with sawtooth signs had significantly higher AHI values, AHI REM, AHI side, and total apnea values, as well as lower mean SaO_2. The small sample size of the

subgroup with sawtooth signs (n=4) limits the statistical power and generalizability of these findings. Therefore, the results should be interpreted with caution, and further studies with larger cohorts are needed to validate these observations.

Flow-volume curves and PFT parameters are not reliable screening tools for differentiating OSAS patients, as features like the sawtooth sign and FEF50/FIF50 ratio >1 were also observed in the control group. However, the lower FEF25–75/FVC ratio observed in OSAS patients may reflect increased upper airway resistance or instability. This spirometric finding may be associated with OSAS; however, further prospective studies are required to evaluate its diagnostic utility. Notably, the sawtooth sign appeared to be useful in identifying more severe cases of OSAS.

Strengths and Limitations

This study addresses a relatively underexplored area by investigating the relationship between spirometric parameters and OSAS. Additionally, the use of correlation analyses between PSG findings and PFT parameters strengthens the clinical relevance of our results. Finally, by focusing on commonly used and cost-effective spirometric measurements, this study highlights a potentially accessible adjunct tool in the evaluation of OSAS. However, several limitations should be considered, including the small control group and single-center design, which may affect generalizability. Additionally, the lack of a detailed evaluation of confounding factors and the retrospective nature of the study may influence data accuracy. Although the study had a retrospective design, the relationship between flow-volume curve abnormalities and OSAS remains an underexplored area in the literature. To our knowledge, few studies have evaluated whether these PFT patterns can help predict the severity of OSAS. Therefore, we believe that our findings still provide relevant clinical insight. Further prospective studies with standardized methods are necessary to better understand the relationship between PFT abnormalities and OSAS.

CONCLUSION

Given the complex pathophysiology of OSAS, identifying abnormalities in flow-volume curves may serve as an early indicator of disorders leading to OSAS. This approach can help in determining which patients should undergo PSG, which is time-consuming and costly. We suggest that multi-center studies, involving a larger number of cases and identifying modifying factors, are needed to assess the reliability of changes in flow-volume curves for predicting OSAS in patients.

Ethics Approval and Consent to Participate

This study was approved by the Konya University Meram Faculty of Medicine Non-Interventional Clinical Research Ethics Committee (Decision No.: 2012/28 and dated 02.03.2012). All procedures performed during data collection, review of patient records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Informed consent was waived because of the retrospective nature of the study and the analysis used anonymous clinical data.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: FAA, ŞY; Study Design: FAA, ŞY; Supervision: FAA, ŞY; Funding: ŞY; Materials: ŞY; Data Collection and/or Processing: FAA; Statistical Analysis and/or Data Interpretation: FAA, ŞY; Literature Review: FAA; Manuscript Preparation: FAA; and Critical Review: FAA, ŞY.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Griffin W, le Roux CW, Heneghan HM, O'Shea D, Garvey JF. The role of weight management in sleep disordered breathing. Breathe (Sheff). 2025;21(3):250182. doi: 10.1183/20734735.0182-2025.
2. Stanczyk M, McNicholas WT, Pevernagie DA, Riha RL, Ryan S. Management of Adult Obstructive Sleep Apnoea: Many Questions, Not Enough Answers! J Sleep Res. 2025;34(5):e70047. doi: 10.1111/jsr.70047.
3. Lee JJ, Sundar KM. Evaluation and Management of Adults with Obstructive Sleep Apnea Syndrome. Lung. 2021;199(2):87-101. doi: 10.1007/s00408-021-00426-w.
4. Lv R, Liu X, Zhang Y, et al. Pathophysiological mechanisms and therapeutic approaches in obstructive sleep apnea syndrome. Signal Transduct Target Ther. 2023;8(1):218. doi: 10.1038/s41392-023-01496-3.
5. Mohning MP, Krefft SD. Use of Spirometry in Pulmonary Function Evaluation. Clin Chest Med. 2025;46(3):401-412. doi: 10.1016/j.ccm.2025.04.001.
6. Campbell AH, Guy PA, Rochford PD, Worsnop CJ, Pierce RJ. Flow-volume curve changes in patients with obstructive sleep apnoea and brief upper airway dysfunction. Respirology. 2000;5(1):11-8. doi: 10.1046/j.1440-1843.2000.00220.x.
7. Ashraf M, Shaffi SA, BaHammam AS. Spirometry and flow-volume curve in patients with obstructive sleep apnea. Saudi Med J. 2008;29(2):198-202.
8. Parker AL, Abu-Hijleh M, McCool FD. Ratio between forced expiratory flow between 25% and 75% of vital capacity and FVC is a determinant of airway reactivity and sensitivity to methacholine. Chest. 2003;124(1):63-69. doi: 10.1378/chest.124.1.63.
9. Liam CK, Lim KH, Wong CM, Lau WM, Tan CT. Awake respiratory function in patients with the obstructive sleep apnoea syndrome. Med J Malaysia. 2001;56(1):10-17.
10. Bourne MH Jr, Scanlon PD, Schroeder DR, Olson EJ. The sawtooth sign is predictive of obstructive sleep apnea. Sleep Breath. 2017;21(2):469-474. doi: 10.1007/s11325-016-1441-x.
11. Sateia MJ. International classification of sleep disorders-third edition: highlights and modifications. Chest. 2014;146(5):1387-1394. doi: 10.1378/chest.14-0970.
12. Ito E, Inoue Y. [The International Classification of Sleep Disorders, third edition. American Academy of Sleep Medicine. Includes bibliographies and index]. Nihon Rinsho. 2015;73(6):916-923. [Article in Japanese]
13. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. Am J Respir Crit Care Med. 2019;200(8):e70-e88. doi: 10.1164/rccm.201908-1590ST.
14. Levent E, Sarıman N. Analysis of obstructive sleep apnea patients with "sawtooth sign" on the flow-volume curve. Sleep Breath. 2011;15(3):357-365. doi: 10.1007/s11325-010-0393-9.
15. Wei D, Wang L, Yu Z, et al. Forced Inspiratory Flow Volume Curve in Patients with Obstructive Sleep Apnea-Hypopnea Syndrome. Int J Clin Med. 2023;14(5):260-273. doi: 10.4236/ijcm.2023.145022
16. Thorve SM, Gupta V, Mandilwar S, Modi N, Prabhudesai P. Spirometry and Flow Volume Loops in Obstructive Sleep Apnea Patients. J Int J Contemp Med Res. 2019;6(12):L5-10. doi: 10.21276/ijcmr.2019.6.12.25
17. Tager IB, Weiss ST, Muñoz A, Welty C, Speizer FE. Determinants of response to eucapneic hyperventilation with cold air in a population-based study. Am Rev Respir Dis. 1986;134(3):502-508. doi: 10.1164/arrd.1986.134.3.502.
18. Yigla M, Tov N, Solomonov A, Rubin AH, Harlev D. Difficult-to-control asthma and obstructive sleep apnea. J Asthma. 2003;40(8):865-871. doi: 10.1081/jas-120023577.
19. Oztürk L, Metin G, Cuhadaroğlu C, Utkusavaş A, Tutluoğlu B. FEF(25-75)/FVC measurements and extrathoracic airway obstruction in obstructive sleep apnea patients. Sleep Breath. 2005;9(1):33-38. doi: 10.1007/s11325-005-0006-1.
20. Mehfooz N, Siraj F, Shabir A, et al. Spirometric abnormalities in patients with sleep-related breathing disorders. J Family Med Prim Care. 2021;10(2):1009-1014. doi: 10.4103/jfmpc.jfmpc_1018_20.
21. Rauscher H, Popp W, Zwick H. Flow-volume curves in obstructive sleep apnea and snoring. Lung. 1990;168(4):209-214. doi: 10.1007/BF02719694.

The European Research Journal 2025;11(6):1167-1175 Original Article

DOI: https://doi.org/10.18621/eurj.1632009 *Primary Health Care*

Comparing the effectiveness of different vaccination regimens using Sinovac and BNT162b2 vaccines among hospitalised patients: A single-centre hospital-based retrospective cohort study

Atakan Turan[1], Hamdi Öğüt[1], Aylin Ayyıldız Varol[2], Nizameddin Koca[3], Hayri Bozkurt[4]

[1]Department of Bioengineering, Faculty of Engineering and Natural Sciences, Bursa Technical University, Bursa, Türkiye; [2]Department of Internal Medicine, Bursa Çekirge State Hospital, Bursa, Türkiye; [3]Department of Internal Medicine, University of Health Sciences, Bursa City Training and Research Hospital, Bursa, Türkiye; [4]Department of Home Care, Bursa Çekirge State Hospital, Bursa, Türkiye

ABSTRACT

Objectives: This study aimed to assess the effectiveness of different vaccination regimens using two distinct SARS-CoV-2 vaccines against mortality risk and the need for intensive care unit (ICU) admission among hospitalised patients.

Methods: The single-centre hospital-based retrospective cohort study was performed with adult COVID-19 patients in a tertiary-level hospital between March 2020 and September 2022. The associations between patients' demographics and clinical features, vaccine status and regimens, in-hospital mortality, and need for ICU admission were evaluated using multivariable regression analyses.

Results: During the study period, 2,373 patients were included. Mortality among unvaccinated patients was 85.0%, which was significantly lower in vaccinated groups (P<0.001), particularly with BNT162b2 than with Sinovac. Vaccination reduced mortality and ICU admission rates, with higher efficacy observed with increased vaccine doses and BNT162b2 regimens. Multivariable analyses confirmed age as a significant determinant and various vaccination schedules showed consistent reductions in mortality and ICU admissions.

Conclusions: A two-dose initial plus one or more-dose booster BNT162b2 regimen effectively reduced mortality risks and ICU admission.

Keywords: COVID-19, vaccine, mortality, intensive care unit

SARS-CoV-2 infection (COVID-19) is characterised by a respiratory syndrome that can vary in severity, ranging from mild effects to life-threatening consequences. Its fatality rate is approximately 2.3%, lower than SARS-CoV and MERS-CoV infections [1]. It is noteworthy that COVID-19 has caused significantly more deaths globally because SARS-CoV-2 is widely transmitted within the community, whereas SARS-CoV and MERS-CoV are primarily transmitted among hospitalised patients [2]. Previous studies have demonstrated that both anticipated annual viral respiratory infections and unfore-

Received: February 3, 2025 **Accepted:** March 9, 2025 **Available Online:** June 1, 2025 **Published:** November 4, 2025

How to cite this article: Turan A, Öğüt H, Ayyıldız Varol A, Koca N, Bozkurt H. Comparing the effectiveness of different vaccination regimens using Sinovac and BNT162b2 vaccines among hospitalised patients: A single-centre hospital-based retrospective cohort study. Eur Res J. 2025;11(6):1167-1175. doi: 10.18621/eurj.1632009

Corresponding author: Atakan Turan, Pharmacist, PhD Student, Phone: +90 224 300 32 32, E-mail: ecz.atakanturan@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

seen pandemic outbreaks result in three to five million cases of severe disease and up to half a million deaths worldwide [3, 4].

The unprecedented pace of vaccine development against COVID−19 and robust global vaccination efforts play critical roles in controlling the pandemic [5]. Few studies have investigated the effectiveness of various vaccine doses and schedules, including the use of different vaccines and complementary administration of a booster dose from another vaccine, in preventing disease severity and mortality during the COVID-19 pandemic. Numerous studies have sought to delineate the number of applications of the same vaccine deemed a full dose and the number of doses considered insufficient within this context [6, 7]. However, the production processes and availability of vaccines have varied across countries during the COVID-19 pandemic, leading to the implementation of mixed vaccination regimens involving different vaccines in many regions. A limited number of studies compared the incidence of severe infections among unvaccinated, partially vaccinated, and fully vaccinated individuals. Therefore, this study aimed to evaluate the efficacy of various vaccines and, more critically, to ascertain which specific vaccines and vaccination regimens are most effective in mitigating the progression to severe disease and reducing mortality rates.

METHODS

This single-centre hospital-based retrospective cohort study included patients with COVID-19 admitted to a tertiary-level hospital. The Clinical Research Ethics Committee approved this study (Approval date: 20.04.2022, Approval number: 2011-KAEK-25 2022/04-18). The study followed the ethical principles outlined in the Declaration of Helsinki and the guidelines for Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement. Informed written consent was not obtained because of the study's retrospective design.

The cohort was defined as "patients aged 18 years and older who were admitted with a diagnosis of COVID-19 between March 2020 and September 2022." According to the cohort definition, the inclusion criteria for the study population were (1) being 18 years or older, (2) having a microbiologically confirmed diagnosis of COVID-19, and (3) being admitted to the hospital. Patients for whom vaccination information, the primary independent variable of the study, was not accessible were excluded from the study. Patients vaccinated with any vaccine other than Sinovac or BNT162b2 were also excluded from the study. However, patients with missing data for other variables were not excluded to provide different types of analyses involving more variables.

Data, including patients' demographics (age and sex), vaccination status, blood group, COVID-19 Reporting and Data System (CO-RADS) stage, Acute Physiology and Chronic Health Evaluation Mortality Prediction Rate (APACHE II-MPR), acute phase reactants [ferritin (ng/mL), procalcitonin (ng/mL), and C-Reactive Protein (CRP) (mg/L)], D-dimer (mg/L FEU), troponin (ng/mL), white blood cell (WBC) count (x109/L), and platelet (×109/L) levels were obtained from electronic medical records.

The vaccination status of the patients was categorised as follows: (1) unvaccinated, (2) one- or two-dose vaccination with Sinovac, (3) two initial doses plus one-dose booster Sinovac, (3) two initial doses plus two- or more-dose booster Sinovac, (4) one- or two-dose BioNTech (BNT162b2), (5) two initial doses plus one-dose booster BNT162b2, (6) two initial doses plus two- or more-dose booster BNT162b2, (7) two initial doses Sinovac plus one-dose booster BNT162b2, and (8) two initial doses Sinovac plus two- or more-dose booster BNT162b2.The underlying rationale behind this categorisation is to consider those who received one or two doses of any vaccine as "inadequately vaccinated." In contrast, individuals who received two initial doses and one booster dose are regarded as "fully vaccinated."

The CO-RADS was used to include the radiological disease status of patients in the analysis [8]. The APACHE II comprises 12 physiological variables: age and previous disease status. The APACHE II score was converted into a percentage to derive the Mortality Prediction Rate (MPR) variable [9]. This approach allowed for the integration and standardised evaluation of multiple variables as a single independent variable.

There were two primary outcomes of the study. These are "in-hospital mortality" and "the need for an intensive care unit (ICU)". "Mortality" was defined as

"all-cause mortality during the stay in the hospital." "The need for ICU" was described as "a need for hospitalization in the intensive care unit for any reason during the stay in the hospital."

Statistical Analysis

No prior sample size was calculated. All eligible patients' data were included in the analysis. Statistical analyses were performed using SPSS version 23 (IBM Corp., Armonk, NY, USA). Descriptive statistics are presented as median with interquartile range (IQR) for numerical data and frequency (n) and percentage (%) for categorical data. Pearson Chi-square Test was used to analyse mortality and ICU admission rates among the study groups with different vaccination statuses. Univariable Logistic Regression Analyses with the Enter Method were employed for univariable analysis to estimate the effects of variables and potential covariates on mortality risk and the need for ICU risk. Several multivariable models were constructed with statistically significant variables from the univariable analyses while avoiding highly correlated independent variables. Odds Ratios (ORs) with a 95% confidence interval (CI) were used to evaluate the risk. Statistical significance was set at $P<0.05$.

RESULTS

The data of 2373 patients were included in the analysis based on the inclusion and exclusion criteria of the study. The number of patients included in the analysis decreased to 916 for some variables. Patients with missing data were not excluded from the analysis to present more comprehensive findings. The median age was 74 years, and 56.9% were male. The median body mass index (BMI) was 26.6 kg/m2, and 48.5% were unvaccinated. Although 5.5% of patients were classified as CO-RADS stage 1, 20.1% as stage 2, 20.2% as stage 3, 24.3% as stage 4, and 29.9% as stage 5, the median APACHE MPR score was 42.8%. Sepsis developed in 13 patients, and the median length of stay was 6.0 days. Notably, the levels of acute-phase reactants were elevated in all patients (Table 1).

Mortality was 85.0% (n=979) among unvaccinated patients, whereas it significantly decreased in all vaccinated groups ($P<0.001$ for the comparisons between the unvaccinated group and all vaccination regimens). Mortality rates were 66.9% in patients vaccinated with one or two-dose Sinovac, 55.9% with two-dose plus one-dose booster Sinovac, 26.1% with two-dose plus two or more-dose boosters of Sinovac, 27.3% with one or two-dose of BNT162b2, 3.3% with two-dose plus one-dose booster of BNT162b2, 3.6% with two-dose plus two or more-dose booster BNT162b2, 47.1% with two-dose Sinovac plus one-dose booster BNT162b2, and 21.7% two-dose Sinovac plus two or more-dose booster BNT162b2. The need for ICU admission followed a similar trend. Among unvaccinated patients, 81.3% (n=937) required ICU care, which was significantly higher compared to vaccinated groups ($P<0.001$ for the comparisons between the unvaccinated group and all vaccination regimens). The ICU admission rates 65.2% in patients vaccinated with one or two-dose Sinovac, 54.0% with two-dose plus one-dose booster Sinovac, 37.7% with two-dose plus two or more-dose boosters of Sinovac, 29.7% with one or two-dose of BNT162b2, 9.8% with two-dose plus one-dose booster of BNT162b2, 17.9% with two-dose plus two or more-dose booster BNT162b2, 49.6% with two-dose Sinovac plus one-dose booster BNT162b2, and 30.2% two-dose Sinovac plus two or more-dose booster BNT162b2. Overall mortality and ICU admission rates for the entire cohort were 64.0% (n=1519) and 63.3% (n=1502), respectively (data not shown).

Age was a statistically significant determinant of mortality and the need for ICU. Compared to unvaccinated individuals, all vaccination schedules were associated with reduced mortality risk. Moreover, as the number of booster doses increased, both the mortality and ICU admission risks decreased. When evaluating between vaccines, the regimen that most significantly reduced the mortality and ICU admission was the BNT162b2 vaccine at both the initial and booster doses, followed by the Sinovac initial regimen with BNT162b2 boosters. The slightest reduction in mortality was observed with the Sinovac regimen for both the initial and booster doses (Table 2). This trend was also the same for the need for an ICU (Table 3).

Three different models were employed to evaluate mortality risk for significant variables in the univariable analyses. The findings from the univariable analyses were preserved across all the models (Table 2). Similar results were obtained in the modelling of ICU

Table 1. Demographics, clinical features and laboratory findings of the patients

Characteristics	Data
Age (years), (n=2373)	74.0 (64.0-82.0)
Sex, n (%) (n=2373)	
Female	1022 (43.1)
Male	1351 (56.9)
BMI (kg/m^2), (n=916)	26.6 (23.4-31.1)
Blood group, n (%) (n=2146)	
0 Rh (-)	105 (4.9)
0 Rh (+)	551 (25.7)
A Rh (-)	101 (4.7)
A Rh (+)	864 (40.3)
B Rh (-)	42 (2.0)
B Rh (+)	320 (14.9)
AB Rh (-)	14 (0.7)
AB Rh (+)	149 (6.9)
Vaccine status, n (%) (n=2373)	
Unvaccinated	1152 (48.5)
One or two-dose Sinovac	408 (17.2)
Two-dose plus one-dose booster Sinovac	202 (8.5)
Two-dose plus two or more-dose booster Sinovac	69 (2.9)
One or two-dose BNT162b2	172 (7.2)
Two-dose plus one-dose booster BNT162b2	92 (3.9)
Two-dose plus two or more-dose booster BNT162b2	28 (1.2)
Two-dose plus one-dose booster mix	121 (5.1)
Two-dose plus two or more-dose booster mix	129 (5.4)
CO-RADS, n (%) (n=2108)	
1	115 (5.5)
2	423 (20.1)
3	426 (20.2)
4	513 (24.3)
5	631 (29.9)
APACHE MPR (%), (n=1557)	42.8 (26.2-70.5)
Occurrence of sepsis, n (%) (n=1427)	13 (0.9)
LOS (days), (n=2373)	6.0 (2.0-11.0)
Ferritin (ng/mL), (n=1916)	602.70 (280.62-1407.41)
Procalcitonin (ng/mL), (n=593)	0.23 (0.10-0.62)
CRP (mg/L), (n=2301)	79.20 (16.20-176.00)
D-Dimer (mg/L FEU), (n=2214)	2.08 (0.95-5.72)
Troponin (ng/mL), (n=2287)	50.90 (10.94-269.60)
WBC count (×10^9/L), (n=2361)	11.36 (7.62-17.42)
Platelet (×10^9/L), (n=1916)	199.00 (123.50-282.00)

Data are shown as median (IQR-Interquartile range) or n (%). BMI=Body mass index, CO-RADS=COVID-19 Reporting and Data System, APACHE MPR=Acute Physiology, Age, and Chronic Health Evaluation-Mortality Prediction Rate, LOS=Length of stay, CRP=C-Reactive Protein, WBC=White blood cell.

Table 2. **Univariable and multivariable analyses in estimating the mortality risk**

Variables	Univariable Analyses[1]		Multivariable Analyses[2]					
			Model 1 (n=2373)		Model 2 (n=2108)		Model 3 (n=1374)	
	OR (95% CI)	P value	aOR (95% CI)	P value	aOR (95% CI)	P value	aOR (95% CI)	P value
Age (years), (n=2373)	1.04 (1.03-1.05)	**<0.001**	1.04 (1.03-1.05)	**<0.001**	1.04 (1.03-1.05)	**<0.001**	1.03 (1.02-1.04)	**<0.001**
Male sex, (n=2373)	1.06 (0.90-1.25)	0.534						
Vaccine status, (n=2373)								
Unvaccinated	Reference		Reference		Reference		Reference	
One or two-dose Sinovac	0.36 (0.28-0.46)	**<0.001**	0.29 (0.22-0.38)	**<0.001**	0.23 (0.28-0.46)	**<0.001**	0.36 (0.21-0.60)	**<0.001**
Two-dose plus one-dose booster Sinovac	0.22 (0.16-0.31)	**<0.001**	0.17 (0.12-0.24)	**<0.001**	0.12 (0.16-0.31)	**<0.001**	0.19 (0.10-0.34)	**<0.001**
Two-dose plus two or more-dose booster Sinovac	0.06 (0.04-0.11)	**<0.001**	0.05 (0.03-0.09)	**<0.001**	0.04 (0.04-0.11)	**<0.001**	0.03 (0.01-0.07)	**<0.001**
One or two-dose BNT162b2	0.07 (0.05-0.10)	**<0.001**	0.08 (0.06-0.12)	**<0.001**	0.07 (0.05-0.10)	**<0.001**	0.15 (0.07-0.29)	**<0.001**
Two-dose plus one-dose booster BNT162b2	0.01 (0.01-0.02)	**<0.001**	0.01 (0.01-0.03)	**<0.001**	0.01 (0.01-0.02)	**<0.001**	0.01 (0.01-0.05)	**<0.001**
Two-dose plus two or more-dose booster BNT162b2	0.01 (0.01-0.05)	**<0.001**	0.01 (0.01-0.06)	**<0.001**	0.01 (0.01-0.05)	**<0.001**	0.01 (0.01-0.14)	**<0.001**
Two-dose Sinovac plus one-dose booster BNT162b2	0.16 (0.11-0.23)	**<0.001**	0.13 (0.09-0.19)	**<0.001**	0.09 (0.11-0.23)	**<0.001**	0.16 (0.07-0.33)	**<0.001**
Two-dose Sinovac plus two or more-dose booster BNT162b2	0.05 (0.03-0.08)	**<0.001**	0.04 (0.03-0.07)	**<0.001**	0.03 (0.03-0.08)	**<0.001**	0.04 (0.02-0.09)	**<0.001**
Blood group, (n=2146)								
0 Rh (-)	2.39 (0.77-7.38)	0.131						
0 Rh (+)	1.77 (0.61-5.12)	0.293						
A Rh (-)	1.97 (0.64-6.08)	0.238						
A Rh (+)	1.79 (0.76-5.14)	0.282						
B Rh (-)	1.63 (0.48-5.50)	0.435						
B Rh (+)	1.94 (0.66-5.66)	0.228						
AB Rh (-)	Reference							
AB Rh (+)	1.66 (0.55-4.98)	0.366						
CO-RADS, (n=2108)								
1					Reference		Reference	
2					1.71 (1.05-2.79)	**0.005**	1.33 (0.54-3.27)	0.542
3					2.47 (1.50-4.06)	**<0.001**	1.51 (0.61-3.76)	0.374
4					2.60 (1.59-4.23)	**<0.001**	1.57 (0.65-3.78)	0.320
5					3.40 (2.10-5.51)	**<0.001**	3.12 (1.27-7.65)	**0.013**
APACHE MPR (%), (n=1557)							1.04 (1.03-1.05)	**<0.001**

OR=Odds Ratio, aOR=Adjusted Odds Ratio, CI=Confidence interval, CO-RADS=COVID-19 Reporting and Data System, APACHE MPR=Acute Physiology and Chronic Health Evaluation Mortality Prediction Rate.

[1]Univariate Logistic Regression Model with the Enter Method was used.

[2]Multivariate Logistic Regression Models with the Enter Method were used. The variables shown in each column were included in the model.

[3]"n" in the first column shows the number of patients included in the univariable analyses.

Table 3. Univariable and multivariable analyses in estimating the need for ICU risks

Variables	Univariable Analyses[1]		Multivariable Analyses[2]											
			Model 1 (n=2373)		Model 2 (n=2146)		Model 3 (n=1896)		Model 4 (n=1314)		Model 5 (n=2108)		Model 6 (n=1374)	
	OR (95% CI)	P value	aOR (95% CI)	P value	aOR (95% CI)	P value	aOR (95% CI)	P value	aOR (95% CI)	P value	aOR (95% CI)	P value	aOR (95% CI)	P value
Age (years), (n=2373)	1.03 (1.02-1.04)	**<0.001**	1.02 (1.02-1.03)	**<0.001**	1.03 (1.02-1.04)	**<0.001**	1.03 (1.02-1.04)	**<0.001**	1.01 (0.99-1.03)	0.197	1.02 (1.02-1.03)	**<0.001**	1.01 (0.99-1.03)	0.280
Male sex, (n=2373)	1.07 (0.90-1.26)	0.445												
Vaccine status, (n=2373)														
Unvaccinated	Reference		Reference		Reference		Reference		Reference		Reference		Reference	
One or two-dose Sinovac	0.43 (0.33-0.55)	**<0.001**	0.38 (0.30-0.49)	**<0.001**	0.34 (0.26-0.45)	**<0.001**	0.37 (0.27-0.50)	**<0.001**	0.27 (0.14-0.53)	**<0.001**	0.41 (0.14-0.53)	**<0.001**	0.31 (0.17-0.56)	**<0.001**
Two-dose plus one-dose booster Sinovac	0.27 (0.20-0.37)	**<0.001**	0.23 (0.17-0.32)	**<0.001**	0.20 (0.14-0.29)	**<0.001**	0.20 (0.14-0.29)	**<0.001**	0.14 (0.07-0.31)	**<0.001**	0.22 (0.07-0.31)	**<0.001**	0.14 (0.07-0.27)	**<0.001**
Two-dose plus two or more-dose booster Sinovac	0.14 (0.08-0.23)	**<0.001**	0.13 (0.08-0.21)	**<0.001**	0.10 (0.06-0.18)	**<0.001**	0.11 (0.06-0.19)	**<0.001**	0.03 (0.01-0.10)	**<0.001**	0.14 (0.01-0.10)	**<0.001**	0.05 (0.02-0.14)	**<0.001**
One or two-dose BNT162b2	0.10 (0.07-0.14)	**<0.001**	0.12 (0.08-0.17)	**<0.001**	0.10 (0.07-0.15)	**<0.001**	0.12 (0.08-0.18)	**<0.001**	0.13 (0.07-0.30)	**<0.001**	0.14 (0.07-0.30)	**<0.001**	0.17 (0.08-0.37)	**<0.001**
Two-dose plus one-dose booster BNT162b2	0.03 (0.01-0.05)	**<0.001**	0.03 (0.02-0.07)	**<0.001**	0.02 (0.01-0.05)	**<0.001**	0.01 (0.01-0.04)	**<0.001**	0.01 (0.01-0.03)	**<0.001**	0.02 (0.01-0.03)	**<0.001**	0.01 (0.01-0.05)	**<0.001**
Two-dose plus two or more-dose booster BNT162b2	0.05 (0.02-0.13)	**<0.001**	0.06 (0.02-0.16)	**<0.001**	0.04 (0.01-0.12)	**<0.001**	0.04 (0.01-0.12)	**<0.001**	0.01 (0.01-0.05)	**<0.001**	0.05 (0.01-0.05)	**<0.001**	0.01 (0.01-0.09)	**<0.001**
Two-dose Sinovac plus one-dose booster BNT162b2	0.23 (0.15-0.33)	**<0.001**	0.20 (0.14-0.30)	**<0.001**	0.18 (0.12-0.28)	**<0.001**	0.21 (0.13-0.33)	**<0.001**	0.26 (0.09-0.74)	**0.011**	0.23 (0.09-0.74)	**<0.001**	0.29 (0.11-0.72)	**0.008**
Two-dose Sinovac plus two or more-dose booster BNT162b2	0.10 (0.07-0.15)	**<0.001**	0.09 (0.06-0.14)	**<0.001**	0.08 (0.05-0.12)	**<0.001**	0.08 (0.05-0.13)	**<0.001**	0.06 (0.03-0.14)	**<0.001**	0.10 (0.03-0.14)	**<0.001**	0.08 (0.04-0.17)	**<0.001**
Blood group, (n=2146)														
0 Rh (-)	1.74 (1.01-3.01)	**0.048**			1.78 (0.95-3.35)	0.073	2.02 (1.01-4.02)	**0.047**	3.03 (0.62-14.74)	0.171				
0 Rh (+)	1.20 (0.83-1.75)	0.339			1.19 (0.77-1.84)	0.443	1.21 (0.78-1.92)	0.430	1.48 (0.60-3.63)	0.395				
A Rh (-)	1.19 (0.70-2.02)	0.526			1.35 (0.73-2.49)	0.335	1.26 (0.67-2.39)	0.473	0.78 (0.24-2.52)	0.680				
A Rh (+)	1.10 (0.77-1.57)	0.610			1.17 (0.77-1.78)	0.468	1.17 (0.75-1.84)	0.488	1.61 (0.67-3.78)	0.274				
B Rh (-)	1.34 (0.65-2.80)	0.430			1.77 (0.71-4.40)	0.221	1.95 (0.72-5.31)	0.189	n.a.	n.a.				
B Rh (+)	1.31 (0.87-1.96)	0.199			1.35 (0.84-2.17)	0.217	1.40 (0.84-2.32)	0.194	2.15 (0.79-5.80)	0.132				
AB Rh (-)	1.08 (0.35-3.40)	0.890			1.79 (0.40-7.96)	0.447	1.71 (0.38-7.67)	0.486	n.a.	n.a.				
AB Rh (+)	Reference				Reference		Reference		Reference					
CO-RADS, (n=2108)														
1	Reference						Reference		Reference		Reference		Reference	
2	1.88 (1.25-2.83)	**0.002**					2.01 (1.21-3.34)	**0.007**	1.99 (0.68-5.81)	0.210	1.97 (1.24-3.14)	**0.004**	2.45 (0.94-6.38)	0.067
3	2.03 (1.34-3.08)	**0.001**					2.25 (1.37-3.72)	**0.001**	1.53 (0.55-4.27)	0.418	1.98 (1.26-3.13)	**0.003**	1.72 (0.70-4.26)	0.239
4	2.32 (1.53-3.53)	**<0.001**					2.46 (1.47-4.10)	**0.001**	2.29 (0.77-6.79)	0.135	2.35 (1.48-3.75)	**<0.001**	2.69 (1.02-7.07)	**0.045**
5	2.31 (1.55-3.46)	**<0.001**					3.07 (1.87-5.03)	**<0.001**	3.50 (1.22-10.04)	**0.020**	2.62 (1.67-4.12)	**<0.001**	3.48 (1.38-8.78)	**0.008**
APACHE MPR (%), (n=1557)	1.07 (1.06-1.09)	**<0.001**							1.08 (1.06-1.10)	**<0.001**			1.07 (1.05-1.09)	**<0.001**

OR=Odds Ratio, aOR=adjusted odds ratio, CI=confidence interval, CO-RADS=COVID-19 Reporting and Data System, APACHE MPR=Acute Physiology and Chronic Health Evaluation Mortality Prediction Rate.

[1]Univariable Logistic Regression Model with the Enter Method was used.

[2]Multivariable Logistic Regression Models with the Enter Method were used. The variables shown in each column were included in the model.

[3]"n" in the first column shows the number of patients included in the univariable analyses.

admission needs. While effect modifications were observed for other variables across the six different models, the trends in the findings related to vaccination regimens remained consistent (Table 3).

DISCUSSION

Our study revealed that all vaccinated groups, including those receiving only one dose of the Sinovac or BNT162b2 regimen, had lower rates of ICU admission during hospitalisation and lower mortality rates than unvaccinated patients. Specifically, ICU admissions and mortality rates were lower in the BNT162b2 vaccinated group compared to the Sinovac-vaccinated group. Moreover, individuals who received a two-dose regimen of BNT162b2 followed by a booster dose (BNT162b2d), or two doses of BNT162b2 with two or more booster doses exhibited the lowest ICU admission rates (9.8% and 17.9%) and mortality rates (3.3% and 3.6%) compared to those who received two doses of Sinovac with one or more booster doses or two doses of Sinovac followed by one BNT162b2 booster dose. Conversely, the most significant reductions in mortality risk and ICU admission were observed in individuals who received BNT162b2 at both the initial and booster doses. Our study also identified age, CO-RADS stage, and APACHE MPR score as statistically significant determinants of mortality and ICU admission.

Individuals of different age groups can be affected by this highly transmissible disease. However, people aged ≥60 years and those with comorbidities such as diabetes, cardiovascular disease, obesity, and chronic lung disease are particularly vulnerable to severe clinical illness and mortality [10]. Additionally, numerous studies indicate that this disease tends to be more severe in men and is associated with higher mortality rates [11, 12]. Our study cohort had a higher median age, and 56.9% were male and considered at high risk for severe infection [7].

The virulence of the virus and the host immune response determine the severity of SARS-CoV-2-induced illness. At the same time, a controlled immune response in mild infections leads to viral clearance, an excessive and uncontrolled immune response, characterised by aberrant cytokine and chemokine activity, results in the infiltration of inflammatory cells, destruction of the respiratory epithelial layer, and consequently respiratory failure in severe cases [13, 14]. Elevated levels of CRP, ferritin, D-dimer, and troponin have been identified as risk factors associated with the severity of the clinical course in patients with COVID-19 [7, 15].

Like other respiratory RNA viruses, SARS-CoV-2 undergoes various mutations as it adapts to its new host, leading to differences from the original strain [5]. Many studies have demonstrated that mRNA and inactivated vaccines are effective against several variants, including Alpha, Beta, Gamma, and Delta, albeit in varying degrees [16, 17]. A population-based study encompassing over seven thousand COVID-19-associated hospitalisations of adults 65 years and older showed that 75% of the cases were unvaccinated, 12% were partially vaccinated, and only 5% were fully vaccinated [18]. In the present study, 50% of the patients were unvaccinated, 40% were partially vaccinated, and 10% were fully vaccinated. A vaccine schedule consisting of two doses plus one or two booster doses was considered fully vaccinated in patients with advanced age and additional comorbidities [7, 19].
A multicentre historical control study investigating the severity and outcomes of COVID-19 among vaccinated compared to unvaccinated patients found that unvaccinated or partially vaccinated individuals experienced more in-hospital complications, severe disease, and death than fully vaccinated individuals [7]. Unvaccinated individuals more frequently show abnormal findings on chest imaging, with more widespread involvement observed at the onset of COVID-19 infection [7, 20]. In addition, these patients were more likely to require ICU admission and experienced a higher death rate than vaccinated individuals. Our experience with the current pandemic indicates that advanced age is a significant risk factor for severe clinical course and disease-related mortality [21, 22].

The severity of the current pandemic was lowest in the fully vaccinated group and highest in the unvaccinated group, and it was between these two groups in partially unvaccinated people, remarking the protective role of mRNA vaccines against severe COVID-19 disease [7]. Heterologous prime-booster and a third dose vaccination induced a robust humoral response in all adult age groups [16, 23]. Our study underscores the importance of delineating effective vaccines and

optimal vaccination schedules, particularly in elderly cohorts. Our findings highlight the potential need for tailored full-dose vaccination protocols based on the specific vaccine in specific patient groups.

Limitations

Since our study was a single-centre hospital-based retrospective analysis, the generalisability of the study is limited. Another potential limitation of the study is the possibility of data source errors, as the study data were obtained from electronic medical records. Some patient characteristics, such as comorbidities, and their potential impact on the outcomes were not assessed. Additionally, the treatment modalities administered before admission and during hospitalisation with discussions on their accessibility and efficacy were not detailed. Readers should keep these limitations in mind when interpreting the results.

CONCLUSION

As a result, a two-dose initial plus one or more-dose booster BNT162b2 regimen showed the highest effectiveness in reducing the risks of mortality and ICU admission compared to partially vaccinated regimens and the unvaccinated group. Our study group consisted of very elderly patients who presented with a severe clinical course, which may explain the high mortality rates.

Ethics Approval and Consent to Participate

This study was approved by the University of Health Sciences, Bursa Yüksek Ihtisas Training and Research Hospital, Clinical Research Ethics Committee (Decision no. 2011-KAEK-25 2022/04-18, date: 20.04.2022). Informed written consent was not obtained because of the study's retrospective design.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: AT, HÖ, AAV; Study Design: AT, HÖ, HB; Supervision: AT, HÖ, NK; Funding: AT, AAV, NK; Materials: AT, AAV, HB; Data Collection and/or Processing: AT, AAV, NK; Statistical Analysis and/or Data Interpretation: AT, HÖ; Literature Review: AT, AAV, HB; Manuscript Preparation: AT, HÖ, AAV, NK, HB and Critical Review: AT, HÖ, AAV, NK, HB.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Fundings

This study was supported by TUBİTAK-BİDEB 2211-A/2020 National PhD Scholarship Program.

Acknowledgements

The authors sincerely thank their colleagues for their unwavering support and collaboration. The authors also thank the participants of this study for their valuable contributions.

We would like to thank Hicran Aslan, the Intensive Care Unit Manager at Çekirge State Hospital, who helped us collect data day and night despite the closed area of our study.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) - China, 2020. China CDC Wkly. 2020.2(8):113-122.

2. Munster VJ, Koopmans M, van Doremalen N, van Riel D, de Wit E. A Novel Coronavirus Emerging in China-Key Questions for Impact Assessment. N Engl J Med. 2020.382(8):692-694. doi: 10.1056/NEJMp2000929.
3. Stöhr K. Influenza--WHO cares. Lancet Infect Dis. 2002.2(9):517. doi: 10.1016/s1473-3099(02)00366-3.
4. Girard MP, Cherian T, Pervikov Y, Kieny MP. A review of vaccine research and development: human acute respiratory infections. Vaccine. 2005.23(50):5708-5724. doi: 10.1016/j.vaccine.2005.07.046.
5. Lotfi H, Mazar MG, Ei NMH, Fahim M, Yazdi NS. Vaccination is the most effective and best way to avoid the disease of COVID-19. Immun Inflamm Dis. 2023;11(8):e946. doi: 10.1002/iid3.946.
6. Kitabatake M, Ouji-Sageshima N, Sonobe S, et al. Transition of Antibody Titers after SARS-CoV-2 mRNA Vaccination in Japanese Healthcare Workers. Jpn J Infect Dis. 2023;76(1):72-76. doi: 10.7883/yoken.JJID.2022.041.
7. Alshanqeeti S, Szpunar S, Anne P, Saravolatz L, Bhargava A. Epidemiology, clinical features and outcomes of hospitalized patients with COVID-19 by vaccination status: a multicenter historical cohort study. Virol J. 2024;21(1):71. doi: 10.1186/s12985-024-02325-x.
8. Prokop M, van Everdingen W, van Rees Vellinga T, et al; COVID-19 Standardized Reporting Working Group of the Dutch Radiological Society. CO-RADS: A Categorical CT Assessment Scheme for Patients Suspected of Having COVID-19-Definition and Evaluation. Radiology. 2020;296(2):E97-104. doi: 10.1148/radiol.2020201473.
9. Bajaj K, Rathee P, Jain P, Panwar VR. Comparison of the Reliability of Anatomic Landmarks based on PA Cephalometric Radiographs and 3D CT Scans in Patients with Facial Asymmetry. Int J Clin Pediatr Dent. 2011;4(3):213-223. doi: 10.5005/jp-journals-10005-1112.
10. Stokes EK, Zambrano LD, Anderson KN, et al. Coronavirus Disease 2019 Case Surveillance - United States, January 22-May 30, 2020. MMWR Morb Mortal Wkly Rep. 2020;69(24):759-765. doi: 10.15585/mmwr.mm6924e2.
11. Jin JM, Bai P, He W, et al. Gender Differences in Patients With COVID-19: Focus on Severity and Mortality. Front Public Health. 2020;8:152. doi: 10.3389/fpubh.2020.00152.
12. Finelli L, Gupta V, Petigara T, Yu K, Bauer KA, Puzniak LA. Mortality Among US Patients Hospitalized With SARS-CoV-2 Infection in 2020. JAMA Netw Open. 2021;4(4):e216556. doi: 10.1001/jamanetworkopen.2021.6556.
13. Akkiz H. The Biological Functions and Clinical Significance of SARS-CoV-2 Variants of Corcern. Front Med (Lausanne). 2022;9:849217. doi: 10.3389/fmed.2022.849217.
14. Yuan Y, Jiao B, Qu L, Yang D, Liu R. The development of COVID-19 treatment. Front Immunol. 2023;14:1125246. doi: 10.3389/fimmu.2023.1125246.
15. Bhargava A, Sharma M, Riederer K, Fukushima EA, Szpunar SM, Saravolatz L. Risk Factors for In-hospital Mortality from Coronavirus Disease 2019 Infection Among Black Patients-An Urban Center Experience. Clin Infect Dis. 2021;73(11):e4005-11. doi: 10.1093/cid/ciaa1468.
16. Fiolet T, Kherabi Y, MacDonald CJ, Ghosn J, Peiffer-Smadja N. Comparing COVID-19 vaccines for their characteristics, efficacy and effectiveness against SARS-CoV-2 and variants of concern: a narrative review. Clin Microbiol Infect. 2022;28(2):202-221. doi: 10.1016/j.cmi.2021.10.005.
17. Yildirim S, Kirakli C, Ozdemir Y, et al. Impact of vaccination on ICU admissions of hospitalized COVID-19 patients in a country with a heterologous vaccine policy. J Infect Dev Ctries. 2024;18(4):513-519. doi: 10.3855/jidc.18342.
18. Moline HL, Whitaker M, Deng L, et al. Effectiveness of COVID-19 Vaccines in Preventing Hospitalization Among Adults Aged ≥65 Years - COVID-NET, 13 States, February-April 2021. MMWR Morb Mortal Wkly Rep. 2021;70(32):1088-1093. doi: 10.15585/mmwr.mm7032e3.
19. Gholinataj Jelodar M, Mirzaei S, Saghafi F, et al. Impact of vaccination status on clinical outcomes of hospitalized COVID-19 patients. BMC Infect Dis. 2024;24(1):254. doi: 10.1186/s12879-024-09139-w.
20. Tenforde MW, Self WH, Adams K, et al; Influenza and Other Viruses in the Acutely Ill (IVY) Network. Association Between mRNA Vaccination and COVID-19 Hospitalization and Disease Severity. JAMA. 2021;326(20):2043-2054. doi: 10.1001/jama.2021.19499.
21. Liu W, Yang C, Liao YG, et al. Risk factors for COVID-19 progression and mortality in hospitalized patients without pre-existing comorbidities. J Infect Public Health. 2022;15(1):13-20. doi: 10.1016/j.jiph.2021.11.012.
22. Gul F, Kasapoglu US, Sabaz MS, et al. The Impact of CoronaVac Vaccination on 28-day Mortality Rate of Critically Ill Patients with COVID-19 in Turkiye. Balkan Med J. 2023;40(6):435-444. doi: 10.4274/balkanmedj.galenos.2023.2023-6-90.
23. Soheili M, Khateri S, Moradpour F, et al. The efficacy and effectiveness of COVID-19 vaccines around the world: a mini-review and meta-analysis. Ann Clin Microbiol Antimicrob. 2023;22(1):42. doi: 10.1186/s12941-023-00594-y.

The European Research Journal 2025;11(6):1176-1183 Original Article

DOI: https://doi.org/10.18621/eurj.1777370 *General Surgery*

Alteration in TSH levels during Ramadan in non fasting levothyroxine treated patients living with fasting family members

Alper Aytekin[1], İpek Köroğlu[2], Zeynel Abidin Sayıner[3], Mehmet Karaçalı[4], Elif Melis Baloğlu Akyol[2], Ersin Akarsu[2]

[1]Department of General Surgery, Gaziantep University School of Medicine, Gaziantep, Türkiye; [2]Department of Endocrinology and Metabolism, Gaziantep University School of Medicine, Gaziantep, Türkiye; [3]Department of Endocrinology and Metabolism, Sanko University School of Medicine, Gaziantep, Türkiye; [4]Department of Internal Medicine, Abdulkadir Yüksel State Hospital, Gaziantep, Türkiye

ABSTRACT

Objectives: During Ramadan, lifestyle changes may affect thyroid function not only in fasting individuals but also in those who do not fast. This study examined the impact of Ramadan-related behavioral shifts on thyroid-stimulating hormone (TSH) levels in patients receiving stable levothyroxine therapy.

Methods: This observational before-and-after study, conducted between January 2023 and January 2024, included 116 hypothyroid patients on stable levothyroxine (LT4) therapy. Serum TSH, free thyroxine (fT4), and free triiodothyronine (fT3) levels were measured one month before and within one week after Ramadan. Patients were categorized as fasting or non-fasting, and completed a structured questionnaire regarding medication timing, fasting behavior, and lifestyle changes. Linear regression analysis was used to identify predictors of post-Ramadan TSH changes.

Results: TSH levels significantly increased after Ramadan in the overall cohort (mean±SD: 3.03±3.24 vs. 4.77±7.83 mU/L; P=0.028), while fT3 and fT4 levels remained unchanged (P=0.14 and P=0.72, respectively). Among fasting patients (n=72), TSH rose from 3.26±3.26 to 4.80±8.58 mU/L (P=0.47), whereas non-fasting patients (n=44) showed a significant increase from 2.65±3.21 to 4.73±6.51 mU/L (P=0.006). Polypharmacy was associated with increased odds of post-Ramadan TSH elevation (OR=2.67, 95% CI: 1.25–5.73, P=0.01). Among non-fasting patients, those who reported changes in sleep or meal patterns during Ramadan (n=28) experienced a significantly higher increase in TSH compared to those without such changes (ΔTSH: 2.47±3.88 vs. 0.81±2.12 mU/L, P=0.02).

Conclusions: Patients receiving LT4 therapy who do not fast during Ramadan may still exhibit significant TSH variability, likely due to behavioral and circadian disruptions. These findings emphasize the importance of proactive monitoring and targeted education for all hypothyroid patients during Ramadan, irrespective of fasting status.

Keywords: Fasting, hypothyroidism, levothyroxine, thyroid function test, thyroidectomy

Received: September 3, 2025 **Accepted:** October 16, 2025 **Available Online:** October 22, 2025 **Published:** November 4, 2025

How to cite this article: Aytekin A, Köroğlu İ, Sayiner ZA, Karaçalı M, Baloğlu Akyol EM, Akarsu E. Alteration in TSH levels during Ramadan in non fasting levothyroxine treated patients living with fasting family members. Eur Res J. 2025;11(6):1176-1183. doi: 10.18621/eurj.1777370

Corresponding author: Alper Aytekin, MD., Assist. Prof., Phone: +90 342 360 07 53, E-mail: aytekinalper83@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

Levothyroxine (LT4) is still the standard of care in the treatment of hypothyroidism and necessitates rigorous compliance with timing protocols to ensure optimal bioavailability. It is generally recommended to be consumed 30–60 minutes before breakfast or at bedtime, ensuring at least three hours have elapsed since the last meal [1]. Differences from this routine may reduce absorption, especially when meal timing and sleep cycles are altered. Ramadan is characterized by a daily fast from dawn to sunset. Religious practices and social norms often cause people to change their dietary habits, cycles of sleep and wakefulness, and medication regimens during this month [2-4]. These changes could cause particular difficulties for patients receiving long-term therapies like LT4.

Food intake is restricted to specific times (Iftar and Suhoor), and patients frequently adjust their medication schedules accordingly, which might affect therapeutic consistency [3]. Fasting has been linked to a number of physiological changes, such as changes in circadian rhythm, changes in melatonin secretion, and more activity at night. These changes have the potential to affect the hypothalamic pituitary-thyroid axis. Furthermore, psychological stress, disrupted sleep architecture, and reduced medication adherence during Ramadan may all have an impact on thyroid hormone dynamics [5-7]. Previous research examining the effects of Ramadan on thyroid function in LT4-treated individuals showed inconsistent results. Some studies indicate significant elevations in serum TSH following Ramadan, frequently ascribed to inappropriate dosing intervals or food-drug interactions [8-9]. Others argue that LT4's extended half-life and physiological buffering mechanisms may safeguard against short-term variations [7-11]. Significantly, the majority of current research concentrates solely on fasting individuals, creating a void in comprehension regarding the impact of Ramadan-related lifestyle disruptions on those who do not fast but live in fasting households [12]. These individuals might experience indirect behavioral and environmental factors, such as disrupted sleep, delayed meals, or irregular medication adherence, which could affect thyroid hormone levels.

This study aimed to assess the impact of Ramadan-related changes in lifestyle on TSH levels in patients undergoing stable LT4 replacement, specifically focusing on those not fasting.

METHODS

Study Design and Setting

This was a prospective, observational, pre-post (before-after), non-randomized cohort study conducted between January 2023 and January 2024 in general surgery and endocrinology outpatient clinics. The study aimed to evaluate the effects of thyroid hormone replacement intake variations on the management of hypothyroidism in patients with Ramadan fasting. Ethical approval was obtained from the relevant institutional review board, and all participants provided written informed consent before participation.

Study Population

The study included 116 patients receiving levothyroxine treatment, regardless of fasting status. The study included 72 fasting patients and 44 non-fasting patients. The criteria for inclusion were as follows: Individuals aged 18 and above, diagnosed with primary hypothyroidism and undergoing stable levothyroxine therapy for a minimum of six months, who are willing to participate and complete the questionnaire. Individuals with a history of central hypothyroidism, malignancy, or disorders impacting thyroid metabolism were excluded. Participants were categorized into two groups: individuals who fasted during Ramadan and those who did not.

Data Collection

Data were gathered utilizing a structured 17-item questionnaire aimed at evaluating sociodemographic characteristics, fasting behaviors, medication compliance, and dietary patterns. A supplementary questionnaire was administered to non-fasting patients to assess changes in sleep and dietary patterns, including whether they woke for Suhoor, had fasting family members, or altered eating, sleep schedules. Binary categorical variables were created based on responses and used for subgroup comparisons of TSH change. This questionnaire was adapted from previously validated Ramadan-related behavioral studies. [2, 3]. Patients' medication timing was recorded and categorized as "before Suhoor," "before Iftar," "after Iftar," or "at variable times." Individuals who reported irregular or inconsistent levothyroxine intake (n=4) were excluded from statistical analysis to ensure reliability. No randomization or intervention was performed.

Table 1. **Sociodemographic and clinical characteristics of patients**

Survey parameters	Options	n	%	Survey parameters	Options	n	%
Gender	Female	101	87.1	**Do you fast?**	Yes	72	62
	Male	15	12.9		No	44	37.9
Age group	30 years old or younger	20	17.2	**How many years have you been using levothyroxine treatment?**	0-5 years	50	43.1
	31-40 years	21	18.q		5-10 years	33	28.4
	41-50 years	32	27.6		10-15 years	22	27.6
	51 years old or older	43	37.1		>15 years	11	18.9
Education level	Illiterate	3	2.6	**When do you take your medication while fasting?**	On an empty stomach at 'Suhoor'	58	80.5
	Primary school graduate	50	43.1		1-2 hours after 'Iftar'	1	1.3
	Middle school graduate	9	7.8		2-4 hours after 'Iftar'	2	2.7
	High school graduate	28	24.1		On an empty stomach before 'Iftar'	10	13.8
	University graduate	24	20.7		At any time	1	1.3
	Master's degree or higher	2	1.7				
Marital status	Married	97	83.6	**Do you attend regular check-ups?**	Yes	105	90.5
	Single	19	16.4		No	11	9.4
Residential Area	Village	10	8.6	**Reason for using levothyroxine (doctor's question)**	Primary hypothyroidism	60	51.7
	District	16	13.8		Total thyroidectomy	56	48.2
	City center	90	77.6				
Employment status	Employed	21	18.1	**Number of medications used**	1	50	43.1
	Unemployed	95	81.9		1 to 5	48	41.3
					>5	18	15.5
Economic status	Low	21	18.1				
	Moderate	84	72.4				
	High	11	9.5				

Assessment of Thyroid Function

Thyroid function tests, comprising serum TSH, free T4 (fT4), and free T3 (fT3) levels, were assessed utilizing standard chemiluminescent immunoassay techniques. Blood samples were obtained one month before and again within one week following Ramadan. The primary outcome was altered serum TSH levels during Ramadan fasting compared with baseline measurements.

Medication Adherence and Timing

Participants were questioned about their compliance with medication and the exact period of levothyroxine administration during Ramadan. The administration timing of levothyroxine was classified as follows: Medication timing was recorded a priori and classified as "before Suhoor," "before Iftar," "after Iftar," or "variable." Individuals reporting irregular or inconsistent LT4 intake were excluded from analyses.

Ethical Considerations

This study was approved by the Gaziantep University Clinical Research Ethics Committee (Decision No: 2023/118; date: 07.06.2023). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Written informed consent was obtained from all participants.

Statistical Analysis

The variables were analyzed using SPSS 25.0 (IBM Corporation, Armonk, New York, USA) and Medcalc 14 (Acacialaan 22, B-8400 Ostend, Belgium) software. The adherence of the data to a normal distribution was assessed using the Shapiro-Wilk test, while the homogeneity of variance was examined through Levene's test. The Monte Carlo method, Mann-Whitney U test, Bootstrap results, and Independent Samples T Test were employed to compare two independent groups based on quantitative variables. A Linear Mixed Model (LMM) analysis was utilized to assess the factors influencing the variation in TSH levels in regression analysis. A P value of less than 0.05 was considered significant.

RESULTS

A total of 116 patients were included in the analysis, of whom 101 (87.1%) were female and 15 (12.9%) were male. Seventy-two patients (62.0%) reported fasting during Ramadan, while 44 (37.9%) did not. Educational status was predominantly at the primary

Table 2. Comparison of thyroid function tests before and after Ramadan by fasting status and polypharmacy

		TSH (mU/L) (0.34-5.6)	fT4 (ng/dL) (0.61-1.45)	fT3 (ng/L) (2.5-4.2)	P value (TSH/fT4/fT3)
All patients (n=116)	Before Ramadan	3.03±3.24)	1.02±0.29	3.26±0.57	**0.028***/0.72/0.14
	After Ramadan	4.77±7.83	1±0.22	3.35±0.59	
Fasting patients (n=72)	Before Ramadan	3.26±3.26	0.98±0.22	3.27±0.44	0.47/0.43/0.35
	After Ramadan	4.80±8.58	0.99±0.22	3.32±0.53	
Non-fasting patients (n=44)	Before Ramadan	2.65±3.21	1.09±0.36	3.25±0.75)	**0.006***/0.59/0.30
	After Ramadan	4.73±6.51	1.03±0.23	3.3±0.68	
Single mediacation (n=50)	Before Ramadan	3.30±3.35	0.96±0.22	3.22±0.42	0.51/0.24/**0.016***
	After Ramadan	4.70±8.65	1±0.24	3.43±0.58	
Multiple medications (n=66)	Before Ramadan	2.82±3.16	1.06±0.33	3.30±0.67	**0.016***/0.57/0.82
	After Ramadan	4.82±7.20	1.00±0.21	3.28±0.59	

Data are shown as mean±standard deviation. TSH=Thyroid-Stimulating Hormone, fT4=Free Thyroxine, fT3=Free Triiodothyronine

*P<0.05 is considered significant

school level (43.1%), and a considerable number were either unemployed (81.9%) or residing in urban areas (77.6%). Of clinical relevance, nearly half of the patients were being treated following total thyroidectomy (48.2%), indicating a significant post-surgical hypothyroidism cohort. Among the patients, 62.0% (n=72) were fasting and 37.9% (n=44) were not fasting. In terms of the duration of levothyroxine treatment, 43.1% had been using levothyroxine for 0-5 years, 28.4% for 5-10 years, 27.6% for 10-15 years and 18.9% for more than 15 years. In terms of the time of use, 80.6% took the drug on an empty stomach at suhoor, 1.4% 1-2 hours after iftar, 2.8% 2-4 hours after iftar, 13.9% on an empty stomach before iftar, and 1.4% at any time. 90.5% of the participants stated that they had regular check-ups during Ramadan (n=105), while 9.5% stated that they did not have regular check-ups (n=11). In the section filled by the physician, 51.7% of the patients were using levothyroxine due to primary hypothyroidism, and 48.3% due to total thyroidectomy. When the number of drugs used was analyzed, 43.1% were using a single drug, 41.5% were using 1-5 drugs, and 15.5% were using more than 5 drugs (Table 1).

Across the entire study population, a statistically significant increase in serum TSH concentrations was observed following Ramadan (pre-Ramadan: 3.03±3.24 mU/L vs. post-Ramadan: 4.77±7.83 mU/L; P=0.028). In contrast, mean serum fT4 and fT3 levels did not differ significantly across the same time interval (P=0.72 and P=0.14, respectively). Subgroup analysis revealed that non-fasting individuals exhibited a statistically significant increase in TSH levels after Ramadan (2.65±3.21 vs. 4.73±6.51 mU/L; P=0.006), whereas the change observed among fasting participants did not reach statistical significance (3.263.26 vs. 4.80±8.58 mU/L; P=0.47). Neither group demonstrated significant alterations in fT4 or fT3 concentrations. The distribution of patients' medication timing was recorded and categorized as "before Suhoor," "before Iftar," "after Iftar," or "at variable times." Individuals who reported irregular or inconsistent levothyroxine intake (n=4) were excluded from statistical analysis to ensure reliabilition of medication timing categories ("before Suhoor," "before Iftar," "after Iftar," "variable") was statistically similar between fasting and non-fasting groups suggesting that medication timing alone did not account for observed group differences (P=0.41). Patients receiving more than one medication (n=66) experienced a statistically

Table 3. Analysis of TSH changes by fasting and medication use

Fasting status	Number of medications used	Variables	Before Ramadan	After Ramadan	P value
Fasting patients (n=72)	Single medication (n=38)	TSH (mU/L)	3.51±3.02	4.70±9.69	0.39
		fT4 (ng/dL)	0.97±0.22	1.01±0.25	0.22
		fT3 (ng/L)	3.25±0.44	3.45±0.62	0.02
	Multiple medications (n=34)	TSH (mU/L)	2.98±3.53	4.90±7.28	**0.06**
		fT4 (ng/dL)	1.00±0.23	0.97±0.18	0.79
		fT3 (ng/L)	3.29±0.45	3.17±0.33	0.33
Non-fasting patients (n=44)	Single medication (n=12)	TSH (mU/L)	2.66±4.33	4.70±4.25	**0.03***
		fT4 (ng/dL)	0.96±0.21	0.99±0.23	0.84
		fT3 (ng/L)	3.12±0.35	3.32±0.45	0.43
	Multiple medications (n=32)	TSH (mU/L)	2.65±2.76	4.74±7.24	0.10
		fT4 (ng/dL)	1.13±0.40	1.04±0.23	0.58
		fT3 (ng/L)	3.31±0.85	3.41±0.75	0.48

Data are shown as mean±standard deviation. TSH=Thyroid-Stimulating Hormone, fT4=Free Thyroxine, fT3=Free Triiodothyronine

*P<0.05 is considered significant

significant elevation in TSH post-Ramadan (2.823.16 vs. 4.82±7.20 mU/L; P=0.016), in contrast to those on monotherapy with levothyroxine (n=50), where the TSH change was not statistically significant (P=0.51). Notably, a modest yet significant increase in fT3 levels was observed among monotherapy users (3.22±0.42 vs. 3.43±0.58 ng/L; P=0.016) (Table 2).

In a stratified analysis based on both fasting status and polypharmacy, only the subgroup of non-fasting patients receiving levothyroxine monotherapy exhibited a significant post-Ramadan increase in TSH (2.66±4.33 vs. 4.70 ± 4.25 mU/L; P=0.03). Although numerically elevated TSH values were also observed in other subgroups, including fasting individuals and those on multiple medications, these differences did not reach statistical significance (Table 3).

DISCUSSION

In this study, a significant increase in serum TSH was observed in all patients receiving LT4 replacement after Ramadan, but the most striking finding was that this increase occurred predominantly in individuals who did not fast and that most of these individuals shared behavioural and environmental changes specific to Ramadan. These results may provide important insights into the interaction between behavioural rhythms, medication compliance, and thyroid homeostasis. Studies have been conducted to investigate the effects of fasting during Ramadan on levothyroxine absorption and TSH dynamics, but the findings are inconsistent. Some suggest that changes in timing and taking the medication with food may impair levothyroxine bioavailability and lead to elevated TSH levels, while others propose that levothyroxine's long half-life and peripheral conversion to T3 may act as a buffer against short-term fluctuations [3, 8-10]. Our findings appear consistent with recent evidence suggesting that circadian rhythm disruption, including changes in sleep-wake cycles, late-night eating, and irregular meal times, may independently affect hypothalamic-pituitary-thyroid (HPT) axis regulation [4, 5]. In this cohort, non-fasting individuals frequently reported shared household schedules, late meals, and sleep disturbances; these factors may have contributed to changes in TSH secretion through alterations in hypothalamic TRH stimulation or impaired medication adherence. Additionally, the significant association between polypharmacy and increased TSH further supports the hypothesis that concomitant medications may impair LT4 absorption [13]. These findings may have clinical significance, as individuals may require personalised counselling and timing adjustments during Ramadan to maintain euthyroidism, regardless of whether they fast or not.

Unlike some previous studies, in our study, the timing of LT4 administration (before suhoor, after iftar, or before bedtime) did not significantly affect TSH dynamics in our population [2]. This suggests that compliance and behavioural routine changes may be more important than dosing time. For example, Mahzari *et al*. [11] also found that TSH increased significantly after Ramadan, independent of administration timing. These observations may highlight that lifestyle changes could be multifactorial determinants of thyroid hormone regulation.

A novel contribution of this study is its emphasis on non-fasting individuals, who are often underrepresented in research on Ramadan and thyroid hormone replacement success. While most previous studies have focused solely on fasting individuals, our findings suggest that even non-fasting individuals may exhibit significant biochemical changes due to shared social and environmental factors. This concept is also supported by a recent observational study by Al-Qahtani *et al*. [12].

There was a significant increase in TSH in individuals who did not fast. This situation can be explained by indirect behavioral and environmental effects among individuals who do not fast living in households where fasting occurs. Although they did not fast themselves, these patients frequently reported sleep disturbances and delayed meal timing, which may be due to circadian misalignment without the adaptive hormonal mechanisms found in fasting individuals. This model may reflect environmental adaptation effects rather than direct fasting physiology. Similar circadian effects on hypothalamic TRH stimulation and TSH dynamics have been previously described in the literature [5].

Strengths and Limitations

This study has several strengths. It is one of the

few that evaluated thyroid hormone changes during Ramadan in both fasting and non-fasting individuals living in the same environment. This approach provides a realistic view of how behavioral and lifestyle factors affect thyroid function. The prospective before–after design and standardized laboratory timing improve the reliability of the results. In addition, medication adherence and timing were carefully documented, and inconsistent users were excluded to ensure data accuracy.

This study has several limitations. First, its observational design limits causal inferences. Second, behavioural data such as compliance, sleep patterns, and food intake are self-reported, which may introduce potential recall bias. Third, we did not assess serum markers or chronobiological parameters of LT4 absorption, which could provide mechanistic insights. Finally, while our sample size was sufficient for primary analyses, some subgroup comparisons were underpowered. Additionally, although patients with inconsistent LT4 intake were excluded, residual confounding due to unmeasured variations in adherence or timing cannot be completely ruled out.

CONCLUSION

Lifestyle changes associated with Ramadan may affect thyroid function in patients taking LT4; however, such changes are heterogeneous and not universal. In this study, the most pronounced TSH variability was observed among non-fasting individuals who reported sleep and meal-schedule disruptions. Considering behavioral, pharmacological, and environmental factors that may disrupt treatment stability regardless of fasting during Ramadan, it may be prudent to recommend personalized follow-up for all patients with hypothyroidism receiving LT4 replacement.

Ethics Approval and Consent to Participate

This study was approved by the Gaziantep University Clinical Research Ethics Committee (Decision No: 2023/118; date: 07.06.2023). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. Written informed consent was obtained from all participants.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: AA, ZAS, İK, EMBA; Study Design: ZAS, AA, İK; Supervision: AA, ZAS, EA; Funding: N/A; Materials: ZAS, AA, İK; Data Collection and/or Processing: İK, EMBA, MK; Statistical Analysis and/or Data Interpretation: AA, ZAS, İK, MK; Literature Review: AA, ZAS, EA; Manuscript Preparation: ZAS, AA, İK; and Critical Review: ZAS, EA, AA.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Jonklaas J, Bianco AC, Bauer AJ, et al. American Thyroid Association Task Force on Thyroid Hormone Replacement. Guide-

lines for the treatment of hypothyroidism: prepared by the american thyroid association task force on thyroid hormone replacement. Thyroid. 2014;24(12):1670-1751. doi: 10.1089/thy.2014.0028.
2. Sheikh A, Mawani M, Mahar SA. Impact of Ramadan fasting on thyroid status and quality of life in patients with primary hypothyroidism: a prospective cohort study from Karachi, Pakistan. Endocr Pract. 2018;24(10):882-888. doi: 10.4158/EP-2018-0038.
3. El-Kaissi S, Dajani R, Lee-St John TJ, et al. Impact of Lifestyle Changes During Ramadan on Thyroid Function Tests in Hypothyroid Patients Taking Levothyroxine. Endocr Pract. 2020;26(7):748-753. doi: 10.4158/EP-2019-0505.
4. Caron P, Grunenwald S, Persani L, Borson-Chazot F, Leroy R, Duntas L. Factors influencing the levothyroxine dose in the hormone replacement therapy of primary hypothyroidism in adults. Rev Endocr Metab Disord. 2022;23(3):463-483. doi: 10.1007/s11154-021-09691-9.
5. Ikegami K, Refetoff S, Van Cauter E, Yoshimura T. Interconnection between circadian clocks and thyroid function. Nat Rev Endocrinol. 2019;15(10):590-600. doi: 10.1038/s41574-019-0237-z.
6. Nazem MR, Bastanhagh E, Emami A, Hedayati M, Samimi S, Karami M. The relationship between thyroid function tests and sleep quality: cross-sectional study. Sleep Sci. 2021;14(3):196-200. doi: 10.5935/1984-0063.20200050.
7. Belal MM, Youssef AR, Baker H, et al. Effect of Ramadan fasting on thyroid functions in hypothyroid patients taking levothyroxine: a systematic review and meta-analysis. Ir J Med Sci. 2024;193(2):741-753. doi: 10.1007/s11845-023-03526-z.
8. Alkaf B, Siddiqui M, Ali T, et al. Ramadan fasting and changes in thyroid function in hypothyroidism: identifying patients at risk. Thyroid. 2022;32(4):368-375. doi: 10.1089/thy.2021.0512.
9. Elsherbiny TM. Impact of fasting on thyrotropin and thyroid status during Ramadan in 292 previously well-controlled hypothyroid patients. IFTAR study. Endocrine. 2023;79(3):484-490. doi: 10.1007/s12020-022-03242-1.
10. Dellal FD, Ogmen B, Ozdemir D, et al. Effect of Ramadan fasting on thyroid hormone levels in patients on levothyroxine treatment. J Coll Physicians Surg Pak. 2020;30(10):1009-1014. doi: 10.29271/jcpsp.2020.10.1009.
11. Mahzari M, Al Remthi F, Ajwah I, et al. Levothyroxine timing during Ramadan: a randomized clinical trial. Int J Endocrinol. 2023;2023:2565031. doi: 10.1155/2023/2565031.
12. Al-Qahtani KM, Aldeeri IA, Alshaibi AM, et al. Optimal timing of thyroid hormone replacement during Ramadan fasting: a randomized controlled trial in patients with prior total thyroidectomy. Thyroid. 2022;32(9):1029-1036. doi: 10.1089/thy.2022.0110.
13. Grindrod K, Alsabbagh W. Managing medications during Ramadan fasting. Can Pharm J (Ott). 2017 Apr 3;150(3):146-9. doi: 10.1177/1715163517700840.

The European Research Journal 2025;11(6):1184-1191 Original Article

DOI: https://doi.org/10.18621/eurj.1651488 *Neurology*

The role of medial plantar nerve conduction studies in the diagnosis of diabetic polyneuropathy: A comparative analysis with sural nerve

Gülçin Koç Yamanyar[1], Hüsniye Aslan[2]

[1]Department of Neurology, Bursa City Hospital, Bursa, Türkiye; [2]Department of Neurology, Health Sciences University, Istanbul Training and Research Hospital, İstanbul, Türkiye

ABSTRACT

Objectives: Diabetic neuropathy, significantly impacts patients' quality of life and may lead to severe morbidity in the long term. Early detection of diabetic neuropathy is crucial in preventing or delaying irreversible damage. In this study, we aimed to evaluate the role of medial plantar nerve conduction studies in the early diagnosis of diabetic polyneuropathy and to determine which of the examined nerve conduction studies demonstrates higher sensitivity.

Methods: Sixty patients with suspected diabetic polyneuropathy and 30 healthy controls were included. Diabetic neuropathy symptoms were assessed using the Diabetic Neuropathy Symptom (DNS) score. Sensory and motor nerve conduction studies, including median, ulnar, posterior tibial, medial plantar, and sural nerves, were performed.

Results: Sensory response amplitudes of both medial plantar and sural nerves were significantly lower in patients compared to controls. Abnormal sensory responses were detected in 23 (38.33%) patients for the sural nerve and 39 (65%) for the medial plantar nerve.

Conclusions: Both nerve conduction studies are valuable in diagnosing diabetic polyneuropathy, but medial plantar nerve conduction studies demonstrated higher sensitivity. Including medial plantar nerve assessments in routine evaluations may improve diagnostic accuracy.

Keywords: Diabetic polyneuropathy, medial plantar nerve, sural nerve

Diabetic neuropathy, a common and chronic complication of Diabetes Mellitus (DM), can significantly impact patients' quality of life and lead to serious long-term morbidity. Detecting diabetic neuropathy at an early stage is crucial to prevent or delay irreversible damage [1].

Patients with diabetic neuropathy may experience dysesthetic and paresthetic symptoms, along with sensory, motor, and autonomic deficits. Beyond clinical symptoms and findings, electrophysiological studies play a key role in diagnosing diabetic neuropathy. In polyneuropathy (PNP), sensory nerves in the feet are frequently affected in the early stages [2, 3].

However, routine nerve conduction studies, which assess the sural and superficial peroneal nerves, do not effectively evaluate the distal regions of the foot [4].

Received: March 7, 2025 **Accepted:** June 5, 2025 **Available Online:** June 13, 2025 **Published:** November 4, 2025

How to cite this article: Koç Yamanyar G, Aslan H. The role of medial plantar nerve conduction studies in the diagnosis of diabetic polyneuropathy: A comparative analysis with sural nerve. Eur Res J. 2025;11(6):1184-1191. doi: 10.18621/eurj.1651488

Corresponding author: Gülçin Koç Yamanyar, MD., Phone: +90 224 975 00 00, E-mail: dr.gulcinkoc@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

Since the medial plantar nerve is one of the most distally located nerves in the foot, it may be affected in the early stages of PNP. In studies comparing medial plantar and sural nerve conduction studies, it has been shown that in some patients clinically suspected of neuropathy, sural nerve conduction findings may remain within normal limits, while medial plantar sensory response amplitudes are reduced or unrecordable during electrophysiological assessment. These findings highlight that medial plantar nerve conduction studies may represent a more sensitive method for the electrophysiological diagnosis of neuropathy [5, 6].

In this study, we aimed to evaluate the role of medial plantar nerve conduction studies in the early diagnosis of diabetic PNP in patients with clinically suspected diabetic PNP and to determine which of the examined nerve conduction studies—sural or medial plantar—demonstrates higher sensitivity.

METHODS

Between September and November 2011, a total of 60 patients who visited the Neurology outpatient clinics at Istanbul Training and Research Hospital and were referred to the Electrophysiology Laboratory with a suspected diagnosis of diabetic PNP were included in the study. The control group consisted of 30 healthy volunteers with no history of diabetes or neuropathic symptoms. Ethical approval for the study was obtained from the Medical Resaerch Ethics Committee of Istanbul Training and Research Hospital upon its dated June 15th, 2012, No: 148. Informed consent forms were obtained from all patients and healthy volunteers.

The study included patients aged 18 to 70 who had been clinically diagnosed with diabetes for at least six months and had a diabetic neuropathy symptom score of 1 or higher. Those with conditions, medications, or substance use that could cause neuropathy as well as pregnant women and individuals with entrapment neuropathy, were excluded.

Evaluation of Cases

Clinical assessment tests, patient history, physical examination findings, age, gender, disease duration, and medications used were recorded.

Patient symptoms were evaluated using the Diabetic Neuropathy Symptom (DNS) Score, a four-variable symptom scoring system. Each symptom is assigned one point, and a score of one or higher is considered indicative of diabetic PNP [7].

The control group consisted of healthy volunteers with no history of diabetes, neuropathic symptoms, medication use, or exposure to toxic substances. Their

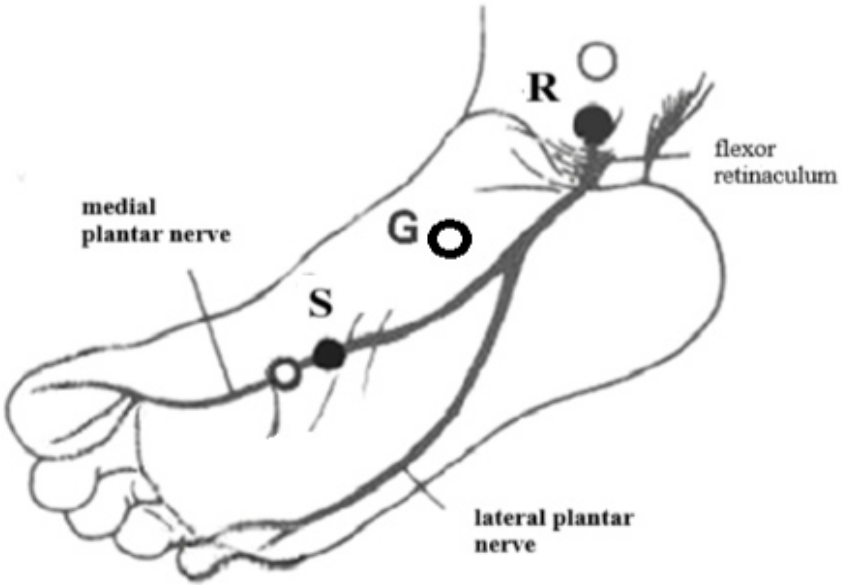

Fig. 1. Medial plantar nerve conduction technique (S=stimulating electrode, R=recording electrode, G=ground electrode (*Adapted with modifications from: Oh SJ. Clinical Electromyography: Nerve Conduction Studies. 2nd ed. Baltimore: Williams & Wilkins; 1993; p. 245.* [8]).

neurological examinations were completely normal. The control group underwent the same nerve conduction studies as the patient group.

Nerve Conduction Studies

Electrophysiological examinations were performed using a Dantec Keypoint Portable EMG device. Throughout the nerve conduction studies, the room temperature was maintained between 22-24°C. Before the examination, the skin was cleansed with alcohol in all cases to minimize skin resistance. Bipolar surface and ring electrodes were used for stimulation and recording during the nerve conduction studies.

In both the patient and control groups, motor and sensory responses of the median and ulnar nerves, posterior tibial motor response, and bilateral medial plantar and sural sensory responses were recorded. The orthodromic method was preferred for sensory nerve conduction studies.

For medial plantar nerve conduction studies, the recording electrode was placed over the flexor retinaculum, while the stimulation electrode was positioned medially on the sole of the foot, between the metatarsal bones. The nerve was stimulated at this site, and responses were recorded over the flexor retinaculum. Distal latency, amplitude, and conduction velocity of the nerve were measured (Fig. 1).

For sural nerve conduction studies, the recording electrode was placed behind the lateral malleolus. The nerve was stimulated 12-15 cm proximally, and responses were recorded. Distal latency, amplitude, and conduction velocity of the nerve were measured (Fig. 2).

Statistical Analysis

Statistical analyses were conducted using the NCSS 2007 & PASS 2008 Statistical Software. In addition to descriptive statistical methods (mean, standard deviation, median, frequency, and ratio), different tests were applied based on the data distribution. For normally distributed parameters, one-way ANOVA was used for comparisons between groups, and Tukey HSD was applied to identify the source of significant differences. For non-normally distributed parameters, Kruskal-Wallis test was used for group comparisons, and Mann-Whitney U test was applied to determine which group caused the difference. For categorical variables, the Yates-corrected Chi-square test was used. A P-value <0.05 was considered statistically significant.

RESULTS

The average age of the patients in the study group was 54.25±8.53 years, while it was 52.93±8.09 years in the control group. There was no significant difference between the two groups in terms of age and gender. Among the patients in the study group, 30% (n=18)

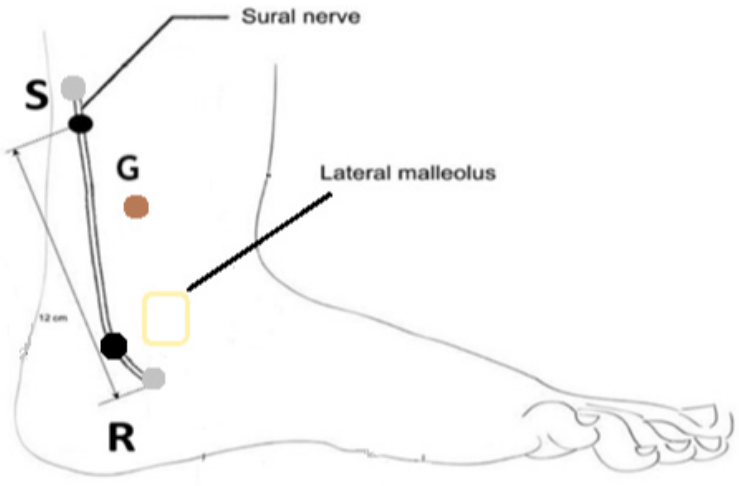

Fig. 2. Sural nerve conduction technique (S=stimulating electrode, R=recording electrode, G=ground electrode) (*Adapted with modifications from: Ricciardi D, Galiero R, Todisco V, et al. Neurophysiological assessment of peripheral neuropathy through whole plantar nerve conduction in type 2 diabetes mellitus and healthy control subjects. Metab Target Organ Damage. 2024;4:24. doi:10.20517/mtod.2024.20.* [9]).

Table 1. Demographic and clinical characteristics of participants

	Patient group (n=60)	Control group (n=30)	P value
Age (years)	54.25±8.53	52.93±8.09	0.484
Sex			
Male	24 (40%)	12 (40%)	
Female	36 (60%)	18 (60%)	
Duration of DM (years)			
<5	18 (30%)		
5-9	14 (23.30%)		
≥10	28 (46.70%)		

Data are shown as mean±standard deviation or n (%). DM=Diabetes Mellitus.
Student t Test, [a]Yates test

had been living with diabetes for less than 5 years, 23.3% (n=14) for 5-9 years, and 46.7% (n=28) for 10 years or more (Table 1).

In sural nerve conduction studies, sensory nerve action potentials could not be recorded in 9 patients. In patients with measurable sural responses, sural nerve distal latency was found to be significantly longer compared to the control group (P<0.01). Additionally, there was a notable decrease in sural nerve sensory response amplitude and conduction velocity (P<0.01) (Table 2).

In medial plantar nerve conduction studies, sensory nerve action potentials could not be recorded in 35 patients. In the 25 patients for whom medial plantar nerve sensory action potentials could be recorded, sensory nerve amplitudes were found to be significantly lower compared to the control group. (P<0.01). However, no significant difference was found in conduction velocity and distal latency measurements (Table 3).

In the nerve conduction studies conducted on healthy controls, bilateral sural nerve and medial plantar sensory responses were successfully recorded in all individuals.

Among the 60 patients in the study group, 23

Table 2. Results of sural nerve conduction studies

Sural nerve	Patient group (n=60)	Control group (n=30)	P value
Right			
Latency (msec)	2.33±0.49	2.03±0.28	**0.001****
Amplitude (µV) (median)	11.56±6.81 (11)	18.13±6.40 (16)	**0.001[a],****
Velocity (m/s)	53.27±7.95	60.15±5.55	**0.001****
Left			
Latency (msec)	2.27±0.41	2.09±0.28	**0.034***
Amplitude (µV) (median)	10.68±5.77 (11)	16.90±5.61 (14.5)	**0.00[a],****
Velocity (m/s)	53.44±7.56	59.52±4.27	**0.001****
Number of no response	9	0	

Data are shown as mean±standard deviation or median
Student-t test, [a]Mann Whitney U Test, *P<0.05, **P<0.01

Table 3. Results of medial plantar nerve conduction studies

Medial plantar nerve	Patient group (n=60)	Control group (n=30)	P value
Right			
Latency (msec)	2.08±0.37	2.16±0.35	0.451
Amplitude (µV) (median)	3.45±1.27 (3.50)	11.09±6.10 (8.25)	0.001[a,]**
Velocity (m/s)	60.08±6.26	58.72±6.95	0.456
Left			
Latency (msec)	2.03±0.33	2.05±0.28	0.754
Amplitude (µV) (median)	3.37±1.27 (3.00)	10.71±5.02 (8.70)	0.001[a,]**
Velocity (m/s)	59.88±6.40	59.87±6.51	0.996
Number of no response	35	0	

Data are shown as mean±standard deviation or median
Student-t test, [a]Mann Whitney U Test, *P<0.05, **P<0.01

(38.33%) had abnormal sural nerve sensory responses, while 39 (65%) had abnormal medial plantar nerve sensory responses (Table 4).

When comparing diabetes duration with nerve conduction findings, a statistically significant relationship was found between diabetes duration and motor and sensory responses of the median nerve, as well as motor responses of the posterior tibial nerve.

However, no significant correlation was found between sural and medial plantar sensory responses and disease duration.

DISCUSSION

In this study, we aimed to investigate the role of medial plantar nerve sensory conduction studies in the early detection of diabetic PNP by comparing sural nerve and medial plantar nerve sensory conduction measurements in patients clinically suspected of having diabetic PNP.

We found that the amplitudes of both the medial plantar and sural nerves were significantly lower in diabetic patients compared to healthy controls. While sural responses were absent in 9 patients, medial plantar sensory responses could not be recorded in 35 patients. In contrast, all healthy controls exhibited normal sensory responses in both the sural and medial plantar nerves.

A study by An *et al*. [8] reported that medial plantar nerve sensory response amplitudes and conduction velocities were significantly lower in diabetic patients compared to healthy controls [10].

In a study conducted by Altun *et al*. [11], medial plantar responses were absent in 19 out of 40 diabetic patients. Notably, in 10 out of 19 of these patients,

Table 4. Sensitivity of bilateral assessment of medial plantar and sural nerves

	Abnormal NCS (n=60)	Sensitivity
Medial plantar nerve	39	65%
Sural nerve	23	38.3%

NCS=nerve conduction studies

sural nerve conduction studies were found to be normal. Among the remaining patients, sural responses were absent in 7, while reduced sural sensory amplitudes were observed in 2 [11].

Similarly, Uluc *et al*. [12] found that among 30 diabetic patients, 9 had absent medial plantar nerve responses, while 11 had absent sural sensory responses. However, in cases where responses were recordable, they observed a significant reduction in medial plantar sensory response amplitudes, consistent with our findings [12].

It is well known that sensory nerve action potentials in the lower extremities tend to diminish and may become unrecordable in individuals over 60 years of age, making it difficult to distinguish between age-related changes and neuropathy in elderly diabetic patients [13]. However, several studies have shown that medial plantar sensory responses remain recordable in healthy controls under the age of 70 [14, 15]. Similarly, in a study by Keskin *et al*. [16] evaluating the reliability of medial plantar nerve conduction studies in healthy elderly individuals over the age of 65, medial plantar responses could not be recorded in only 2 cases, both over the age of 72, among a total of 81 participants. These results emphasize that medial plantar nerve conduction studies are reliable in healthy elderly individuals under the age of 72 [16].

In our study, the oldest patient was 66 years old, and the absence of sensory responses in some patients was attributed to the severity of diabetic PNP rather than aging. Further studies with larger patient cohorts are needed to better understand the impact of age, but it is important to note that age affects nerve conduction studies in both diabetic and healthy individuals. In our study, all healthy controls exhibited normal medial plantar and sural sensory responses, supporting this perspective.

While there was a statistically significant difference between patient and control groups in terms of medial plantar sensory response amplitudes, there was no significant difference in distal latency and conduction velocity values. This could be explained by the predominant axonal degeneration in diabetic polyneuropathy, which primarily affects sensory amplitudes rather than conduction velocities. However, since medial plantar sensory responses were absent in 35 patients, statistical comparisons were only possible in the 25 patients with measurable responses. When considering the entire patient group, it is evident that medial plantar nerve conduction is significantly impaired.

In the patient group, isolated impairment of medial plantar nerve sensory conduction was observed in 15 patients, whereas 24 patients exhibited abnormalities in multiple nerve conduction studies. Without medial plantar nerve conduction studies, the rate of electrophysiologically detected neuropathy in the patient group would be 40%, but when medial plantar nerve conduction is included, this rate increases to 65%.

When comparing sural and medial plantar nerve conduction studies, sural sensory conduction abnormalities were observed in 23 (38.3%) out of 60 diabetic patients, whereas medial plantar nerve abnormalities were detected in 39 (65%) patients. Importantly, all patients with abnormal sural sensory responses also had abnormal medial plantar nerve responses. However, 15 patients with abnormal medial plantar responses had normal sural nerve conduction studies, suggesting that medial plantar nerve abnormalities may be an earlier indicator of diabetic polyneuropathy.

Similarly, a study by Løseth *et al*. [15] reported that 59% of diabetic patients had abnormal medial plantar nerve conduction studies, while only 24% exhibited sural nerve abnormalities. Another study by An *et al*. [10] found that 46.7% of symptomatic diabetic patients and 14.3% of asymptomatic patients had abnormal medial plantar sensory action potentials, highlighting the predominant axonal degeneration in the medial plantar nerve in diabetic PNP.

These findings suggest that medial plantar nerve conduction studies are more sensitive for the electrophysiological diagnosis of diabetic PNP. This increased sensitivity may be due to the early involvement of the medial plantar nerve, which can be affected even when other sensory nerves remain intact. The early vulnerability of this nerve might be explained by its distal localization.

In light of these findings, we believe that including medial plantar nerve conduction studies in routine nerve conduction assessments for diabetic PNP can significantly improve diagnostic sensitivity.

Limitations

This study has several limitations. First, it was con-

ducted in a single center, which may limit the generalizability of the findings to broader clinical settings. Second, the study was conducted on a relatively small patient group. Future studies involving larger cohorts may provide more robust statistical results and help validate our findings. Third, the findings were not confirmed with skin biopsy, which might have provided additional insight into small fiber involvement. Finally, the cross-sectional design of the study precludes any conclusions about the progression or longitudinal changes in nerve conduction in diabetic patients over time.

CONCLUSION

This study highlights the importance of medial plantar nerve conduction studies in the early diagnosis of diabetic PNP. Our findings indicate that medial plantar nerve conduction studies are more sensitive than sural nerve conduction studies in detecting early neuropathic changes in diabetic patients. While both tests are valuable in the electrophysiological assessment of diabetic PNP, the medial plantar nerve was found to be more frequently affected.

Given the increased sensitivity, integrating medial plantar nerve conduction studies into routine diagnostic protocols may enhance the detection of diabetic neuropathy, particularly in cases where standard assessments yield normal results.

Future studies with larger patient cohorts are needed to further validate these findings and explore potential age-related influences on nerve conduction studies. However, our results suggest that including medial plantar nerve conduction studies in standard evaluations can significantly improve diagnostic accuracy of diabetic PNP.

Ethics Approval and Consent to Participate

The study was approved by the Istanbul Training and Research Hospital Clinical Research Ethics Committee (Decision no.: 148 and date: 15.06.2012). All procedures performed during data collection, review of patient records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Informed consent forms were obtained from all patients and healthy volunteers.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: GKY, HA; Study Design: GKY, HA; Supervision: HA; Funding: HA, GKY; Materials: HA, GKY; Data Collection and/or Processing: GKY; Statistical Analysis and/or Data Interpretation: GKY; Literature Review: GKY; Manuscript Preparation: GKY and Critical Review: GKY, HA.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the author(s) and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Sumner CJ, Sheth S, Griffin JW, Cornblath DR, Polydefkis M. The spectrum of neuropathy in diabetes and impaired glucose tolerance. Neurology. 2003;60(1):108-111. doi: 10.1212/wnl.60.1.108.

2. Singleton JR. Evaluation and treatment of painful peripheral polyneuropathy. Semin Neurol. 2005;25(2):185-195. doi: 10.1055/s-2005-871327.
3. Oh SJ, Melo AC, Lee DK, et al. Large-fiber neuropathy in distal sensory neuropathy with normal routine nerve conduction. Neurology. 2001;56(11):1570-1572. doi: 10.1212/wnl.56.11.1570.
4. Killian JM, Foreman PJ. Clinical utility of dorsal sural nerve conduction studies. Muscle Nerve. 2001;24(6):817-820. doi: 10.1002/mus.1074.
5. Nodera H, Logigian EL, Herrmann DN. Class of nerve fiber involvement in sensory neuropathies: clinical characterization and utility of the plantar nerve action potential. Muscle Nerve. 2002;26(2):212-217. doi: 10.1002/mus.10196.
6. Herrmann DN, Ferguson ML, Pannoni V, Barbano RL, Stanton M, Logigian EL. Plantar nerve AP and skin biopsy in sensory neuropathies with normal routine conduction studies. Neurology. 2004;63(5):879-885. doi: 10.1212/01.wnl.0000137036.26601.84.
7. Meijer JWG, Bosma E, Lefrandt JD, et al. Clinical diagnosis of diabetic polyneuropathy with the diabetic neuropathy symptoms and diabetic neuropathy examination scores. Diabetes Care. 2003;26(3):697-701. doi: 10.2337/diacare.26.3.697.
8. Oh SJ. Clinical Electromyography: Nerve Conduction Studies. 2nd ed. Baltimore: Williams & Wilkins; 1993.
9. Ricciardi D, Galiero R, Todisco V, et al. Neurophysiological assessment of peripheral neuropathy through whole plantar nerve conduction in type 2 diabetes mellitus and healthy control subjects. Metab Target Organ Damage. 2024;4(3):24. doi: 10.20517/mtod.2024.20.
10. An JY, Park MS, Kim JS, et al. Comparison of diabetic neuropathy symptom score and medial plantar sensory nerve conduction studies in diabetic patients showing normal routine nerve conduction studies. Intern Med. 2008;47(15):1395-1398. doi: 10.2169/internalmedicine.47.0901.
11. Altun Y, Demirkol A, Tumay Y, et al. The medial plantar and medial peroneal cutaneous nerve conduction studies for diabetic polyneuropathy. Neurol Sci. 2011;32(5):849-854. doi: 10.1007/s10072-011-0669-2.
12. Uluc K, Isak B, Borucu D, et al. Medial plantar and dorsal sural nerve conduction studies increase the sensitivity in the detection of neuropathy in diabetic patients. Clin Neurophysiol. 2008;119(4):880-885. doi: 10.1016/j.clinph.2008.01.001.
13. Mendell JR, Kissel JT, Cornblath DR. Editors. Diagnosis and Management of Peripheral Nerve Disorders. Oxford University Press: Oxford, 2001.
14. Lee KW, Oh SJ. Early appearance of aging phenomenon in the interdigital nerves of the foot. Muscle Nerve. 1994;17(1):58-63. doi: 10.1002/mus.880170108.
15. Loseth S, Nebuchennykh M, Stalberg E, Mellgren SI. Medial plantar nerve conduction studies in healthy controls and diabetics. Clin Neurophysiol. 2007;118(5):1155-1161. doi: 10.1016/j.clinph.2007.01.008.
16. Keskin G, Kahraman Koytak P, Bastan B, Tanridag T, Us O, Uluc K. The reliability of medial and lateral plantar nerve recordings in healthy elderly individuals. Neurol Sci. 2015;36(6):883-888. doi: 10.1007/s10072-014-2056-2.

The European Research Journal 2025;11(6):1192-1203 Original Article

DOI: https://doi.org/10.18621/eurj.1648649 *Public Health*

Impact of earthquakes on adolescent future expectations: Insights from a study in Türkiye

Osman Küçükkelepçe[1], Fatma Sena Konyalıoğlu[2], Osman Kurt[3], Fedli Emre Kılıç[4], Erdoğan Öz[2]

[1]Department of Public Health, Adıyaman Provincial Health Directorate, Adıyaman, Türkiye; [2]Department of Public Health, Ministry of Health of the Republic of Türkiye, Ankara, Türkiye; [3]Department of Public Health, İnönü University, Faculty of Medicine, Malatya, Türkiye; [4]Department of Pediatrics, Adıyaman University, Faculty of Medicine, Adıyaman, Türkiye

ABSTRACT

Objectives: While earthquakes impact the future expectations of entire societies, adolescents bear a hefty burden due to their transitional stage of life and the intricacies of adolescence itself. Our study examined the factors influencing the future expectations of adolescents residing in the region six months after the earthquake.

Methods: A face-to-face survey was conducted involving 385 individuals aged 11 to 17 affected by the February 6 earthquake in Turkey. The survey included sociodemographic inquiries, questions regarding earthquake experiences, and the administration of the Future Expectation scale to all participants.

Results: Participants who lost their family members, who were trapped under the debris, and whose family members were trapped under the debris had significantly lower future expectations. Moreover, happier individuals, who perceived themselves as healthier and more socialized, had significantly higher future expectations. Female individuals had significantly lower future expectations in all subdimensions of the scale. Lastly, adolescents who changed their career aspirations post-earthquake exhibited higher scores in all subdimensions of the scale compared to those who did not undergo such changes.

Conclusions: Adolescents confronted with traumatic events like earthquakes, particularly those who have lost relatives, often experience a negative impact on their future expectations. However, during subsequent phases, adolescents who report happiness, improved health, and increased socialization tend to harbor more positive outlooks for the future. This underscores the significance of implementing psychosocial rehabilitation efforts following earthquakes.

Keywords: Adolescent, disaster, earthquake, future expectation, psychosocial impact

Earthquakes, among the most frequent natural disasters, profoundly impact both the physical infrastructure and inhabitants of affected regions, resulting in building collapses, structural damage, injuries, and fatalities. In one of the world's most active earthquake zones, Turkey has endured numerous devastating tremors throughout its history [1]. The most recent of these occurred on February 6, 2023, when seismic events measuring 7.7 and 7.6 magnitudes struck Kahramanmaraş in southeastern Turkey and the adjacent territories of northern Syria [2]. These earthquakes left an indelible mark on the affected

Received: February 28, 2025 **Accepted:** May 11, 2025 **Available Online:** June 3, 2025 **Published:** November 4, 2025

How to cite this article: Küçükkelepçe O, Konyalıoğlu FS, Kurt O, Kılıç FE, Öz E. Impact of earthquakes on adolescent future expectations: Insights from a study in Türkiye. Eur Res J. 2025;11(6):1192-1203. doi: 10.18621/eurj.1648649

Corresponding author: Osman Küçükkelepçe, MD., PhD., Phone: +90 416 225 01 95, E-mail: osmankkelepce@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

communities, comprising approximately 15 million inhabitants. The resultant devastation led to the loss of over 50 thousand lives and left more than 115 thousand individuals injured [3]. Moreover, the psychological toll on survivors, grappling with the trauma of the catastrophe and the loss of loved ones, remains profound. Adolescents, in particular, constitute a vulnerable demographic profoundly impacted by these adversities.

According to the World Health Organization, adolescence encompasses the transitional phase between childhood and adulthood, typically spanning from ages 10 to 19. This developmental stage is characterized by ongoing cognitive, emotional, physical, and psychological growth [4]. Adolescents shift from concrete to abstract thinking, grapple with concerns about the future, seek autonomy from their families, and begin to formulate life plans [5, 6].

Future expectations denote an individual's anticipation of forthcoming events and their likelihood of occurrence, shaping their behavior and developmental trajectory [7-9]. This anticipation is particularly salient during adolescence, characterized by heightened focus on the future. When contemplating the future, adolescents predominantly ponder aspects such as education, career, family, material possessions, and lifestyle choices [10]. While many adolescents maintain a positive outlook, socioeconomically disadvantaged groups, individuals facing stressors, and those grappling with mental health issues tend to harbor more pessimistic views [10-12]. Moreover, future expectations exhibit associations with variables such as race and gender [12]. Notably, adolescents harboring positive future expectations tend to perceive life as more meaningful, demonstrate greater internal locus of control in decision-making, and cultivate warmer interpersonal relationships [13]. Conversely, negative future expectations can precipitate adverse outcomes, including heightened engagement in risky behaviors such as substance abuse and suicidal tendencies among adolescents [14].

Positive experiences can increase individuals' levels of future expectations [15]. However, traumatic experiences with wide-ranging effects, such as war, disasters, and pandemics, can damage adolescents' expectations for the future [16]. Additionally, it is possible to have positive expectations due to the desire to minimize the effects of the experienced event. Conversely, some adolescents may cultivate positive expectations as a coping mechanism to mitigate the impact of such adversities. Consequently, following catastrophic events like earthquakes, there arises a pertinent need to scrutinize adolescents' outlook on the future and explore its associated factors.

In this study, we aim to assess the future expectations of adolescents residing in the Adıyaman province, which bore the brunt of the earthquakes that struck on February 6, 2023, centered in Kahramanmaraş. This study hypothesizes that the future aspirations and outlooks of adolescents are significantly influenced by the multifaceted impact of seismic events, particularly on their social environments, demographic characteristics, and the extent to which they were personally affected by the earthquake.

METHODS

Study Design and Sampling

This cross-sectional study was conducted from August to October 2023 in Adıyaman province, in the Southeastern Anatolia region of Turkey. The study population comprised adolescents aged 11-17 who had experienced the earthquake in Adıyaman city center or surrounding districts. The sample size was calculated as 384 using the Epi Info program (version 7.2.5) with a 95% confidence interval, 5% margin of error, 50% expected frequency, 1.0 design effect and unknown population. The inclusion criteria encompassed individuals residing in tents or containers and those present in Adıyaman at the time of the earthquake. However, individuals absent from Adıyaman during the earthquake were excluded. In the study, 420 individuals were interviewed to achieve the required sample size, and 385 participants agreed to participate. Once the necessary sample size was reached, the interviews were concluded. 29 individuals were excluded from the study due to their absence from Adıyaman at the time of the earthquake, while an additional six individuals withdrew their consent to participate. No other exclusion criteria were applied. Data collection took place in the residences of earthquake survivors whose homes sustained minimal or no damage in Adıyaman city center and districts and in temporary settlement areas housing individuals whose homes were destroyed or severely damaged. Addition-

ally, adolescents seeking assistance from health units or family health centers in temporary settlement areas were included.

Questionnaire Design, Validity, and Reliability

Before starting the study, participants are first given verbal information about the study and then their written consent is obtained. Following obtaining written consent from participating earthquake survivors, survey forms were administered via face-to-face interviews. These forms, covered sociodemographic characteristics (such as age, gender, parental education and employment status, household income, marital status of parents, number of siblings, and current residence), earthquake-related experiences (including being inside during the earthquake, being trapped in debris, experiencing the loss of a family member, changing residence post-earthquake, alterations in career aspirations, and emotional responses), and other factors (such as social interactions, monetary and in-kind social support, and perceived health status). Additionally, the survey included the Future Expectations Scale for Adolescents (FESA) to assess participants' outlook on the future.

The Future Expectations Scale for Adolescents (FESA), developed by McWhirter *et al*. in 2008, is a Likert-type scale comprising 25 items. It encompasses four sub-dimensions: 'Work and Education', 'Marriage and Family', 'Religion and Society', 'Health and Life'. Responses to scale items range from 'I definitely do not believe' (scored 1 point) to 'I definitely believe' (scored 7 points). [17] The Turkish adaptation of the scale, validated by Tuncer in 2011, also comprises 25 items organized into four sub-dimensions. Specifically, 'Work and Education' encompasses 11 items, 'Marriage and Family' includes seven items, 'Religion and Society' comprises three, and 'Health and Life' contains four items. The Cronbach's alpha coefficient for the Turkish adaptation of the scale was calculated to be 0.925. [15]. An increase in the score obtained from the scale reflects a more optimistic outlook towards the future.

Ethical Approval

Ethics committee approval was obtained from Non-Interventional Research Ethics Committee of Fırat University, with a decision dated 27.07.2023 and numbered 2023/10-27. Written consent was obtained from all participants, and the study followed the principles outlined in the Declaration of Helsinki.

Statistical Analysis

Data analyses were performed using SPSS (Statistical Package for the Social Sciences; SPSS Inc., Chicago, IL) version 22. Descriptive statistics are presented as frequencies and percentages for categorical variables, and as mean±standard deviation (Mean±SD) for continuous variables. The normality of continuous variables was assessed using the Kolmogorov-Smirnov test. Group comparisons were conducted using the Mann-Whitney U-test for two groups and the Kruskal-Wallis test for more than two groups. The Spearman correlation test was utilized to explore relationships between continuous variables. Statistical significance was set at <0.05 for all analyses.

RESULTS

The study included a total of 385 participants, comprising 158 (41%) female and 227 (59%) male, with an average age of 14.2±1.9 years (range: 11 to 17 years). During the earthquake, 95.1% of the adolescents were inside a building, while 7% were trapped under debris. Additionally, 58.4% of participants experienced the loss of a family member. Regarding housing, 52.2% of adolescents resided in their own homes, 28.1% rented accommodation, and 19.7% lived in containers. Further participant characteristics are outlined in Table 1.

The predominant emotion reported by adolescents in the past month was stress, with 29.6% indicating this feeling, while the least common emotion expressed was excitement, reported by only 2.6% (Fig. 1).

Significant gender differences were observed, with girls scoring lower than boys in the domains of marriage and family (P=0.011), health and life (P=0.001), and total score (P=0.016). Additionally, adolescents with mothers educated up to secondary school level or below exhibited higher health and life scores compared to those with mothers educated to high school level or above (P=0.005). Scores in religion and society (P=0.02) and health and life (P=0.002) were significantly lower for adolescents with working mothers than those with nonworking mothers. A disparity was also noted based on income

Table 1. **All characteristics of participants**

		Data
Age (years)		14.2±1.9
Gender	Female	158 (41.0)
	Male	227 (59.0)
Mothers' education status	Middle school and below	181 (47.0)
	High school and above	204 (53.0)
Fathers' education status	Middle school and below	106 (27.5)
	High school and above	279 (72.5)
Mothers' employment	Yes	103 (26.8)
	No	282 (73.2)
Fathers' employment	Yes	322 (83.6)
	No	63 (16.4)
Income	Less than expenses	126 (32.7)
	Equal to expenses	184 (47.8)
	More than expenses	75 (19.5)
Parents' marital status	Married	363 (94.3)
	Seperated	22 (5.7)
Number of siblings	None	7 (1.8)
	1	32 88.3)
	2	103 (26.8)
	3	138 (35.8)
	≥4	105 (27.3)
Place of presence during earthquake	Inside	366 (95.1)
	Outside	19 (4.9)
Being trapped under the debris	Yes	27 (7.0)
	No	358 (93.0)
Family member being trapped under the debris	Yes	190 (49.4)
	No	195 (50.6)
Loss of family member due to earthquake	Yes	225 (58.4)
	No	160 (41.6)
Change of residence after the earthquake	Yes	171 (44.4)
	No	214 (55.6)
Current place of residence	Owner	201 (52.2)
	Rental	108 (28.1)
	Container	76 (19.7)
Level of fear during earthquake	Not afraid	11 (2.9)
	Little afraid	23 (6.0)
	Afraid	37 (9.6)
	Very afraid	65 (16.9)
	Extremely afraid	249 (64.7)

Table 1 contunied. All characteristics of participants

		Data
Change of career aspirations	Yes	89 (23.1)
	No	296 (76.9)
Frequency of socializing with friends or family	None	72 (18.7)
	1-2 times a week	244 (63.4)
	3-4 times a week	50 (13.0)
	5-6 times a week	12 (3.1)
	Every day	7 (1.8)
Perception of good health status	Very bad	20 (5.2)
	Bad	30 (7.8)
	Not good or bad	98 (25.5)
	Good	121 (31.4)
	Very good	116 (30.1)
Social support	Yes	69 (17.9)
	No	316 (82.1)
Work and Education		57.5±17.8
Marriage and Family		32.9±10.6
Religion and Society		15.7±5.7
Health and Life		20.3±6.1
Total score		126.3±35.2

Data are shown as mean±standard deviation or number (percent).

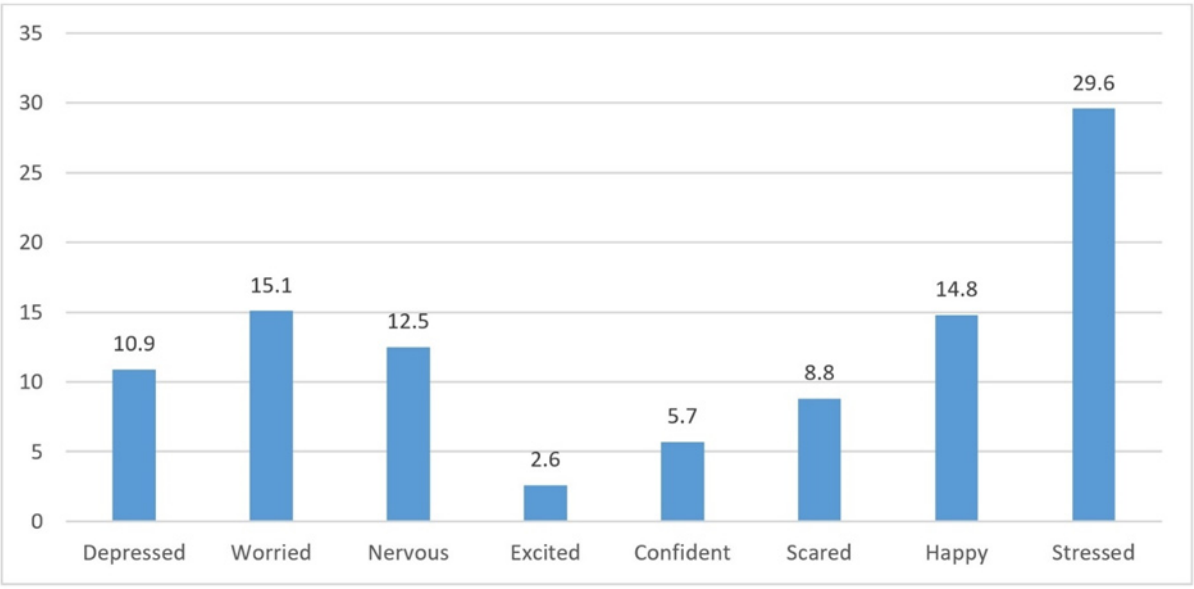

Fig. 1. Mood perception in the last month.

Table 2. Comparison of scale scores according to all characteristics

		Work and education		Marriage and family		Religion and society		Health and life		Total score	
		Mean±SD	P value*	Mean±SD	P value*	Mean±SD	P value*	Mean±SD	P value*	Mean±SD	P value*
Gender	Female	55.7±19.1	0.216	31.2±11.0	**0.011**	15.1±5.8	0.066	19.0±6.4	**0.001**	120.9±37.5	**0.016**
	Male	58.8±16.7		34.0±10.2		16.1±5.5		21.2±5.6		130.1±33.0	
Mothers' education status	Middle school and below	57.7±18.4	0.611	33.8±10.7	0.124	15.9±5.5	0.386	21.2±5.8	**0.005**	128.6±35.4	0.141
	High school and above	57.3±17.2		32.1±10.5		15.5±5.8		19.5±6.2		124.3±34.9	
Fathers' education status	Middle school and below	56.5±18.5	0.515	32.5±10.5	0.538	16.1±5.4	0.397	20.0±5.8	0.320	125.0±33.6	0.371
	High school and above	57.9±17.5		33.0±10.7		15.5±5.7		20.4±6.2		126.8±35.8	
Mothers' employment	Yes	56.1±19.3	0.652	31.7±10.9	0.209	14.6±6.0	**0.02**	18.5±6.8	**0.002**	120.9±39.3	0.143
	No	58.0±17.2		33.3±10.5		16.1±5.5		20.9±5.6		128.3±33.3	
Fathers' employment	Yes	58.2±17.3	0.162	32.9±10.8	0.665	15.7±5.6	0.629	20.3±6.1	0.517	127.1±35.3	0.243
	No	54.2±19.7		32.6±9.6		15.4±5.7		20.0±5.7		122.2±34.6	
Income	Less than expenses	56.1±19.4	0.788**	33.1±11.6	0.727**	15.3±6.3	0.673**	21.3±6.0[a]	**0.029****	125.8±38.3	0.931**
	Equal to expenses	58.0±17.4		33.0±10.5		16.0±5.3		19.9±6.3[b]		126.9±34.8	
	More than expenses	58.8±15.9		32.1±9.5		15.6±5.3		19.5±5.5[b]		125.9±30.6	
Parents' marital status	Married	58.0±17.6	**0.047**	33.2±10.6	**0.006**	15.8±5.6	0.261	20.4±6.0	**0.046**	127.4±34.8	**0.016**
	Seperated	49.7±19.8		26.8±10.3		14.1±6.6		17.7±6.5		108.3±36.6	
Place of presence during earthquake	Inside	58.1±17.1	0.066	33.1±10.5	**0.049**	15.9±5.5	**0.005**	20.4±5.9	0.134	127.6±34.0	**0.016**
	Outside	45.8±25.6		27.6±12.6		11.2±7.1		17.8±7.7		102.3±47.5	
Being trapped under the debris	Yes	43.6±24.3	**0.003**	25.9±12.6	**0.001**	11.9±6.3	**0.001**	15.1±8.6	**0.001**	96.4±49.6	**0.001**
	No	58.6±16.8		33.4±10.3		16.0±5.5		20.7±5.6		128.6±32.8	
Family member being trapped under the debris	Yes	54.4±18.3	**<0.001**	31.5±10.0	**0.016**	15.2±5.7	0.057	19.4±5.9	**0.002**	120.6±34.5	**0.002**
	No	60.5±16.8		34.2±11.1		16.1±5.6		21.1±6.1		131.9±34.9	
Loss of family member due to earthquake	Yes	53.8±18.0	**<0.001**	30.2±9.6	**<0.001**	14.8±5.8	**<0.001**	19.0±5.8	**<0.001**	117.7±32.8	**<0.001**
	No	62.7±16.1		36.6±10.9		17.0±5.2		22.1±6.0		138.4±34.8	

Table 2 contunied. Comparison of scale scores according to all characteristics

		Work and education		Marriage and family		Religion and society		Health and life		Total score	
		Mean±SD	P value*	Mean±SD	P value*	Mean±SD	P value*	Mean±SD	P value*	Mean±SD	P value*
Change of residence after the earthquake	Yes	53.7±19.6	**0.001**	31.2±10.9	**0.008**	14.9±6.0	**0.036**	19.2±6.3	**0.002**	118.9±37.8	**0.001**
	No	60.6±15.6		34.2±10.2		16.3±5.3		21.2±5.7		132.2±31.8	
Current place of residence	Owner	59.5±16.2	0.189	33.5±10.3	0.533	16.1±5.3	0.312	20.8±5.4	0.381	129.9±32.1	0.267
	Rental	55.9±16.9		32.2±10.0		15.2±5.4		19.6±6.5		122.9±33.9	
	Container	54.7±22.2		32.1±12.3		15.2±6.7		19.7±7.0		121.7±43.2	
Change of career aspirations	Yes	48.3±19.2	**<0.001**	29.6±9.9	**0.001**	14.2±6.0	**0.002**	18.7±6.1	**0.005**	110.8±36.3	**<0.001**
	No	60.3±16.4		33.8±10.7		16.1±5.5		20.7±6.0		131.0±33.5	
Social support	Yes	55.1±21.3	0.703	33.0±11.9	0.629	15.4±6.2	0.844	19.8±7.1	0.970	123.4±42.2	0.956
	No	58.0±16.9		32.8±10.3		15.7±5.5		20.4±5.8		126.9±33.5	

SD=standard deviation, Mann Whitney U test, **Kruskal Wallis analysis was applied. [a,b]The group where the difference originates, a>b

level, with those experiencing income deficits scoring higher in health and life (P=0.029). Furthermore, significant differences were found concerning parental marital status, with higher scores observed in work and education (P=0.047), marriage and family (P=0.006), health and life (P=0.046), and total score (P=0.016) for adolescents whose parents were married compared to those with separated parents.

Adolescents who were present inside buildings during the earthquake reported higher scores in marriage and family (P=0.049), religion and society (P=0.005), and total scores (P=0.016) compared to those outside. Conversely, those trapped under debris exhibited lower scores in work and education (P=0.003), marriage and family (P=0.001), religion and society (P=0.001), health and life (P=0.001), and total score (P=0.001) compared to their counterparts.

Similarly, family members trapped under debris were associated with lower scores in work and education (P<0.001), marriage and family (P=0.016), health and life (P=0.002), and total score (P=0.002) compared to those with untrapped family members. Furthermore, adolescents who lost relatives in the earthquake reported lower scores in work and education (P<0.001), marriage and family (P<0.001), religion and society (P<0.001), health and life (P<0.001), and total score (P<0.001) compared to those without such losses. Lastly, adolescents who changed their career aspirations post-earthquake exhibited higher scores in work and education (P<0.001), marriage and family (P=0.001), religion and society (P=0.002), health and life (P=0.005), and total score (P<0.001) compared to those who did not undergo such changes (Table 2).

Significant variations in emotional experiences over the last month were observed, particularly about the total score of future expectations (P<0.001). Notably, individuals reporting excitement and happiness exhibited the highest scores, with a disparity between this group and others (Fig. 2).

A significant positive correlation was observed between the subdimensions of the scale scores and the total score. Additionally, age showed a significant negative correlation between the health and life sub-dimensions and the total score. Moreover, a significant positive correlation was found between the number of siblings and the health and life sub-dimension. Notably, an increase in the fear experienced during the

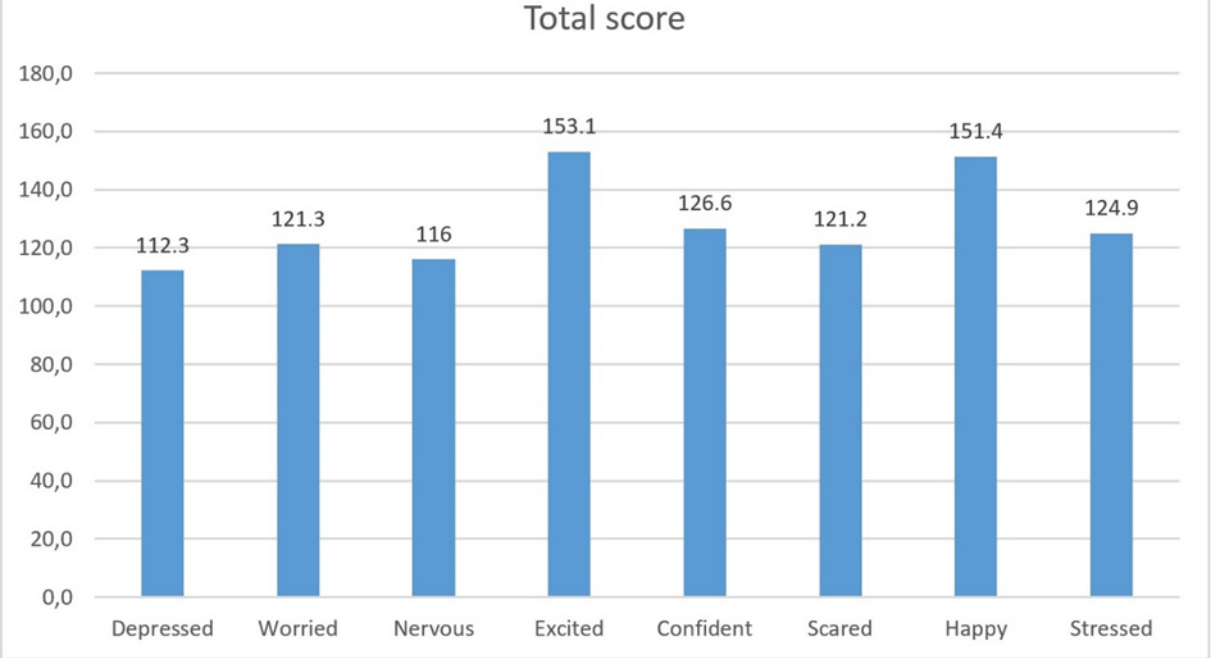

Fig. 2. **Distribution of the total scores according to the mood perception in the last month.**

Table 3. **Correlation of scale scores**

		Work and education	Marriage and family	Religion and society	Health and life	Total score
Marriage and family	r	**.638**				
	P value	**.000**				
Religion and society	r	**.599**	**.725**			
	P value	**.000**	**.000**			
Health and life	r	**.694**	**.711**	**.624**		
	P value	**.000**	**.000**	**.000**		
Total score	r	**.889**	**.880**	**.792**	**.827**	
	P value	**.000**	**.000**	**.000**	**.000**	
Age	r	-.083	-.087	-.030	**-.138**	**-.101**
	P value	.104	.090	.559	**.007**	**.049**
Number of siblings	r	.065	.085	.063	**.110**	.071
	P value	.201	.097	.220	**.031**	.162
Level of fear during earthquake	r	.030	.077	.090	**.118**	.078
	P value	.557	.134	.077	**.020**	.127
Frequency of socializing with friends or family	r	**.147**	**.196**	**.119**	**.159**	**.178**
	P value	**.004**	**.000**	**.019**	**.002**	**.000**
Perception of good health status	r	**.434**	**.323**	**.213**	**.279**	**.392**
	P value	**.000**	**.000**	**.000**	**.000**	**.000**

earthquake was associated with a significant increase in the health and life subscale score. Furthermore, an increase in the frequency of socializing with friends or family was linked to significant increases in the scale subscale and total score. Finally, an enhanced perception of good health status corresponded to significant increases in the scale sub-dimension and total score (Table 3).

DISCUSSION

This study assessed the future expectations of adolescents living in Adıyaman after the February 6 earthquake. It was observed that adolescents exposed to the earthquake had lower levels of future expectations compared to studies conducted in Turkey using the same scale but not related to earthquakes [18-20]. This situation could be evidence that the earthquake negatively impacts adolescents' expectations for the future in this country.

A literature review reveals no direct studies examining the impact of earthquakes on adolescents' future expectations. However, there are studies examining the relationship between earthquakes and hopelessness. For example, a study conducted in a different city affected by the same earthquake observed a dominant sense of hopelessness among adolescents [21]. Another study found that hopelessness regarding the future increased after the earthquake [22]. In another study involving earthquake exposure, it was found that individuals' trauma levels were positively correlated with their levels of hopelessness [23]. According to these findings, hopelessness and lack of expectations for the future in adolescents are significant issues that can arise from traumatic life events.

The February 6 earthquakes predominantly struck at night, catching most individuals indoors. Those outside during the earthquake exhibited lower expectations regarding marriage, children, overall health, and life, and general future prospects compared to those inside buildings at the time of the event. Furthermore, individuals who found themselves trapped under debris experienced even lower levels of future expectations than those who avoided entrapment. This trend is unsurprising, given the traumatic experiences endured by those trapped, including fear of death and panic. Contrary to expectations, individuals not inside buildings during the earthquake also reported diminished future expectations. Factors such as being in chaotic post-earthquake environments, witnessing search and rescue efforts, and severely injured or deceased individuals may have contributed to a heightened sense of hopelessness regarding the future among this group. Additionally, individuals with family members trapped under debris, those who lost relatives in the earthquake, and those compelled to relocate their residences afterward also expressed lower future expectations. It is plausible that these individuals, facing the insecurity of their safety and that of their loved ones, harbored pessimistic outlooks on the future. In a study conducted on adolescents affected by the Marmara earthquake, it was found that those who had family members who died or were injured due to the earthquake had higher levels of hopelessness compared to those who did not have such experiences. However, no difference was observed in hopelessness levels based on the damage to the house [24]. A qualitative study reported that losing relatives and one's home harmed future expectations [21]. Another study found that those who were trapped under rubble, lost a family member, witnessed the collapse of buildings, or saw someone severely injured during the earthquake had higher levels of trauma post-earthquake [25]. Exposure to trauma negatively affects individuals' expectations for the future [26].

In this study, conducted within 6 to 8 months after the earthquake, the majority of adolescents reported feeling stressed in the past month. Those who felt stressed, depressed, tense, anxious, or scared had lower levels of future expectations compared to those who felt happy or excited. The negative emotions felt by adolescents are likely symptoms of mental health issues. According to the literature, mental health problems such as depression and post-traumatic stress disorder (PTSD) can increase in adolescents after an earthquake, affecting their psychological functioning, perceptions, and expectations for the future [18, 27, 28]. Another study conducted on university students in Turkey who were exposed to a different earthquake also reported that those with higher levels of hopelessness about the future experienced higher levels of stress [29].

In this study, experiencing intense fear during the earthquake was associated with higher expectations for a healthy and long life. Consistent with this result,

the literature also shows that fear of death is related to the desire for a long life [30]. Additionally, in the current study, spending time outside with close friends and family, socializing, and perceiving oneself as healthier were associated with higher future expectations. Similarly, the literature indicates that social support from friends is related to higher hopes for the future [31]. Improving adolescents' socialization and peer relationships in the school environment, where they spend a significant portion of their day, and encouraging the social support they provide will play an important role in enhancing adolescents' future expectations. Socializing and spending quality time with parents positively affect the well-being of children and adolescents [32]. Well-being is a condition that enhances hope for the future [33]. Therefore, spending more time outside with close ones and high self-rated health scores is expected to be associated with higher future expectations in the present study.

Gender disparities were evident in the total future expectation score, as well as in the marriage and family, and health and life sub-dimensions, with lower scores observed among girls. This suggests that girls have lower expectations regarding marriage, family, and overall health and life than boys. Previous literature presents conflicting findings on gender-based future expectations [17, 34]. While some studies indicate that young women prioritize family-oriented goals. In contrast, young men focus more on career aspirations, while others suggest that girls may have lower expectations regarding marriage and family than boys [7, 17, 35, 36]. Such discrepancies may stem from variations in gender roles across different communities. When examining the relationship between age and future expectations, it was found that general future expectations and the expectation of a long and healthy life increase as age decreases. Similarly, a study conducted during the pandemic found that younger adolescents had higher overall future expectations and expectations for a long and healthy life [37]. Another study highlighted that older adolescents prioritize future goals related to education, career, and family more than younger adolescents.7 In the current study, more siblings were also associated with the expectation of a healthy and long life. However, another study found that, although not statistically significant, those with siblings had lower future expectation levels [18]. While siblings can be a source of support or conflict, they are mostly considered supportive family members [34].

In the current study, the education and employment status of the mother influenced adolescents' levels of future expectations. At the same time, no difference was found based on the father's education and employment status. A study using the same scale found no difference based on the mother's education status [18]. Another study found that adolescents with working parents had higher future goals and educational planning [38]. In this study, those with lower household income had higher health and life expectations. Economic difficulties did not hinder adolescents' hopes for a healthy and long life. However, other studies have shown that lower socio-economic status of families led adolescents to view the future more fatalistically and pessimistically [15, 39]. In the present study, a difference was observed based on the parents' marital status, with those having married parents showing higher future expectations, except in the religion and community sub-dimension. Although no direct study examines this relationship, the literature shows that adolescents with both parents are more resilient and optimistic than those with separated parents [40, 41].

Limitations

Although there are studies in the literature evaluating the hopelessness levels of adolescents after earthquakes, this study is the first to assess future expectations directly. Because the face-to-face interview method was used, issues such as skipping questions or not understanding them were avoided. However, the study's cross-sectional design limits the ability to determine the direction of causality. Since the study was conducted in a single province, the results cannot be generalized to all adolescents nationwide.

CONCLUSION

Plans, goals, and expectations for the future are essential concepts during an adolescent's developmental stages. Positive future expectations encourage high resilience and a successful transition into adulthood. Conversely, negative future expectations are associated with risky behaviors in adolescents. Therefore, interventions are necessary to improve the future expectations of adolescents affected by earthquakes.

School guidance and psychological counselors can be beneficial by helping adolescents replace negative emotions with positive ones and increasing their hope for the future. Based on the findings of this study, interventions can prioritize those more affected by the earthquake (those experiencing negative emotions like stress, those who have lost family members, or those who were trapped under rubble) and girls. Additionally, future studies should evaluate the future expectations of adolescents at later stages after the earthquake to assess changes over time.

Ethics Approval and Consent to Participate

This study was approved by the Fırat University Non-Interventional Research Ethics Committee (Decision no. 2023/10-27, date: 27.07.2023). Written consent was obtained from all participants, and the study followed the principles outlined in the Declaration of Helsinki.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: OK, FSK, EÖ; Study Design: OK, FSK, EÖ, FEK; Supervision: FSK, EÖ; Funding: N/A; Materials: OK, FSK, EÖ; Data Collection and/or Processing: OK, OKurt, FEK; Statistical Analysis and/or Data Interpretation: OK, OKurt, FSK; Literature Review: OK, FSK, EÖ; Manuscript Preparation: OK, FSK, OKurt, FEK and Critical Review: OK, FSK, FEK.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Emre Ö, Duman TY, Özalp S, et al. Active fault database of Turkey. Bull Earthq Eng. 2018;16(8):3229-3275. doi: 10.1007/s10518-016-0041-2.
2. World Health Organization. Earthquake response in Türkiye and whole of Syria [Internet]. 2023 [cited 2024 January 20]. Available from: https://www.who.int/publications/m/item/who-flash-appeal--earthquake-response-in-t-rkiye-and-whole-of-syria.
3. Ergönül Ö, Keske Ş, Ksinzik A, et al. The challenges in the monitoring of infectious diseases after the earthquake in Türkiye in 2023. Lancet Infect Dis. 2023;23(11):e482-488. doi: 10.1016/S1473-3099(23)00362-6.
4. World Health Organization. Adolescent health [Internet]. [cited 2024 January 20]. Available from: https://www.who.int/health-topics/adolescent-health#tab=tab_1.
5. Ahioğlu-Lindberg EN. [Piaget and Cognitive Development in Adolescence]. Kastamonu Egitim Derg. 2011;19(1):1-10. [Article in Turkish]
6. Derman O. [Psychosocial Development in Adolescents]. İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Sürekli Tıp Eğitimi Etkinlikleri. 2008;63:19-21. [Article in Turkish]
7. Nurmi J-E. How do adolescents see their future? a review of the development of future orientation and planning. Dev Rev. 1991;11(1):1-59. doi: 10.1016/0273-2297(91)90002-6.
8. Bandura A, Barbaranelli C, Caprara GV, Pastorelli C. Self-efficacy beliefs as shapers of children's aspirations and career trajectories. Child Dev. 2001;72(1):187-206. doi: 10.1111/1467-8624.00273.
9. Seginer R. Future orientation in times of threat and challenge: how resilient adolescents construct their future. Int J Behav Dev. 2008(4);32:272-282. doi: 10.1177/0165025408090970
10. Iovu MB. Future expectations of senior high schoolers in Romania. Int J Adolesc Youth. 2015;20(4):518-527. doi: 10.1080/02673843.2013.810162.
11. Raffaelli M, Koller SH. Future expectations of Brasilian street youth. J Adolesc. 2005;28(2):249-262. doi: 10.1016/j.adoles-

cence.2005.02.007.
12. Thompson R, Wiley TRA, Lewis T, et al. Links between traumatic experiences and expectations about the future in high risk youth. Psychol Trauma Theory Res Pract Policy. 2012;4(3):293-302. doi: 10.1037/a0023867.
13. Çalışkan S, Dilmaç B. [Predictive Relationships Between Adolescents'Future Expectation, Hope and Psychological Well-Being]. Necmettin Erbakan Üniversitesi Ereğli Eğitim Fakültesi Derg. 2021;3(2):131-143. doi: 10.51119/ereegf.202216. [Article in Turkish]
14. Stickley A, Koposov R, Kamio Y, et al. Attention deficit/hyperactivity disorder and future expectations in Russian adolescents. ADHD Atten Deficit Hyperact Disord. 2019;11(3):279-287. doi: 10.1007/s12402-019-00292-w.
15. Tuncer M. [Adaptation of adolescent future expectations scale]. J Turkish Stud. 2011;6(3):1265-1275. doi: 10.7827/turkishstudies.2335. [Article in Turkish]
16. Gökler Danışman I, Okay D. [Disasters, Children and Adolescents: The Impacts and Psychological Interventions]. Türkiye Klin J Psychol Spec Top. 2017;2(3):189-197. [Article in Turkish]
17. McWhirter EH, McWhirter BT. Adolescent future expectations of work, education, family, and community development of a new measure. Youth Soc. 2008;40(2):182-202. doi: 10.1177/0044118X08314257.
18. Topaktaş B, Dündar C, Pekşen Y. A cross-sectional analysis of the relationship among adolescents' perceived social support, psychological state and future expectations among Turkish students. Isr J Psychiatry. 2017;54(2):25-30.
19. Uzun K, Tagay Ö, Cırcır O. Childhood traumas and future expectations in adolescents: examination of the role of perceived social support and attitudes to seeking psychological help. Curr Psychol. 2024;43(8):7116-7130. doi: 10.1007/s12144-023-04925-2.
20. Özçevik Subaşi D, Akça Sümengen A, Ekim A, Ocakçı AF, Beşer A. The relation between quality of life and future expectations for refugee adolescents. J Child Adolesc Psychiatr Nurs. 2021;34(3):206-211. doi: 10.1111/jcap.12314.
21. Yanardağ R. [Discovering the Basic Problems and Needs of Earthquake Victim Adolescents Living in A Container City]. Türkiye Sosyal Araştırmalar Derg. 2024;28(1):1-22. doi: 10.20296/tsadergisi.1369894. [Article in Turkish]
22. Artar M. [Contents of Future Orientations ol Adoiescents Who Lived an Earthquake in 1999]. Kriz Derg. 2003;11(3):21-7. doi: 10.1501/Kriz_0000000195. [Article in Turkish]
23. Kardaş F, Tanhan F. [Investigating Hopelessness and Trauma Levels of High School Students Exposed to the Van Earthquake]. Sakarya Univ J Educ. 2014;4(1):102-115. doi: 10.19126/suje.52552. [Article in Turkish]
24. Dizer D. Sakarya ilindeki liseli ergenlerin 1999 Marmara Depremi sonrası travmayı algılama, sosyal destek sistemleri ve umutsuzluk belirtilerinin incelenmesi [dissertation]. Maltepe Üniversitesi; 2008.
25. Akçay S, Beydili Gürbüz E, Koca E. Trauma levels after the 2023 Kahramanmaraş earthquakes in Türkiye. J Soc Serv Res. 2024;50(4):1-12. doi: 10.1080/01488376.2024.2339967.
26. Center for Substance Abuse Treatment. Chapter 3, understanding the impact of trauma. In: Trauma-informed care in behavioral health services. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014:59-85.
27. Ekşi A, Braun KL, Ertem-Vehid H, et al. Risk factors for the development of PTSD and depression among child and adolescent victims following a 7.4 magnitude earthquake. Int J Psychiatry Clin Pract. 2007;11(3):190-199. doi: 10.1080/13651500601017548.
28. Şalcıoğlu E, Başoğlu M. Psychological effects of earthquakes in children: prospects for brief behavioral treatment. World J Pediatr. 2008;4(3):165-172. doi: 10.1007/s12519-008-0032-8.
29. Kardaş F, Tanhan F. [Investigating Posttraumatic Stress, Posttraumatic Growth and Hopelessness Levels of University Students Exposed to the Van Earthquake]. YYÜ Eğitim Fakültesi Derg. 2018;15(1):1-36. doi: 10.23891/efdyyu.2018.60. [Article in Turkish]
30. Rupprecht FS, Martin K, Lang FR. Aging-related fears and their associations with ideal life expectancy. Eur J Ageing. 2022;19(3):587-597. doi: 10.1007/s10433-021-00661-3.
31. Kwok SYCL, Gu M, Lai KY. A longitudinal study of perceived social support from friends and hope in adolescents: emotional intelligence as the mediator. Curr Psychol. 2024;43(25):21518-21529. doi: 10.1007/s12144-024-05875-z.
32. Li D, Guo X. The effect of the time parents spend with children on children's well-being. Front Psychol. 2023;14:1096128. doi: 10.3389/fpsyg.2023.1096128.
33. Karaca S, Karakoc A, Bingol F, Eren N, Andsoy II. Comparison of subjective wellbeing and positive future expectations between working and nonworking adolescents in Turkey. Iran Red Crescent Med J. 2016;18(2):e21055. doi: 10.5812/ircmj.21055.
34. Sulimani-Aidan Y, Benbenishty R. Future expectations of adolescents in residential care in Israel. Child Youth Serv Rev. 2011;33(7):1134-1141. doi: 10.1016/j.childyouth.2011.02.006.
35. Seginer R. Future orientation: developmental and ecological perspectives. 1st ed. New York: Springer; 2009.
36. Zeira A, Dekel R. The self-image of adolescents and its relationship to their perceptions of the future. Int Soc Work. 2005;48(2):177-191. doi: 10.1177/0020872805050208.
37. Küçükkelepçe O, Parlak ME, Öz E, Kapıcı Y, Ener D. Future expectation levels of adolescents during the COVID-19 pandemic. Front Public Health. 2023;11:1199280. doi: 10.3389/fpubh.2023.1199280.
38. Frasquilho D, Gaspar de Matos M, Gaspar T, Caldas de Almeida JM. Do adolescents' future expectations differ by parental employment situation? Highlights from the Portuguese HBSC/WHO survey. Int J Dev Res. 2014;4(10):1981-1986.
39. Yılmaz C, Yalçın H. [Attitudes Towards Marriage and Future Expectations of Single Parent Adolescent]. Karatay Sosyal Araştırmalar Derg. 2021;6:51-72. [Article in Turkish]
40. Şahin H. Ergenlerin psikolojik sağlamlık düzeyleri ve hayat amaçları arasındaki ilişkisinin incelenmesi [dissertation]. Haliç Üniversitesi Sosyal Bilimler Enstitüsü; 2018.
41. Güney E, Yalçın SB. [The Resilience Levels and Perceived Social Support Levels of Adolescents whose Parents Are Divorced and Not Divorced]. Necmettin Erbakan Univ Eregli Egitim Fak Derg. 2020;2(2):217-229. doi: 10.51119/ereegf.2020.7. [Article in Turkish]

The European Research Journal 2025;11(6):1204-1212 Original Article

DOI: https://doi.org/10.18621/eurj.1641700 *Emergency Medicine*

Mortality risk prediction in emergency department patients: Modeling approaches and performance analysis with gradient boosting

Erkan Boğa

Department of Emergency Medicine, Esenyurt Necmi Kadıoğlu State Hospital, İstanbul, Türkiye

ABSTRACT

Objectives: The aim of this study is to evaluate the effectiveness of the Gradient Boosting algorithm in predicting mortality risk among emergency department patients and to identify the most critical demographic, clinical, and physiological data for these predictions. This study is designed to support early identification and enhance clinical decision support systems.

Methods: This retrospective study analyzed data from 1,500 patients who visited a state hospital's emergency department between January 1 and August 31, 2024. Data were collected based on multidimensional features such as demographic information, vital signs, laboratory results, and clinical history. The Gradient Boosting algorithm was used to develop the model, and its performance was evaluated using metrics such as accuracy, sensitivity, specificity, and F1 score.

Results: The Gradient Boosting model identified oxygen saturation, age, and heart rate as the most significant predictors of mortality. The CatBoost algorithm demonstrated the highest performance with an accuracy of 88.8% and an F1 score of 85%. The model was proven to be highly accurate in predicting mortality risk.

Conclusions: Gradient Boosting algorithms, particularly CatBoost, emerged as a reliable and effective tool for predicting mortality risk. This model can contribute to the development of clinical decision support systems in emergency department settings.

Keywords: Emergency Department, gradient boosting, mortality prediction, machine learning, clinical decision support

Emergency departments are critical components of healthcare systems and conduct evaluations and treatment for patients with high mortality risk in a short period of time. The timely implementation of appropriate treatment strategies is crucial for improving clinical outcomes and optimizing resource management; therefore, early identification of such patients is vital [1]. Identifying those patients at risk of mortality is not without its challenges in the emergency setting, where the diversity of patients, complex clinical presentations, and rapid data flow place further demands on earlier identification [2]. Here, data science and machine learning techniques offer useful methods for extracting the knowledge hidden in healthcare data and generating prescriptive decision support systems [3].

Received: February 17, 2025 **Accepted:** May 23, 2025 **Available Online:** June 12, 2025 **Published:** November 4, 2025

How to cite this article: Boğa E. Mortality risk prediction in emergency department patients: Modeling approaches and performance analysis with gradient boosting. Eur Res J. 2025;11(6):1204-1212. doi: 10.18621/eurj.1641700

Corresponding author: Erkan Boğa, MD., Phone: +90 212 596 19 99, E-mail: drerkanboga@gmail.com

However, there is a machine learning method that outperforms the rest - the Gradient Boosting algorithm for classification and regression problems. It is widely used in healthcare data analytics because it can capture complex relationships within a dataset and provide high prediction accuracy [4]. Gradient Boosting also has strong generalization ability and flexibility, which makes it a particularly competitive modeling tool for predicting mortality risk in the emergency department [5].

This study has two main objectives: to prepare clinical data for use with the Gradient Boosting algorithm in order to build a model able to predict risk of mortality from patients who attended the emergency department, and to assess the performance of such a predictive model. This is accomplished through the use of multidimensional data features such as demographic details, vital signs, lab values, and clinical history [6]. Metrics such as accuracy, sensitivity, specificity, and F1 score will be used to assess model performance, and results will be compared with similar studies in the literature [7].

This study aims to evaluate the effectiveness of the Gradient Boosting algorithm in predicting mortality risk among emergency department patients and to identify the most critical demographic, clinical, and physiological data used in these predictions [8]. The results obtained will support early intervention processes and contribute to the development of clinical decision support systems [9].

This study investigates whether the Gradient Boosting algorithm can be effectively used to predict mortality risk in emergency department patients and which demographic, clinical, and physiological data are most critical for mortality prediction. The hypotheses tested in this study are as follows: the Gradient Boosting algorithm can achieve high accuracy in predicting the mortality risk of emergency department patients; physiological factors such as low oxygen saturation, low blood pressure, and high respiratory rate are strongly associated with mortality risk; and demographic data and clinical history can improve the accuracy of prediction models. Based on these questions and hypotheses, the aim of this study is to develop a model using the Gradient Boosting algorithm to predict mortality risk and to evaluate the performance of this model. The results obtained will contribute to improving clinical decision support systems in the emergency department setting.

METHODS

This study has a retrospective design and analyzes data from 1,500 patients who visited the Emergency Department of State Hospital between January 1, 2024, and August 31, 2024. The primary aim of the study is to develop a model using the Gradient Boosting algorithm to predict the mortality risk of these patients and to evaluate the performance of the model.

Patients aged 18 years and older with complete datasets, including demographic information, vital signs, laboratory results, and clinical history, were included in the study. Patients with incomplete or erroneous data, missing treatment or clinical evaluation records, and those under 18 years of age were excluded.

Patient data were retrospectively collected from the hospital information management system (HIMS) and categorized into the following groups: demographic data (age, gender), vital signs (heart rate, blood pressure, oxygen saturation, body temperature, respiratory rate), laboratory results (complete blood count, electrolyte levels, liver and kidney function tests, inflammatory markers, blood gas analyses), clinical history (comorbidities, current diagnoses, previous hospitalizations, surgical histories), treatment information (treatments administered and medications given in the emergency department), and outcome data (discharge status, discharge duration, and mortality within 24 hours, 48 hours, or 30 days).

Missing data were handled using median imputation. Outliers were detected and managed using Z-scores and interquartile range (IQR) methods. Continuous variables were normalized, and categorical variables were processed using one-hot encoding.

For model development, the dataset was randomly split into 70% training and 30% test subsets. The Gradient Boosting algorithm was employed, and hyperparameter optimization was performed using GridSearchCV. The tested hyperparameters included the learning rate (0.01, 0.1, 0.2), the number of weak learners (100, 200, 300), the maximum depth of trees (3, 5, 7), subsample ratio (0.5, 0.7, 1.0), and the proportion of features used for each tree (0.5, 1.0). The optimal combination was determined to be a learning rate of 0.1, 200 weak learners, a maximum depth of 5, a subsample ratio of 0.7, and a colsample_bytree of 0.8.

This study was conducted in accordance with the

Declaration of Helsinki and approved by the Ethics Committee of Medipol University. All data were anonymized and used solely for scientific purposes.

Statistical Analysis

Statistical analysis included exploratory data analysis (EDA) to examine the general characteristics of the dataset and its relationship with the target variable. Chi-square tests were used for categorical variables, while T-tests or Mann-Whitney U tests were applied for continuous variables. The performance of the model was evaluated using metrics such as accuracy, sensitivity, specificity, F1 score, and the area under the ROC curve (AUC).

RESULTS

The dataset analysis revealed no missing data or outliers in key variables such as age, oxygen saturation level, and heart rate. Normality tests showed that these variables did not follow a normal distribution ($P<0.05$). Descriptive statistics indicated a mean age of 52.87±20.86 years, a mean oxygen saturation level of 83.22±8.84%, and a mean heart rate of 89.66±17.16 beat/min. The mortality rate was 52.40%, with 47.60% of patients surviving (Fig. 1).

Statistical tests found significant differences in oxygen saturation levels between survivors and deceased patients ($P<0.05$), while age and heart rate showed no significant differences. Categorical variables such as low oxygen saturation, low blood pressure, and high respiratory rate were significantly associated with mortality ($P<0.05$). Correlation analysis revealed a strong negative correlation between oxygen saturation and mortality, while high respiratory rate and heart rate showed moderate positive correlations with mortality.

The most significant predictors of mortality in the Gradient Boosting model were oxygen saturation, age, and heart rate. Among Gradient Boosting algorithms, CatBoost demonstrated the highest performance, achieving an accuracy of 88.8%, an F1 score of 85%, and a ROC-AUC of 0.911. Hyperparameter optimiza-

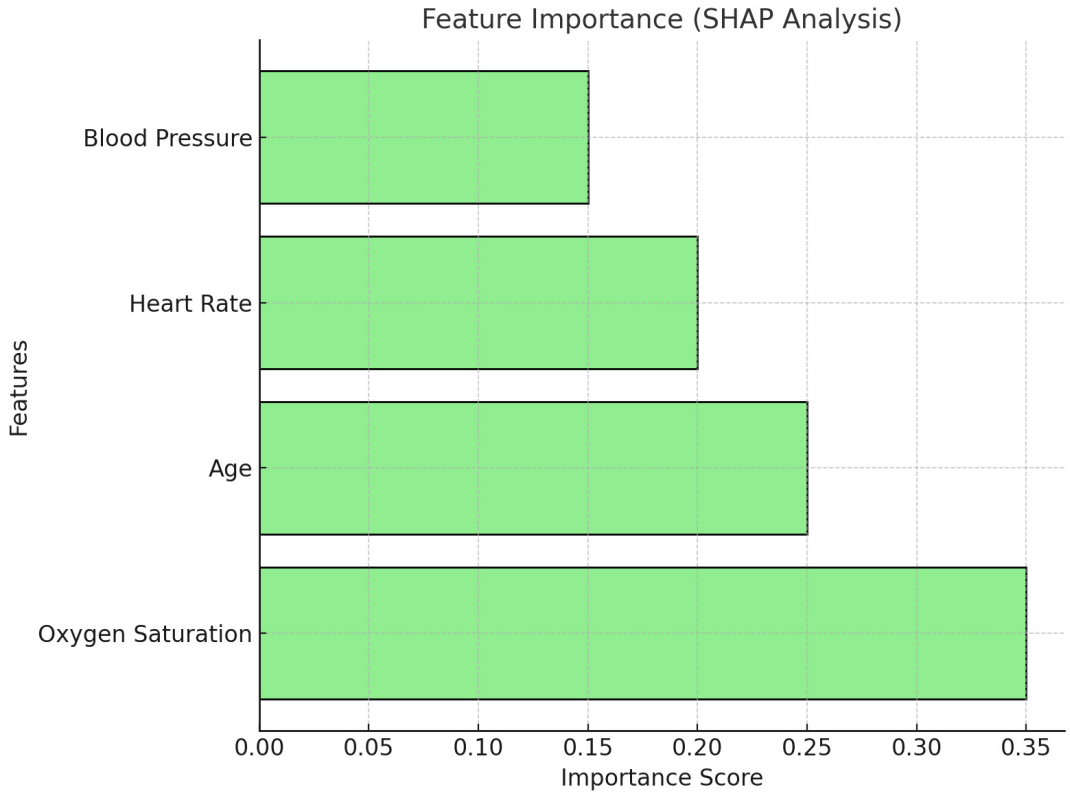

Fig. 1. Feature importance (SHAP analysis). The horizontal bar chart displays the importance of various features in predicting mortality. Oxygen saturation was the most significant factor, followed by age, heart rate, and blood pressure. These features significantly influenced the model's predictions.

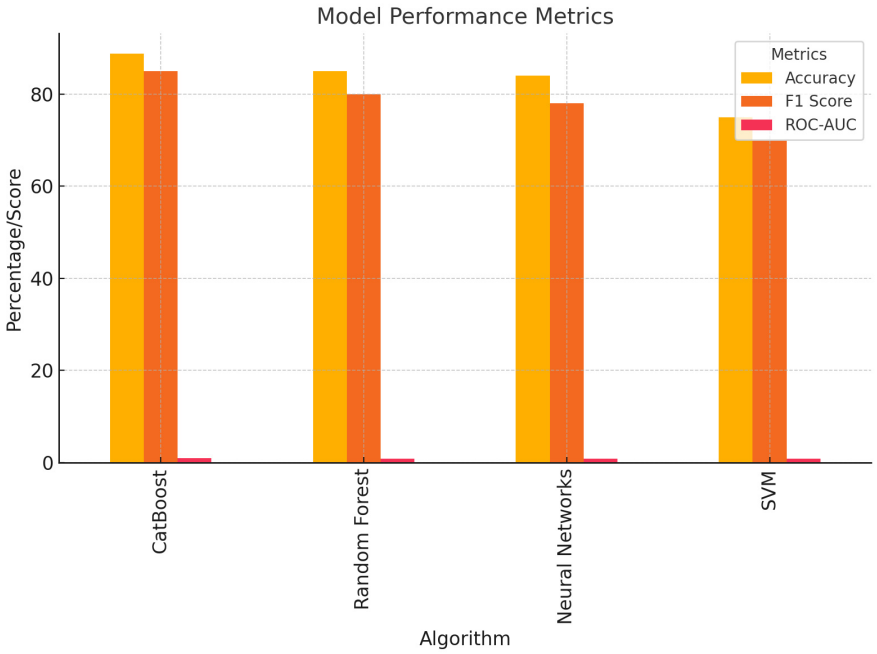

Fig. 2. Model performance metrics. The bar chart illustrates the performance of various machine learning algorithms used for mortality prediction. CatBoost outperformed other models with an accuracy of 88.8%, an F1 score of 85.0%, and a ROC-AUC of 0.911. Random Forest and Neural Networks also performed well, while SVM had comparatively lower metrics.

tion improved performance by 1-3%. SHAP analysis confirmed that low oxygen saturation had the greatest impact on mortality prediction, followed by age and low blood pressure.

Additional machine learning algorithms, including Random Forest, Neural Networks, and Support Vector Machines (SVM), were compared. CatBoost outperformed these models with superior accuracy,

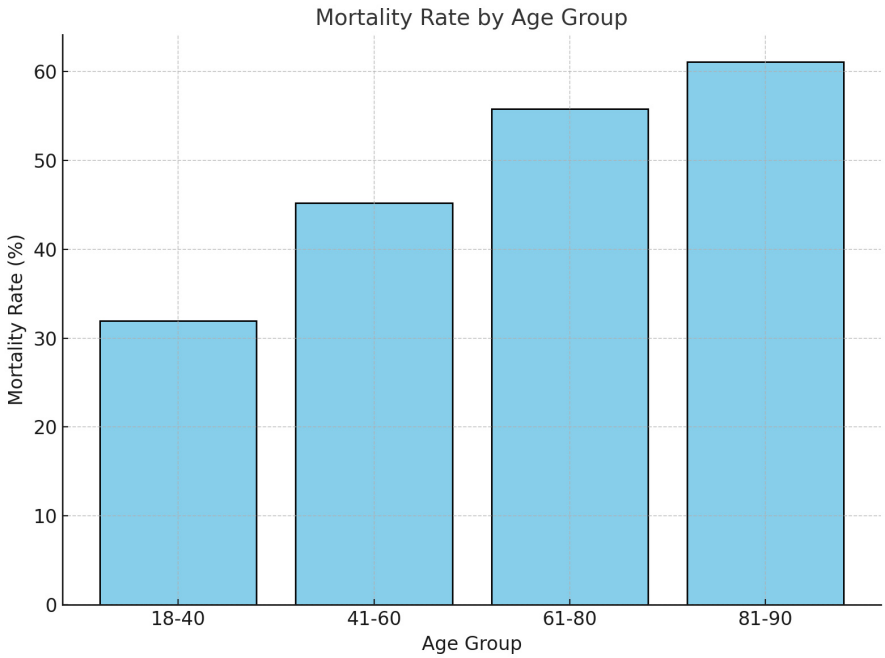

Fig. 3. Mortality rate by age group. This bar chart highlights the distribution of mortality rates across different age groups. The mortality rate increases with age, starting from 31.97% in the 18-40 age group to 61.11% in the 81-90 age group.

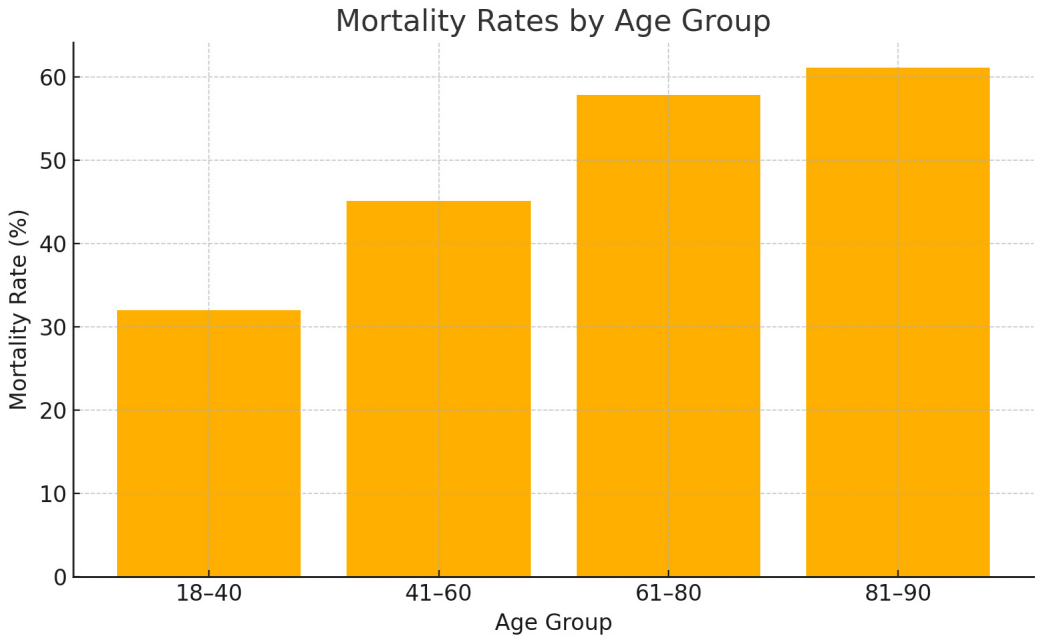

Fig. 4. Correlation matrix. The heatmap shows the correlations between key variables. There is a strong negative correlation between oxygen saturation and mortality, while heart rate and respiratory rate have moderate positive correlations with mortality. This provides insight into the relationships among variables influencing mortality risk.

sensitivity, specificity, and ROC-AUC. Random Forest and Neural Networks also exhibited strong performance but required longer training times, while SVM showed comparatively lower accuracy and sensitivity (Fig. 2).

Subgroup analysis revealed that mortality rates increased with age, rising from 31.97% in the 18-40 age group to 61.11% in the 81-90 age group. Low oxygen saturation, low blood pressure, and high respiratory rate were critical predictors across all age groups (Fig. 3).

The dataset comprised 1,500 patients, with a mean age of 52.87±20.86 years. Among these, the overall mortality rate was 52.40%, with survivors accounting for 47.60%. A detailed analysis of mortality rates by age groups and patient diagnoses was performed to provide more context to the results.

Mortality Rates by Age Groups

•18-40 years: Mortality rate of 31.97%.
•41-60 years: Mortality rate of 45.12%.
•61-80 years: Mortality rate of 57.83%.
•81-90 years: Mortality rate of 61.11%.

The results indicate a significant increase in mortality rates with advancing age, demonstrating the age-related risk factors in emergency department settings (Fig. 4).

Mortality Rates by Patient Diagnoses

Patient diagnoses were categorized into the following groups:

1. Cardiovascular diseases: Mortality rate of 65.4%.
2. Respiratory diseases: Mortality rate of 58.7%.
3. Infectious diseases: Mortality rate of 42.3%.
4. Trauma-related cases: Mortality rate of 28.9%.

These categories highlight the variation in mortality risks across different clinical conditions (Fig. 5).

Key Predictors of Mortality

The Gradient Boosting model identified the following variables as the most significant predictors:

1. Oxygen saturation (low levels strongly correlated with mortality).
2. Age (higher age groups associated with increased mortality risk).
3. Heart rate (elevated rates showed moderate positive correlation with mortality).

The CatBoost algorithm demonstrated superior performance compared to other models, achieving an accuracy of 88.8%, an F1 score of 85.0%, and a ROC-AUC of 0.911. Subgroup analysis further validated the model's reliability in predicting mortality across various demographics and clinical presentations.

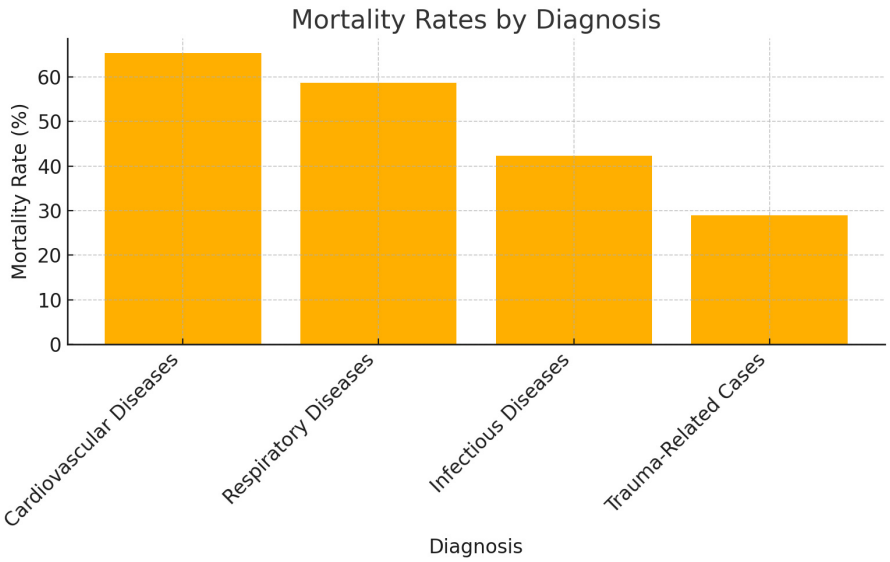

Fig. 5. **Mortality rates by patient diagnoses.**

Enhanced Data Presentation

To support the findings, additional tables and visualizations summarizing patient characteristics, age groups, and diagnoses with their corresponding mortality rates have been included. This approach narrows the broad term "mortality," providing a clear connection between patient profiles and model predictions.

The CatBoost model demonstrated high reliability and accuracy in predicting mortality risk, particularly highlighting the importance of oxygen saturation, blood pressure, and respiratory rate. These findings provide a strong foundation for the implementation of predictive models to enhance clinical decision support systems in emergency department settings (Fig. 6).

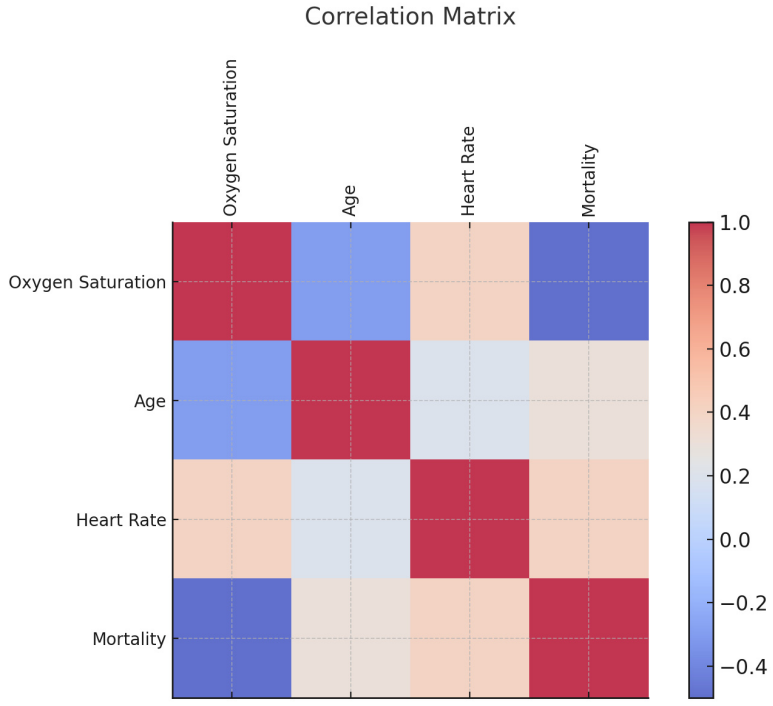

Fig. 6. **Correlation matrix.**

DISCUSSION

This study demonstrated the effectiveness of Gradient Boosting algorithms, particularly CatBoost, in predicting mortality risk in emergency department patients. The CatBoost model achieved a high F1 score (85.0%) and AUC (0.911), consistent with previous findings that underscore the strong performance of Gradient Boosting in healthcare prediction tasks [10, 11].

The model identified low oxygen saturation, low blood pressure, and high respiratory rate as the most significant predictors of mortality. These results align with prior studies that emphasized the prognostic importance of hypoxia and respiratory compromise in critically ill patients [12, 13]. For example, Topol [13] emphasized the role of key physiological indicators such as oxygen saturation and respiratory rate in enhancing predictive performance in emergency care through artificial intelligence applications.

Among the evaluated algorithms, CatBoost outperformed Random Forest and Neural Networks, confirming previous literature suggesting that Gradient Boosting methods offer superior generalization and accuracy in complex, high-dimensional clinical datasets [11, 14]. While Neural Networks demonstrated competitive performance, their interpretability and training demands remain limitations in emergency settings [15].

A key challenge in the study was class imbalance, with a mortality prevalence of 52.4%. To mitigate this, we applied synthetic oversampling and weighted loss functions, improving the model's sensitivity and F1 score. These findings are supported by broader literature emphasizing the necessity of addressing imbalance for reliable healthcare prediction [16, 17].

The study also revealed limitations associated with false negative predictions. These misclassifications often stem from borderline physiological values or multivariate complexities not fully captured by the model. Such errors could delay critical interventions, echoing concerns raised in previous studies that emphasized the risks of delayed care and the importance of transparency and reliability in clinical AI applications [10, 18].

In terms of clinical implications, integrating such predictive models into triage systems could enhance early risk identification, resource allocation, and patient management. However, the model’s dependence on retrospective data and its single-center design limit its generalizability. Therefore, external validation on multicenter datasets is needed to strengthen confidence in model applicability [19].

Future directions should prioritize real-time integration of models into hospital information management systems (HIMS), requiring robust APIs, low-latency data pipelines, and adaptive imputation strategies for missing data. Developing clinician-friendly dashboards that present predictions with SHAP-based explanations would also enhance usability and trust [20, 21].

Interdisciplinary collaboration is essential for aligning ML developments with clinical workflows. Moreover, ensuring privacy via federated learning and regulatory-compliant encryption protocols will be critical for ethical deployment [22]. Ethical concerns, such as algorithmic bias and accountability, must also be proactively addressed to prevent unintended harm [23].

Pilot studies in real-world clinical environments should be conducted to assess operational feasibility and collect clinician feedback. This iterative evaluation process will ensure that models are not only technically accurate but also practically beneficial for emergency department decision-making.

Limitations

This study is limited by its retrospective, single-center design, which may affect the generalizability of the findings. The dataset, although large, represents a specific patient population, and external validation is necessary to ensure broader applicability. Another limitation is the potential for bias in the training data, which may influence model predictions. Additionally, the performance of the model in real-time clinical settings remains untested. Integration into HIMS and assessment in prospective studies are required to evaluate its true clinical utility.

CONCLUSION

This study demonstrated that Gradient Boosting algorithms, particularly CatBoost, can serve as effective tools for predicting mortality risk in emergency department patients. The model showed high accuracy and reliability, highlighting the importance of key predictors such as oxygen saturation, age, and heart rate.

These findings support the potential integration of machine learning-based models into clinical decision support systems to aid early risk identification and optimize patient management. However, challenges such as false negative predictions and data imbalance remain critical areas for improvement. Future research should focus on real-time implementation, interdisciplinary collaboration, and iterative validation to ensure clinical applicability and ethical deployment of such predictive tools.

Ethics Approval and Consent to Participate

The study was approved by the Medipol University Non-Interventional Clinical Research Ethics Committee (Decision no.: 1138 and date: 28.11.2024). It was conducted in accordance with the ethical standards established in the Declaration of Helsinki and all data were anonymized and used solely for scientific purposes.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: EB; Study Design: EB; Supervision: EB; Funding: EB; Materials: EB; Data Collection and/or Processing: EB; Statistical Analysis and/or Data Interpretation: EB; Literature Review: EB; Manuscript Preparation: EB and Critical Review: EB.

Conflict of interest

The author disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the author(s) and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Wiler JL, Welch S, Pines JM, Schuur JD, Jouriles NJ, Stone-Griffith S. Emergency department performance measures updates: proceedings of the 2014 Emergency Department Benchmarking Alliance Consensus Summit. Acad Emerg Med. 2015;22(5):542-553. doi: 10.1111/acem.12654.
2. Hsia RY, Sarkar N, Shen YC. Impact of emergency department crowding on care for severe pain and patient satisfaction. J Emerg Med. 2019;57(3):291-298. doi: 10.1016/j.jemermed.2018.12.023.
3. Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. Nat Med. 2019;25(1):24-29. doi: 10.1038/s41591-018-0316-z.
4. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. Proc 22nd ACM SIGKDD Int Conf Knowl Discov Data Min. 2016;785-794. doi: 10.1145/2939672.2939785.
5. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30:4768-4777. doi: 10.48550/arXiv.1705.07874.
6. van Walraven C, Wong J, Forster AJ. Predicting postdischarge death or readmission: derivation and validation of a clinical prediction tool. BMC Med Inform Decis Mak. 2012;12:89. doi: 10.1186/1472-6947-12-89.
7. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med. 2018;24(11):1716-1720. doi: 10.1038/s41591-018-0213-5.
8. Sezik S, Cingiz MÖ, İbiş E. Machine Learning-Based Model for Emergency Department Disposition at a Public Hospital. Appl Sci. 2025;15(3):1628. doi: 10.3390/app15031628.
9. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med. 2016;375(13):1216-1219. doi: 10.1056/NEJMp1606181.
10. Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. Sci Data. 2016;3:160035. doi: 10.1038/sdata.2016.35.
11. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med. 2019;380(14):1347-1358. doi: 10.1056/NEJMra1814259.
12 Goldstein BA, Navar AM, Carter RE. Moving beyond regres-

sion techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. Eur Heart J. 2017;38(23):1805-1814. doi: 10.1093/eurheartj/ehw302.
13. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56. doi: 10.1038/s41591-018-0300-7.
14. Beam AL, Kohan IS. Big data and machine learning in health care. JAMA. 2018;319(13):1317-1318. doi: 10.1001/jama.2017.18391.
15. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. IEEE J Biomed Health Inform. 2018;22(5):1589-1604. doi: 10.1109/JBHI.2017.2767063.
16. Holzinger A, Biemann C, Pattichis CS, Kell DB. What do we need to build explainable AI systems for the medical domain? 2017. arXiv:1712.09923. doi: 10.48550/arXiv.1712.09923.
17. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: Review, opportunities and challenges. Brief Bioinform. 2018;19(6);1236-1246. doi: 10.1093/bib/bbx044.
18. Amann J, Blasimme A, Vayena E, Frey D, Madai VI; Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020;20(1):310. doi: 10.1186/s12911-020-01332-6.
19. Sendak MP, Ratliff W, Sarro D, et al. Real-World Integration of a Sepsis Deep Learning Technology Into Routine Clinical Care: Implementation Study. JMIR Med Inform. 2020;15;8(7):e15182. doi: 10.2196/15182.
20. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Medicine. 2019;17(1):195. doi: 10.1186/s12916-019-1426-2.
21. Wiens J, Saria S, Sendak M, et al. Do no harm: A roadmap for responsible machine learning for health care. Nat Med. 2019;25;1337-1340. doi: 10.1038/s41591-019-0548-6.
22. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use. Proceedings of the 4th Machine Learning for Healthcare Conference. Proceedings of Machine Learning Research. 2019;106:359-380. doi: 10.48550/arXiv.1905.05134.
23. Lin SY, Mahoney MR, Sinsky CA. Ten Ways Artificial Intelligence Will Transform Primary Care. J Gen Intern Med. 2019;34(8):1626-1630. doi: 10.1007/s11606-019-05035-1.

The European Research Journal 2025;11(6):1213-1226 Original Article

DOI: https://doi.org/10.18621/eurj.1794309 *Neurology and Artificial Intelligence*

Artificial intelligence-based handwriting analysis for non-invasive multiple sclerosis detection: A preliminary study

Yelda Fırat[1], Meral Seferoğlu[2], Hakan Kılıçaslan[2], Ali Özhan Sıvacı[2], Murat Kaan Yılmaz[1], Yılmaz Kılıçaslan[1]

[1]Department of Computer Engineering, Faculty of Engineering, Architecture and Design, Mudanya University, Bursa, Türkiye; [2]Department of Neurology, University of Health Sciences, Bursa Yüksek Ihtisas Training and Research Hospital, Bursa, Türkiye

ABSTRACT

Objectives: Multiple sclerosis (MS) is a chronic central nervous system disorder that causes demyelination, inflammation, and axonal damage, leading to permanent disabilities in motor, sensory, visual, and balance functions. This study aimed to develop an artificial intelligence (AI)-based, non-invasive diagnostic approach for MS detection using handwriting analysis, leveraging deep learning methods to identify disease-specific handwriting patterns.

Methods: A classification model was designed using a convolutional neural network (CNN) based on the VGG16 architecture with transfer learning. The dataset consisted of 426 handwriting samples, including 213 from MS patients and 213 from healthy individuals. Data augmentation and early stopping techniques were employed to improve model generalization capability.

Results: The proposed model achieved a validation accuracy of 83.72% and a test accuracy of 85%, indicating its robustness in distinguishing MS patients from healthy subjects. The confusion matrix analysis demonstrated a sensitivity of 86% and a specificity of 84%, indicating moderate discriminatory performance.

Conclusions: The findings suggest that the developed AI-based model offers an effective, non-invasive diagnostic tool for MS detection. This approach provides a promising foundation for future research on monitoring disease progression and developing clinically applicable AI-supported diagnostic systems.

Keywords: Multiple sclerosis, handwriting, VGG16, Transfer learning, convolutional neural network

Multiple sclerosis (MS) is a chronic, autoimmune, demyelinating disease affecting the central nervous system (CNS) [1-9]. The disease is characterized by inflammation and neurodegeneration processes [10-12], with diverse clinical outcomes based on the location and severity of the damage [13]. Common clinical symptoms of MS include limb weakness, optic neuritis, sensory deficits, double vision, dizziness, bladder dysfunction, chronic fatigue, balance and coordination disturbances, gait difficulties, cognitive impairment, and depression [14].

Accurate diagnosis of MS is critical as it provides

Received: September 30, 2025 **Accepted:** October 30, 2025 **Available Online:** .October 31, 2025 **Published:** November 4, 2025

How to cite this article: Fırat Y, Seferoğlu M, Kılıçaslan H, Sıvacı AÖ, Yılmaz MK, Kılıçaslan Y. Artificial intelligence-based handwriting analysis for non-invasive multiple sclerosis detection: A preliminary study. Eur Res J. 2025;11(6):1213-1226. doi: 10.18621/eurj.1794309

Corresponding author: Yelda Fırat, PhD., Assist. Prof., Phone: +90 224 224 20 22, E-mail: yelda.firat@mudanya.edu.tr

Available Online at https://dergipark.org.tr/en/pub/eurj

opportunities for early intervention [15], which can significantly influence the disease's progression [16]. Since there is no definitive laboratory parameter for diagnosing MS [17, 18], the 2017 McDonald criteria [19, 20] are widely utilized in the diagnostic process. These criteria integrate clinical evaluations, imaging findings, and laboratory data to standardize the diagnostic procedure [21, 16]. Magnetic Resonance Imaging (MRI) is particularly prominent in MS diagnosis and monitoring its progression, being one of the most effective methods available [22-25]. However, similar lesions caused by other diseases can sometimes appear in MRI scans, and errors in MRI interpretation are possible. For this reason, artificial intelligence (AI)-based methods have been proposed to identify early signs of MS and perform differential diagnoses of similar neurological conditions [26-29]. Despite the growing use of imaging and clinical data in MS diagnosis, a critical gap remains in the development of non-invasive, AI-powered biomarkers tailored specifically for MS. The question of whether AI, particularly through machine learning and its subfield deep learning, can accelerate the MS diagnostic process has become increasingly relevant [2, 30-33]. In recent years, transfer learning, a method that adapts pre-trained deep learning models to new tasks, has gained widespread adoption. This approach allows for high performance with less data and shorter training times. While AI studies have made significant progress in MS diagnosis, a review of the literature reveals no studies employing handwriting analysis as a basis for MS detection.

In addition to imaging-based approaches, motor performance and fine motor coordination are increasingly recognized as potential biomarkers of neurological dysfunction. Since MS affects both central motor pathways and visual feedback mechanisms, subtle impairments in handwriting can reflect disease-related motor, sensory, or cognitive changes. Studies on Parkinson's and Alzheimer's diseases have already shown that handwriting Dynamics - such as tremor amplitude, irregular stroke pressure, or inconsistent spacing - can serve as sensitive indicators of neural impairment. These findings from related neurological conditions provide a meaningful foundation to investigate whether handwriting dynamics could also reflect MS-related neurophysiological changes. Therefore, exploring handwriting as a non-invasive and quantifiable marker of MS-related motor dysfunction provides a logical and clinically relevant extension of AI-based diagnostic research.

In this study, we aimed to propose an innovative approach to MS diagnosis by developing a deep learning model based on handwriting image analysis.

METHODS

This cross-sectional study was conducted at the MS center of Bursa Yüksek İhtisas Training and Research Hospital, focusing on patients diagnosed with MS (Relapsing-Remitting Multiple Sclerosis [RRMS], Sec-

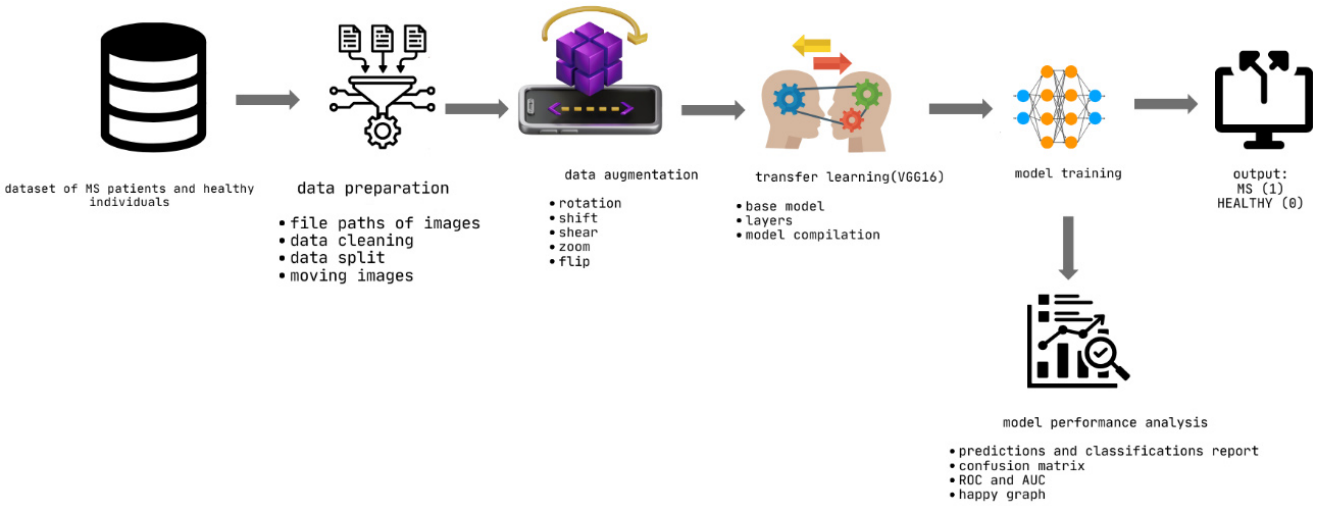

Fig. 1. Proposed model architecture.

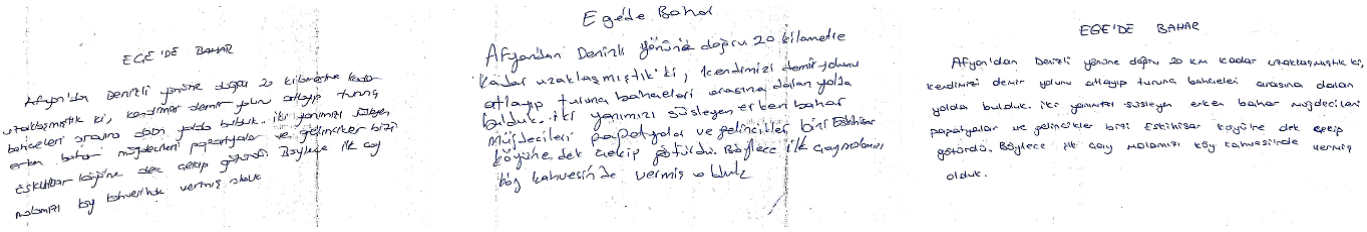

Fig. 2. Handwriting samples from healthy individuals in the MSHaW dataset.

ondary Progressive Multiple Sclerosis [SPMS], Primary Progressive Multiple Sclerosis [PPMS]) according to the 2017 McDonald criteria. The study was conducted from November to December 2024, involving consecutive patients and healthy controls aged 18 to 75 years who provided informed consent. Ethical approval was obtained from the local ethics committee (approval number: 2024-TBEK 2024/11-12), ensuring compliance with the Helsinki Declaration guidelines. All participants provided informed consent before inclusion, confirming their understanding and willingness to participate under clearly defined conditions. Illiterate individuals and those with an Expanded Disability Status Scale (EDSS) score of 7 or higher, patients with cognitive impairments that hindered the capacity to provide consent, uncontrolled psychiatric conditions were excluded from the study. Comprehensive data collection included participants' demographic, clinical data collected, including education, age, sex, EDSS score, and clinical subtype.

The dataset was created using handwriting samples collected from MS patients and healthy individuals. For model training, a transfer learning approach was employed, leveraging a VGG16-based convolutional neural network (CNN) architecture.

Fig. 1 provides an overview of the experimental methodology. As shown in Fig. 1, the model utilized handwriting data and incorporated transfer learning and data augmentation techniques for MS diagnosis. The VGG16-based CNN architecture produces a binary classification output, where "0" indicates a healthy individual, and "1" indicates MS. The steps involved in model development are outlined below.

Data Preparation and Preprocessing

After obtaining informed consent from patients monitored in the MS clinic, they were asked to copy a provided paragraph of text onto a plain white sheet of paper on a flat surface. Similarly, healthy volunteers who agreed to participate in the study were asked to write the same text. As a result, a dataset comprising 426 handwriting images was created, including 213 samples from healthy individuals and 213 from MS patients.

This dataset, named MSHaW (Multiple Sclerosis Handwriting Analysis and Writing dataset), serves as the primary data source for this study. Only handwriting images served as the foundation for developing deep learning models for MS diagnosis.

Fig. 2 shows handwriting samples from healthy individuals in the MSHaW dataset, while Fig. 3 presents handwriting samples from MS patients. The handwriting images in the MSHaW dataset, as shown in Figs. 2 and 3, underwent several steps of data preparation and preprocessing before being fed into the model.

First, the images in the dataset were divided into

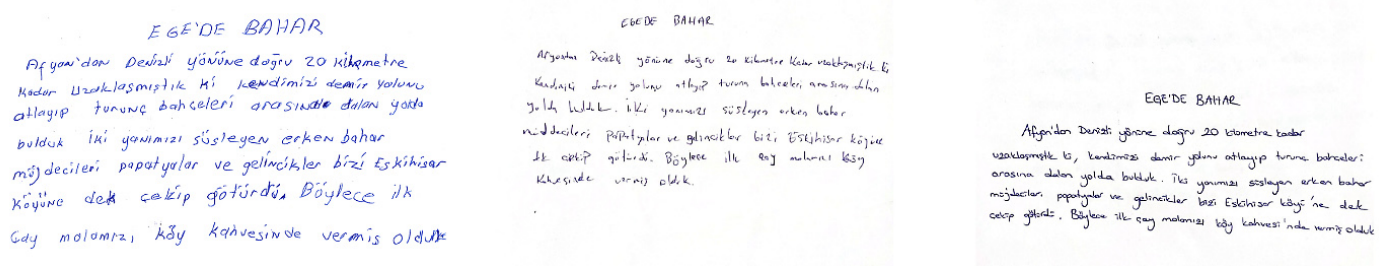

Fig. 3. Handwriting samples from multiple sclerosis patients in the MSHaW dataset.

two classes: those of patients with MS and those of healthy individuals. These classes were further subdivided into training and test subsets. In creating the dataset, 80% of the images were allocated for training and 20% for testing. During model training, 20% of the training data was reserved as a validation set for hyperparameter tuning and early stopping, resulting in an effective split of 64% training, 16% validation, and 20% test data. This procedure ensured that the model was trained and validated on distinct subsets to prevent data leakage and enhance generalizability. To facilitate accurate classification and processing, images belonging to each class were organized into appropriate directories.

After data preparation, the preprocessing phase began. One of the key steps in this process was image resizing. Since images of varying dimensions cannot be processed by the same model, all images needed to be resized to a fixed dimension compatible with the model's input layer. Accordingly, all images were resized to 128×128 pixels.

Moreover, data augmentation was a crucial part of the preprocessing phase. Due to the relatively small size of the dataset used for model development, the training dataset was enriched using techniques such as rotation, zooming, shifting, and horizontal flipping. To prevent overfitting, transformations were applied randomly rather than all at once, ensuring that each image underwent a unique combination of augmentations. Specific augmentations included rotation (up to 40 degrees), horizontal and vertical shifting (up to 30%), zooming (up to 30%), and shearing. Additionally, pixel values were normalized to the range of 0–1 to facilitate more effective model training.

To further enhance generalization, data augmentation was applied dynamically during each epoch, exposing the model to varying transformations in every iteration. This approach allowed the model to learn the underlying structure of the data rather than memorizing fixed patterns. By ensuring a diverse range of augmented samples throughout training, the model's robustness and adaptability were improved. These data preparation and augmentation processes were implemented as a critical preprocessing stage to enhance the model’s performance and generalization capacity.

Implementation Details

The model was implemented in Python 3.12 using TensorFlow and Keras deep learning frameworks. The following Python libraries were employed for data preprocessing, augmentation, and model training: *os, shutil, TensorFlow, Keras, scikit-learn, NumPy, Matplotlib, and glob.* Model training and evaluation were performed on a workstation equipped with an Intel Core i5-1235U CPU (1.40 GHz, 10 cores), 16 GB DDR4 RAM, an NVIDIA GeForce MX550 GPU (2 GB VRAM), and running Windows 11 Pro 64-bit operating system.

Utilizing the VGG16 Transfer Learning Model

The model used in this study is based on the VGG16 architecture, one of the transfer learning methods. VGG16 is an advanced deep learning model pre-trained on the ImageNet dataset. Its core feature is being a CNN capable of extracting meaningful features from images. With a deep network structure consisting of 16 layers, VGG16 exhibits a high ability to learn various visual features due to its training on a large dataset [1].

In this study, the core feature extraction layers of VGG16 were retained while its classification layers were redesigned. The base layers (convolutional and pooling layers) were frozen, halting their training to preserve the pre-learned visual features. On this foundational structure, a flatten layer was first added. The flatten layer converted the multidimensional tensor data from the convolutional and pooling layers into a one-dimensional vector suitable for fully connected layers. Subsequently, two fully connected layers with 256 and 128 neurons were added. To improve the generalization ability of the model and prevent overfitting, a dropout layer with a 60% dropout rate and an L2 regularization method were applied. Finally, a sigmoid activation function was added to the output layer, enabling binary classification into 0 (healthy) and 1 (diseased).

For this study, images were resized to 128×128 pixels to match the input layer requirements of the VGG16 model. The model was optimized using the Adam optimization algorithm with a learning rate of 0.00005. The loss function was set to binary crossentropy, which is suitable for binary classification tasks. During training, the early stopping method was employed to terminate the training process if validation loss did not improve for 10 consecutive epochs. To improve convergence and avoid local minima, the num-

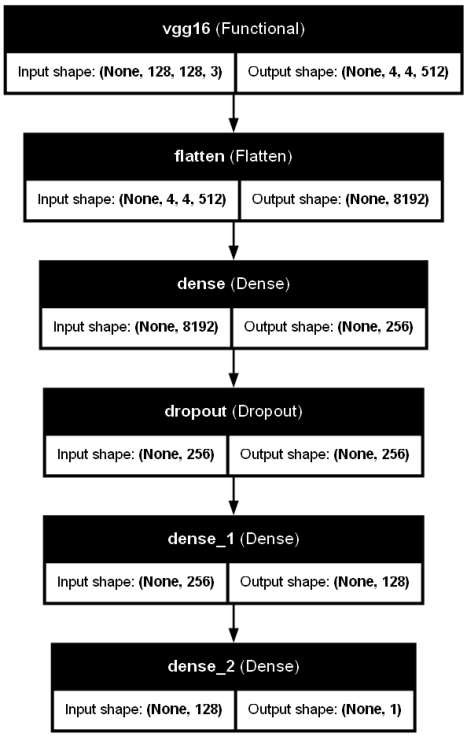

Fig. 4. Architecture of the VGG16-based CNN model.

ber of training epochs was increased from 100 to 150, and the early stopping patience value was adjusted from 10 to 15. These adjustments allowed the model to continue learning for a longer period while preventing overfitting. As a result, the model achieved a validation accuracy of 83.72% and a test accuracy of 85%, demonstrating its effectiveness of this approach. This structure allowed the VGG16 model to adapt its robust pre-learned visual features for detecting MS, achieving effective classification.

The architectural structure of the VGG16-based CNN model developed in this study is presented in Fig. 4.

Statistical Analysis

Accuracy, sensitivity, specificity, precision, F1-score, and area under the ROC curve (AUC) metrics were used to assess model performance. Sensitivity was calculated as the ratio of true positives (TP) to the total number of patients: Sensitivity = TP/(TP+FN), where FN represents false negatives. Specificity was calculated as the ratio of true negatives (TN) to the total number of healthy individuals: Specificity= TN/(TN+FP), where FP represents false positives. Precision (TP/[TP+FP]) and F1-score (harmonic mean of sensitivity and precision) were used for balanced performance evaluation. The 95% confidence intervals for binary classification metrics were calculated using the Wilson score method. Standard deviations for training and validation accuracies and missing values were obtained from epoch-to-epoch variations. The statistical significance of the difference between sensitivity and specificity was assessed using the McNemar test.

RESULTS

The sociodemographic and clinical information of the individuals from whom handwriting samples were taken were analysed. Among the patient group, 52% (n=111) were female and 48% (n=102) were male. Their age ranged from 17 to 71 years (median=40.54).

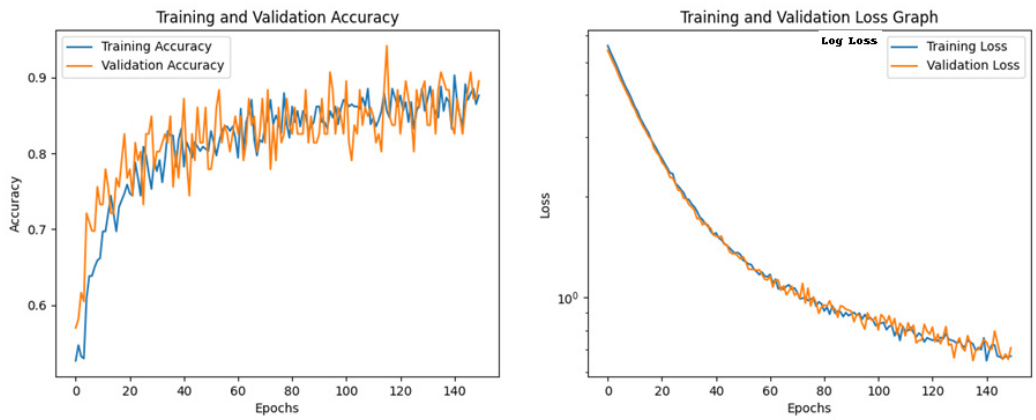

Fig. 5. Accuracy and loss curves of the model.

The educational levels were distributed as follows: 34.8% (n=74) primary school, 30.1% (n=64) high school, 31.4% (n=67) undergraduate, and 3.7% (n=8) postgraduate. The average EDSS score of the patient group was 2.37. The distribution of EDSS scores was as follows: 46.4% (n=99) had scores between 0-1.5, 30% (n=64) between 2-3.5, 23% (n=49) between 4-6.5, and 0.5% (n=1) scored 7 or higher. Among the patients, 87.8% (n=187) had RRMS, 6.6% (n=14) had PPMS, and 5.6% (n=12) had SPMS. In the healthy control group, 54% were female (n=115) and 46% were male (n=98). Their age ranged from 18 to 60 years (median=32.53). The educational levels were distributed as follows: 4.7% (n=10) primary school, 14.1% (n=30) high school, 73.2% (n=156) undergraduate, and 8% (n=17) postgraduate.

Model Performance

The performance of the model was thoroughly evaluated using both visual and numerical analyses obtained during the training and testing processes. Fig. 5 provides insights into the model's learning behavior through accuracy and loss curves, while Figs. 6 and 7 illustrate its classification capability via the confusion matrix and the receiver operating characteristic (ROC) curve, respectively. These combined visual and quantitative assessments allowed for a comprehensive evaluation of the model's effectiveness.

The progression of accuracy and loss values in the training and validation phases, as shown in Fig. 5, strongly indicates that the model underwent an effective learning process. The continuous increase in training accuracy, along with a parallel increase in validation accuracy, demonstrates the model's ability to generalize successfully not only to the training data but also to the validation data. This indicates a high generalization capacity and resilience against undesirable issues such as overfitting or underfitting.

Similarly, the consistent decrease in training and validation loss values illustrates the model's ability to optimize its performance by reducing error rates during the learning process. The early stopping method applied during training effectively identified the points where the model exhibited a tendency toward overfitting, allowing the training process to terminate at the appropriate time. Statistical analysis using Wilson score intervals revealed that the training accuracy was 82.71% ± 1.50% and validation accuracy was 83.72%

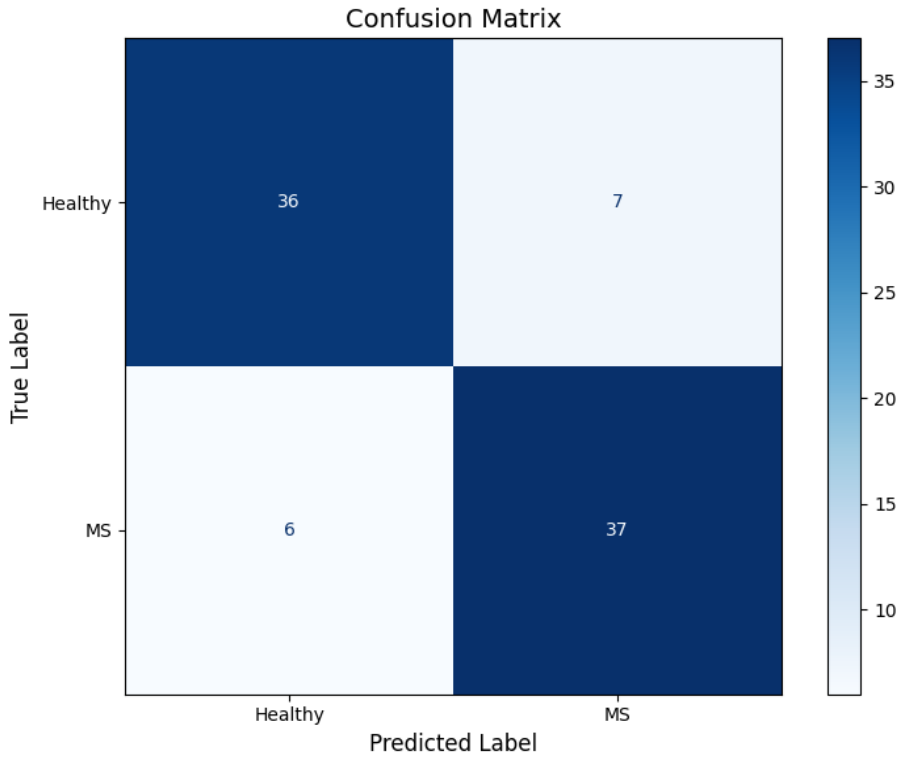

Fig. 6. Confusion matrix of the model.

±2.40% (95% CI: 75.84%, 90.95%), while training loss was 0.8901±0.0450 and validation loss was 0.8618±0.0520.

Fig. 5 illustrates the training and validation accuracy as well as the training and validation loss. The accuracy graph shows a stable increase in both training and validation accuracy, confirming that the model is learning effectively without overfitting. The log-scale loss graph demonstrates a consistent decrease in loss, highlighting the model's ability to steadily optimize its parameters and avoid local minima. The smooth decline in loss values further supports the model's convergence without instability.

The confusion matrix in Fig. 6 shows that the model has a strong discrimination capacity between healthy individuals and MS patients. The balanced rates of correctly classifying both classes indicate that the model does not focus on only one class and works equally effectively between the two classes. The low false positive and false negative rates are evidence that the model performs well in terms of both specificity and sensitivity.

Based on the confusion matrix, the model achieved a sensitivity of 86.05% (95% confidence interval [CI]: 72.74%, 93.44%) and specificity of 83.72% (95% CI: 70.03%, 91.88%), with precision of 84.09% (95% CI: 70.63%, 92.07%) and F1-score of 85.06% (95% CI: 76.04%, 91.08%).

To comprehensively evaluate the model's classification performance, multiple metrics were used, including accuracy, precision, recall (sensitivity), F1-score, specificity, and the area under the ROC curve (AUC). Accuracy represents the overall proportion of correctly classified instances. Precision measures the proportion of true positive predictions among all positive predictions, while recall (sensitivity) indicates the proportion of actual positive cases correctly identified by the model. The F1-score is the harmonic mean of precision and recall, providing a balanced measure when there is class imbalance. Specificity measures the proportion of true negative cases correctly identified, highlighting the model's ability to avoid false positives. Finally, the AUC value reflects the model's overall ability to discriminate between MS patients and healthy individuals, with higher values indicating better performance. These metrics provide a comprehensive understanding of the model's strengths and potential limitations in detecting MS from handwriting samples.

As seen in Fig. 7, the ROC curve and AUC analysis present the details of the model's classification capacity. The ROC curve visualizes the model's true positive rates against false positive rates, clearly demonstrating its ability to distinguish between classes. A curve that approaches the top-left corner indicates a high accuracy rate and low error rate, which can be readily observed visually. The AUC value rep-

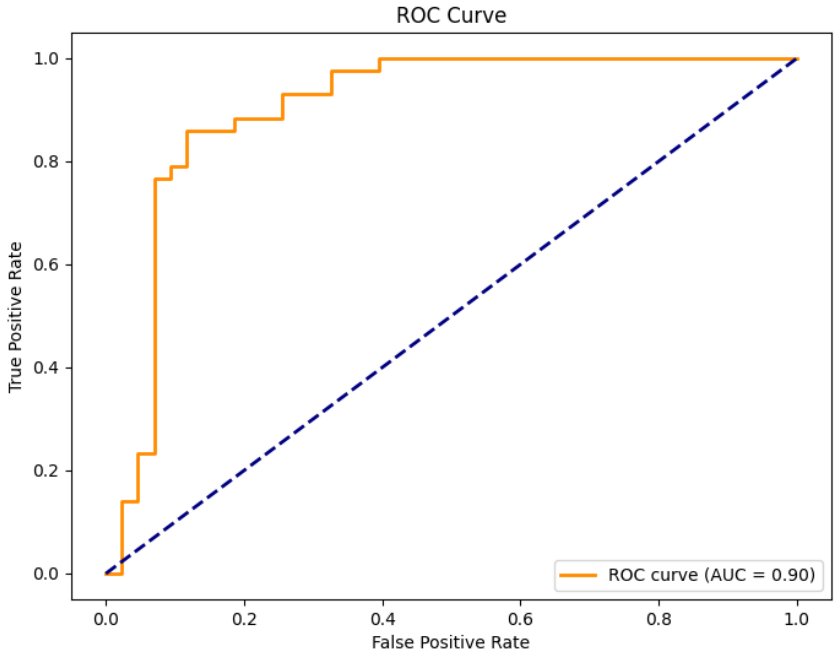

Fig. 7. ROC curve and AUC score of the model.

Table 1. Training and validation results

Metric	Training	Validation
Accuracy	82.71%	83.72%
Loss	0.8901	0.8618

resents the area under the ROC curve and serves as a measure of the model's overall classification capacity. The closer the AUC value is to 1, the better the model's ability to correctly distinguish between positive and negative classes. The ROC curve is also useful for understanding the impact of threshold values on model performance. Lower threshold values increase the detection of positive instances (higher sensitivity) but may result in a higher false positive rate. Conversely, higher threshold values reduce the false positive rate, thereby increasing specificity, but can lead to a risk of lower sensitivity. These analyses help determine the most suitable threshold value for specific applications while emphasizing that the model provides a reliable approach to diagnosing MS with precision and sensitivity. The ROC curve was generated using the validation dataset, ensuring that the performance assessment reflects the model's ability to generalize to unseen data.

Experimental Results

The results obtained from various metrics and visualizations in this study are shared in detail. Tables 1, 2, and 3 provide an in-depth analysis of the model's performance based on key evaluation metrics, including accuracy and loss values, the confusion matrix, and the ROC curve.

As shown in Table 1, the metrics from the model's training and validation processes indicate a well-balanced performance. The training accuracy was recorded at 82.71%±1.50%, while the validation accuracy reached 83.72%±2.40% (95% CI: 75.84%, 90.95%). These results demonstrate the ability of the model to effectively generalize what it learned from the training set to the validation set. The close alignment between training and validation accuracies highlights the absence of overfitting or underfitting, emphasizing the model's strong generalization capability and robust performance. When examining loss values, the training loss was calculated as 0.8901, and the validation loss was 0.8618. These low loss values indicate that the model effectively minimized errors across both datasets. These low loss values indicate that the model effectively minimized errors across both datasets. Validation accuracy is often considered a better metric for assessing how the model will perform in real-world scenarios, as it measures the model's effectiveness on unseen data. The alignment between training and validation accuracies also indicates the model's resilience to undesirable outcomes. For instance, if the training accuracy were significantly lower than the validation accuracy, it would suggest that the model was either insufficiently powerful or using inappropriate parameters, causing it to struggle even with the training data. Conversely, if the training accuracy were much higher than the validation accuracy, it would indicate overfitting, where the model adapts too closely to the training data at the expense of its generalization ability. However, our model appears to avoid both scenarios, demonstrating consistent performance. In this context, the obtained accuracy and loss values clearly indicate that the model underwent an effective learning process and possesses strong generalization capabilities.

Table 2 details the model's classification performance. The results presented in Table 2 were obtained using the validation dataset, ensuring that the reported performance reflects the model's ability to generalize to unseen data. The model correctly identified 37 MS patients but incorrectly classified 6 MS patients as healthy. Similarly, it accurately classified 36 healthy individuals as NOT MS but misclassified 7 healthy individuals as MS. These results suggest that the model performs effectively overall, demonstrating a balanced performance in detecting MS patients (recall) and distinguishing healthy individuals (specificity). However, further reduction of false positive and false negative rates is essential to prevent missed diagnoses of MS patients and to avoid unnecessary misdiagnoses. This

Table 2. Prediction results

		Predictions		
		MS	Not MS	Total
True	MS	37	6	43
	Not MS	7	36	43
	Total	44	42	86

MS=Multiple sclerosis

Table 3. Classification report

	Precision	Recall	F1-Score
Healthy	0.86	0.84	0.85
MS	0.84	0.86	0.85
Accuracy	-	-	0.85
Macro Avg	0.85	0.85	0.85
Weighted Avg	0.85	0.85	0.85

MS=Multiple sclerosis, Avg=Average

highlights the need for optimizing the model with larger and more diverse datasets to further enhance its performance. Recall and specificity were directly computed from Table 2. The model achieved a sensitivity of 86.05% (95% CI: 72.74%, 93.44%), meaning that it correctly identified 86% of actual MS patients. Similarly, the specificity was 83.72% (95% CI: 70.03%, 91.88%), indicating that the model correctly classified 84% of healthy individuals without mistakenly labeling them as MS patients. These results confirm the model's high capability in distinguishing between MS and healthy individuals and further validate its effectiveness in clinical applications.

Table 3 provides a detailed evaluation of the model's classification performance using various metrics. For healthy individuals, precision was calculated as 86%, recall as 84%, and the F1-score as 85%. Similarly, for MS patients, precision was 84%, recall 86%, and the F1-score also 85%.

Fig. 8 shows the composite performance metric graph of the model, illustrating its overall performance and generalization capacity across datasets of varying sizes.

The composite metric analysis, referred to as the composite performance metric graph, shown in Fig. 8, evaluates the model's overall performance across different datasets from a broader perspective. The composite performance metric graph was constructed by computing a composite metric score based on the weighted combination of accuracy, precision, recall, and F1-score for datasets of varying sizes. This analysis reveals that the model consistently achieves high performance and exhibits strong generalization capabilities across various datasets. The datasets contain 50, 100, 150, 200, and 213 samples, respectively, and the composite metric score increases steadily across these datasets, reaching a maximum score of 0.77. However, a performance dip is observed between 150 and 200 samples. This decline suggests that the model may have experienced temporary difficulty adapting to the diversity within the data groups or that the heterogeneity of the dataset challenged its generalization capacity. Despite this, the model's performance rebounded as the dataset size increased to 213 samples, ultimately achieving the maximum score. This im-

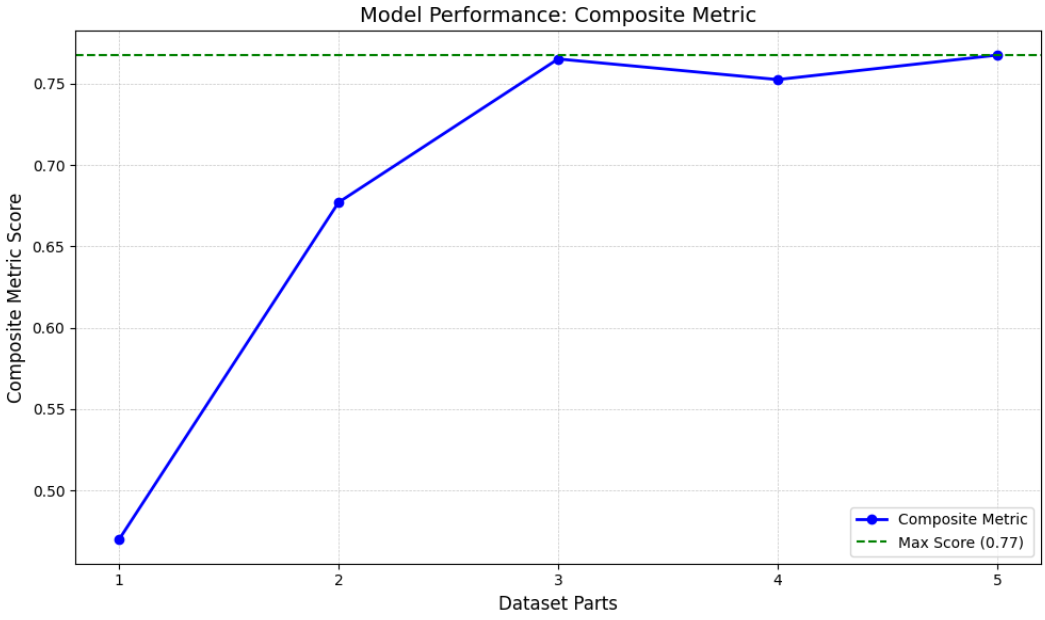

Fig. 8. Composite performance metric graph of the model.

provement highlights the model's ability to adapt to data diversity and optimize its overall performance. The findings from the composite metric analysis confirm that the model not only excels in terms of accuracy and F1-score but also effectively reflects its generalization capacity.

Considering the ROC curve in Fig. 7, the model's AUC value was determined to be 0.90 (95% CI: 0.83, 0.97). An AUC value of 0.90 indicates that the model has a 90% probability of correctly ranking a positive instance higher than a randomly chosen negative instance. This high AUC score underscores the model's balanced performance in terms of both sensitivity and specificity. The trajectory of the ROC curve that is near the top-left corner demonstrates a high true positive rate and a low false positive rate, providing clear evidence of the model's reliable and consistent classification ability.

DISCUSSION

AI techniques have become a vital tool in supporting physicians in disease diagnosis [34, 35]. Numerous studies have been conducted on AI applications in medicine and healthcare, offering opportunities to make disease diagnosis and management processes faster and more accurate [36-39]. In recent years, the use of machine learning and deep learning algorithms for detecting neurological diseases, such as MS, has seen significant growth [40-48]. Machine learning is a branch of AI that enables computers to make data-driven predictions and decisions without explicit programming. Deep learning, a subfield of machine learning, analyzes large datasets and solves more complex problems through layered artificial neural networks. AI models, especially those applied to clinical data such as MRI and cerebrospinal fluid (CSF) analyses, have become critical in medical processes by offering faster and more accurate diagnostic capabilities. Additionally, advanced techniques such as transfer learning, which allows for high performance even with limited datasets, have been widely adopted in the diagnosis of neurological diseases. Moreover, innovative approaches such as handwriting analysis [49] for diagnosing neurological disorders have also attracted growing attention in recent research.

AI-based MS diagnosis yields significant results using various data types. MRI-based studies achieve high accuracy rates, with Daqqaq *et al*. [50] reporting 98.81% accuracy using 2D and 3D CNN models with T2-weighted and FLAIR imaging, though the Dice Similarity Coefficient for segmentation remained at 63.78%. Lopatina *et al*. [51] achieved 95% accuracy using Susceptibility-Weighted Imaging and CNN, with explainable AI methods identifying critical brain regions, though limited dataset size challenged generalization. Vázquez-Marrufo *et al*. [52] comprehensively explored machine learning applications in MS, evaluating CNN, Support Vector Machine (SVM), Random Forest, and k-NN models, achieving 70.6-96% accuracy in MRI analyses and 61.8-92.9% for predicting disease progression, while highlighting that dataset heterogeneity limits generalizability. Montolío *et al*. [53] combined clinical data and Optical Coherence Tomography (OCT) images, achieving 87.7% accuracy with Ensemble Classifier and 81.7% for disease progression prediction using Long Short-Term Memory (LSTM), demonstrating effective integration of clinical data with deep learning algorithms.

Transfer learning has emerged as a powerful method for handling limited datasets. Jannat *et al*. [54] achieved 98.24% accuracy using VGG16-based transfer learning, while Shoeibi *et al*. [2] reported 96.88% accuracy with 2D-CNN-based transfer learning. Wahlig *et al*. [3] employed 3D U-Net with transfer learning for automatic MS lesion detection, achieving 74% sensitivity and 69% positive predictive value for active demyelination segmentation, demonstrating effectiveness with small datasets while revealing that dataset diversity and overfitting risk can limit generalization.

Handwriting analysis has gained attention as a non-invasive, cost-effective diagnostic method for neurological diseases [55, 56]. By assessing motor skill impacts through kinematic, mechanical, and visual features, this approach shows promise for early diagnosis and progression prediction. Öcal [57] achieved 97.14% accuracy for Alzheimer's diagnosis using handwriting-based ensemble models. Aouraghe *et al*. [58] and Huang *et al*. [59] analyzed handwriting data from Parkinson's patients using digital tablets and biometric pens, examining kinematic features (speed, acceleration), mechanical features (pressure), and visual features via CNNs with SVM classification, achieving 83% accuracy. However, limited dataset diversity and

task variation constrain generalization capacity.

We leveraged transfer learning to overcome small dataset challenges, extracting meaningful features from handwriting data and improving generalization. Data augmentation techniques (rotation, shifting, zooming) enhanced dataset diversity and model robustness. The model achieved 83.72% validation accuracy, 86.05% sensitivity, and 83.72% specificity, successfully distinguishing MS patients from healthy individuals. The model's accuracy, sensitivity, and specificity had 95% confidence intervals of 75.84–90.95%, 72.74–93.44%, and 70.03–91.88%, respectively. The area under the ROC curve (AUC) was 0.90 (95% CI: 0.83–0.97), confirming the model's high discriminative ability. Advancements in deep learning architectures and hyperparameter optimization could further enhance model performance, supporting widespread clinical application.

Strengths and Limitations

This study presents several notable strengths. To the best of our knowledge, the application of handwriting analysis to MS diagnosis has not been systematically investigated. Our study addresses this gap by developing a novel, non-invasive, and low-cost AI biomarker for MS detection through handwriting pattern analysis. We leveraged transfer learning to overcome challenges with small datasets, extracting meaningful features from handwriting data, and improving generalization. Data augmentation techniques (rotation, shifting, zooming) enhanced dataset diversity and model robustness. The model achieved 83.72% validation accuracy, 86.05% sensitivity, and 83.72% specificity, successfully distinguishing MS patients from healthy individuals. The 95% confidence intervals for accuracy, sensitivity, and specificity were 75.84–90.95%, 72.74–93.44%, and 70.03–91.88%, respectively. No significant difference was found between sensitivity and specificity ($P = 0.76$), and balanced accuracy was approximately 85%. This performance is comparable to MRI-based methods while offering significant advantages: handwriting analysis is non-invasive, requires no specialized equipment, incurs minimal cost, and can be easily conducted in patients' daily lives, making it particularly valuable for rural and resource-limited regions. Our findings demonstrate that handwriting analysis-based models can serve as reliable, accessible, and cost-effective tools for MS diagnosis, making a unique contribution as non-invasive AI biomarkers. This model could be utilized as a supportive system in initial evaluation, screening, and follow-up processes, accelerating clinical decision-making and promoting efficient use of healthcare resources.

However, this study also has limitations. Only handwriting images were included in the study; clinical and sociodemographic information, such as age, gender, educational level, EDSS scores, and MS types were not included in the evaluation. This decision was made to focus the model solely on handwriting classification and to adopt an approach that aligns with other studies. The small size of the dataset may have restricted the model's generalization capacity. This is particularly significant due to the limited diversity of handwriting samples collected from different individuals, which may hinder the model's ability to fully represent the complex variations encountered in real-world scenarios. A larger, more diverse dataset is anticipated to enhance the model's generalization and yield more robust, reliable results. Additionally, the large datasets used in transfer learning methods are often composed of natural images, which may be irrelevant to the domain of handwriting analysis. This mismatch can lead to the inadequacy of certain features when applied to handwriting analysis, particularly in capturing subtle motor variations. For instance, handwriting analysis focuses on features such as writing style, patterns, and disease-specific indicators, whereas natural images typically exhibit broader, more general visual patterns. This discrepancy may be a factor limiting the model's performance. Because our dataset came from a single centre in Turkey and uses only the Turkish alphabet, the handwriting styles and MS subtypes represented here may not generalise to other populations; multi-centre, multi-lingual external validation studies are therefore essential before clinical use. Although ImageNet-pretrained VGG16 provides a baseline, its features may not capture handwriting-specific nuances. Future work should investigate models trained from scratch on large handwritten datasets or pretrained on handwriting recognition tasks (e.g., Optical Character Recognition), which may yield more relevant representations. Future studies are encouraged to use larger, more di-

verse datasets. Enriching the dataset with samples from various age groups, cultures, and handwriting habits could enhance the model's generalization and increase its sensitivity to disease-specific features, such as those in MS.

CONCLUSION

Our study presents a preliminary study of an AI and deep learning-based approach for the diagnosis of MS. This proof-of-concept study demonstrates the feasibility of using handwriting analysis for MS detection, achieving 83.72% accuracy. However, validation in larger, multi-center cohorts is essential before clinical application can be considered. The study offers a non-invasive, cost-effective, and easily applicable diagnostic and monitoring method. The model's success was validated by results obtained during the training and validation processes, where it exhibited balanced performance across metrics such as sensitivity and specificity. Additionally, analyses of the confusion matrix and ROC curve confirmed the model's strong ability to distinguish between healthy individuals and MS patients. In the future, the use of larger and more diverse datasets, as well as the exploration of alternative deep learning architectures, could further enhance the model's performance.

Ethics Approval and Consent to Participate

This study was approved by the University of Health Sciences Bursa Yüksek Training and Research Hospital Medical Sciences Ethics Committee (Decision No: 2024-TBEK 2024/11-12; date: 06.11.2024). All procedures were conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. All participants provided informed consent before inclusion, confirming their understanding and willingness to participate under clearly defined conditions.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: MS; Study Design: YF, MS; Supervision: MS; Funding: N/A; Materials: HK, AÖS; Data Collection and/or Processing: HK, AÖS; Statistical Analysis and/or Data Interpretation: YF, MKY; Literature Review: YK, YF; Manuscript Preparation: YF, MKY; and Critical Review: YK, MS.

Conflict of Interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles. During the preparation of this work, the authors used ChatGPT to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Editor's Note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. ArXiv. 2014;1409.1556. doi: 10.48550/arXiv.1409.1556.
2. Shoeibi A, Khodatars M, Jafari M, et al. Applications of deep learning techniques for automated multiple sclerosis detection using magnetic resonance imaging: a review. Comput Biol Med. 2021;136:104697. doi: 10.1016/j.compbiomed.2021.104697.
3. Wahlig SG, Nedelec P, Weiss DA, Rudie JD, Sugrue LP,

Rauschecker AM. 3D U-Net for automated detection of multiple sclerosis lesions: utility of transfer learning from other pathologies. Front Neurosci. 2023;17:1188336. doi: 10.3389/fnins.2023.1188336.
4. Ekmekyapar T, Taşcı B. Exemplar MobileNetV2-Based Artificial Intelligence for Robust and Accurate Diagnosis of Multiple Sclerosis. Diagnostics (Basel). 2023;13(19):3030. doi: 10.3390/diagnostics13193030.
5. Ebers GC. Environmental factors and multiple sclerosis. Lancet Neurol. 2008;7(3):268-277. doi: 10.1016/S1474-4422(08)70042-5.
6. Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology. 2014;83(3):278-286. doi: 10.1212/WNL.0000000000000560.
7. Walton C, King R, Rechtman L, et al. Rising prevalence of multiple sclerosis worldwide: insights from the Atlas of MS, third edition. Mult Scler. 2020;26(14):1816-1821. doi: 10.1177/1352458520970841.
8. Oh J, Vidal-Jordana A, Montalban X. Multiple sclerosis: clinical aspects. Curr Opin Neurol. 2018;31(6):752-759. doi: 10.1097/WCO.0000000000000622.
9. Compston A, Coles A. Multiple sclerosis. Lancet. 2008;372(9648):1502-1517. doi: 10.1016/S0140-6736(08)61620-7.
10. Miller AE, Vermersch P, Kappos L, et al; TOPIC study group. Long-term outcomes with teriflunomide in patients with clinically isolated syndrome: Results of the TOPIC extension study. Mult Scler Relat Disord. 2019;33:131-138. doi: 10.1016/j.msard.2019.05.014.
11. García-Lorenzo D, Francis S, Narayanan S, Arnold DL, Collins DL. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. Med Image Anal. 2013;17(1):1-18. doi: 10.1016/j.media.2012.09.004.
12. Hartmann M, Fenton N, Dobson R. Current review and next steps for artificial intelligence in multiple sclerosis risk research. Comput Biol Med. 2021;132:104337. doi: 10.1016/j.compbiomed.2021.104337.
13. Harirchian MH, Fatehi F, Sarraf P, Honarvar NM, Bitarafan S. Worldwide prevalence of familial multiple sclerosis: A systematic review and meta-analysis. Mult Scler Relat Disord. 2018;20:43-47. doi: 10.1016/j.msard.2017.12.015.
14. Goldenberg MM. Multiple sclerosis review. P T. 2012r;37(3):175-814.
15. Kobelt G, Thompson A, Berg J, Gannedahl M, Eriksson J; MSCOI Study Group; European Multiple Sclerosis Platform. New insights into the burden and costs of multiple sclerosis in Europe. Mult Scler. 2017;23(8):1123-1136. doi: 10.1177/1352458517694432.
16. Kaisey M, Solomon AJ. Multiple Sclerosis Diagnostic Delay and Misdiagnosis. Neurol Clin. 2024;42(1):1-13. doi: 10.1016/j.ncl.2023.07.001.
17. Miller DH, Weinshenker BG, Filippi M, et al. Differential diagnosis of suspected multiple sclerosis: a consensus approach. Mult Scler. 2008;14(9):1157-1174. doi: 10.1177/1352458508096878.
18. Fangerau T, Schimrigk S, Haupts M, et al; Multiple Sclerosis Study Group. Diagnosis of multiple sclerosis: comparison of the Poser criteria and the new McDonald criteria. Acta Neurol Scand. 2004;109(6):385-389. doi: 10.1111/j.1600-0404.2004.00246.x.
19. Murray TJ. Diagnosis and treatment of multiple sclerosis. BMJ. 2006;332(7540):525-527. doi: 10.1136/bmj.332.7540.525.
20. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. 2018;17(2):162-173. doi: 10.1016/S1474-4422(17)30470-2.
21. Gill S, Agarwal M. Multiple Sclerosis Part 1: Essentials and the McDonald Criteria. Magn Reson Imaging Clin N Am. 2024;32(2):207-220. doi: 10.1016/j.mric.2023.11.002.
22. Wildner P, Stasiołek M, Matysiak M. Differential diagnosis of multiple sclerosis and other inflammatory CNS diseases. Mult Scler Relat Disord. 2020;37:101452. doi: 10.1016/j.msard.2019.101452.
23. Matias-Guiu JA, Díaz-Álvarez J, Ayala JL, et al. Clustering Analysis of FDG-PET Imaging in Primary Progressive Aphasia. Front Aging Neurosci. 2018;10:230. doi: 10.3389/fnagi.2018.00230.
24. Pontillo G, Cocozza S, Lanzillo R, et al. Determinants of Deep Gray Matter Atrophy in Multiple Sclerosis: A Multimodal MRI Study. AJNR Am J Neuroradiol. 2019;40(1):99-106. doi: 10.3174/ajnr.A5915.
25. Brownlee WJ, Hardy TA, Fazekas F, Miller DH. Diagnosis of multiple sclerosis: progress and challenges. Lancet. 2017;389(10076):1336-1346. doi: 10.1016/S0140-6736(16)30959-X.
26. Moetesum M, Siddiqi I, Vincent N, Cloppet F. Assessing visual attributes of handwriting for prediction of neurological disorders - A case study on Parkinson's disease. Pattern Recognit Lett. 2018;121:19-27. doi: 10.1016/j.patrec.2018.04.008.
27. Mohammadpoor M, Shoeibi A, Zare H, Shojaee H. A Hierarchical Classification Method for Breast Tumor Detection. Iran J Med Phys. 2016;13(4):261-268. doi: 10.22038/ijmp.2016.8453.
28. Alizadehsani R, Khosravi A, Roshanzamir M, et al. Coronary artery disease detection using artificial intelligence techniques: A survey of trends, geographical differences and diagnostic features 1991-2020. Comput Biol Med. 2021;128:104095. doi: 10.1016/j.compbiomed.2020.104095.
29. Amin M, Martínez-Heras E, Ontaneda D, Prados Carrasco F. Artificial Intelligence and Multiple Sclerosis. Curr Neurol Neurosci Rep. 2024;24(8):233-243. doi: 10.1007/s11910-024-01354-x.
30. Bonacchi R, Filippi M, Rocca MA. Role of artificial intelligence in MS clinical practice. Neuroimage Clin. 2022;35:103065. doi: 10.1016/j.nicl.2022.103065.
31. Sadeghi D, Shoeibi A, Ghassemi N, et al. An overview of artificial intelligence techniques for diagnosis of Schizophrenia based on magnetic resonance imaging modalities: Methods, challenges, and future works. Comput Biol Med. 2022;146:105554. doi: 10.1016/j.compbiomed.2022.105554.
32. Martín-Noguerol T, Paulano-Godino F, López-Ortega R, Górriz JM, Riascos RF, Luna A. Artificial intelligence in radiology: relevance of collaborative work between radiologists and engineers for building a multidisciplinary team. Clin Radiol. 2021;76(5):317-324. doi: 10.1016/j.crad.2020.11.113.
33. Vandewinckele L, Claessens M, Dinkla A, et al. Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance. Radiother Oncol. 2020;153:55-66. doi: 10.1016/j.radonc.2020.09.008.

34. Liu J, Pan Y, Li M, et al. Applications of deep learning to MRI images: a survey. Big Data Min Anal. 2018;1(1):1-18. doi: 10.26599/BDMA.2018.9020001.
35. Akkus Z, Galimzianova A, Hoogi A, Rubin DL, Erickson BJ. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. J Digit Imaging. 2017;30(4):449-459. doi: 10.1007/s10278-017-9983-4.
36. Huseyn E. Deep Learning Based Early Diagnostics of Parkinsons Disease. arXiv. 2020;2008.01792. doi: 10.48550/arXiv.2008.01792.
37. Shaban M. Deep Learning for Parkinson's Disease Diagnosis: A Short Survey. Computers. 2023;12(3):58. doi: 10.3390/computers12030058.
38. Tanveer M, Richhariya B, Khan RU, et al. Machine Learning Techniques for the Diagnosis of Alzheimer's Disease: A review. ACM Trans Multimed Comput Commun Appl. 2020;16(1):1-35. doi: 10.1145/3344998.
39. Zhou Z, Kanwal A, Chaturvedi K, Agrawal A, Nayan AA, Huang Y. Deep learning-based classification of neurodegenerative diseases using gait dataset: a comparative study. In: Proc Int Conf Robot Control Vis Eng. Tokyo, Japan; July 21-23, 2023:59-64. doi: 10.1145/3608143.3608154.
40. Chandaran SR, Muthusamy G, Sevalaiappan LR, Senthilkumaran N. Deep Learning-based Transfer Learning Model in Diagnosis of Diseases with Brain Magnetic Resonance Imaging. Acta Polytech Hung. 2022;19(5):127-147. doi: 10.12700/APH.19.5.2022.5.7.
41. Suganthe RC, Latha RS, Geetha M, Sreekanth GR. Diagnosis of Alzheimer's disease from brain magnetic resonance imaging images using deep learning algorithms. Adv Electr Comput Eng. 2020;20(3):57-64. doi:10.4316/AECE.2020.03007.
42. Swati ZNK, Zhao Q, Kabir M, et al. Brain tumor classification for MR images using transfer learning and fine-tuning. Comput Med Imaging Graph. 2019;75:34-46. doi: 10.1016/j.compmedimag.2019.05.001.
43. Sivaranjini S, Sujatha CM. Deep learning based diagnosis of Parkinson's disease using convolutional neural network. Multimed Tools Appl. 2020;79:15467-15479. doi: 10.1007/s11042-019-7469-8.
44. Lua S, Lua Z, Zhang YD. Pathological brain detection based on AlexNet and transfer learning. J Comput Sci. 2019;30:41-47. doi: 10.1016/j.jocs.2018.11.008.
45. Oh K, Chung YC, Kim KW, Kim WS, Oh IS. Classification and Visualization of Alzheimer's Disease using Volumetric Convolutional Neural Network and Transfer Learning. Sci Rep. 2019;9(1):18150. doi: 10.1038/s41598-019-54548-6.
46. Kim CM, Lee W. Classification of Alzheimer's disease using ensemble convolutional neural network with lfa algorithm. IEEE Access. 2023;11:143004-15. doi: 10.1109/ACCESS.2023.3342917.
47. Myszczynska MA, Ojamies PN, Lacoste AMB, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nat Rev Neurol. 2020;16(8):440-456. doi: 10.1038/s41582-020-0377-8.
48. Yakkundi A, Gupta R, Ramesh K, Verma A, Khan U, Ansari MA. Implications of Convolutional Neural Network for Brain MRI Image Classification to Identify Alzheimer's Disease. Parkinsons Dis. 2024;2024:6111483. doi: 10.1155/2024/6111483.
49. Chai J, Wu R, Li A, et al. Classification of mild cognitive impairment based on handwriting dynamics and qEEG. Comput Biol Med. 2023;152:106418. doi: 10.1016/j.compbiomed.2022.106418.
50. Daqqaq TS, Alhasan AS, Ghunaim HA. Diagnostic effectiveness of deep learning-based MRI in predicting multiple sclerosis: A meta-analysis. Neurosciences (Riyadh). 2024;29(2):77-89. doi: 10.17712/nsj.2024.2.20230103.
51. Lopatina A, Ropele S, Sibgatulin R, Reichenbach JR, Güllmar D. Investigation of Deep-Learning-Driven Identification of Multiple Sclerosis Patients Based on Susceptibility-Weighted Images Using Relevance Analysis. Front Neurosci. 2020;14:609468. doi: 10.3389/fnins.2020.609468.
52. Vázquez-Marrufo M, Sarrias-Arrabal E, García-Torres M, Martín-Clemente R, Izquierdo G. A systematic review of the application of machine-learning algorithms in multiple sclerosis. Neurologia (Engl Ed). 2023;38(8):577-590. doi: 10.1016/j.nrleng.2020.10.013.
53. Montolío A, Martín-Gallego A, Cegonino J, et al. Machine learning in diagnosis and disability prediction of multiple sclerosis using optical coherence tomography. Comput Biol Med. 2021;133:104416. doi: 10.1016/j.compbiomed.2021.104416.
54. Jannat SA, Hoque T, Supti NA, Alam MA. Detection of multiple sclerosis using deep learning. In: Proc Asian Conf Innov Technol (ASIANCON). Pune, India; Aug 27–29, 2021. doi: 10.1109/ASIANCON51346.2021.9544601.
55. Cilia ND, De Gregorio G, De Stefano C, Fontanella F, Marcelli A, Parziale A. Diagnosing Alzheimer's disease from on-line handwriting: A novel dataset and performance benchmarking. Eng Appl Artif Intell. 2022;111:104822. doi: 10.1016/j.engappai.2022.104822.
56. Sarin K, Bardamova M, Svetlakov M, et al. A three-stage fuzzy classifier method for Parkinson's disease diagnosis using dynamic handwriting analysis. Decis Anal J. 2023;8:100274. doi: 10.1016/j.dajour.2023.100274.
57. Öcal H. A novel approach to detection of Alzheimer's disease from handwriting: triple ensemble learning model. Gazi Univ J Sci Part C Des Technol. 2024;12(1):214-223. doi: 10.29109/gujsc.1386416.
58. Aouraghe I, Khaissidi G, Mrabti M. A literature review of online handwriting analysis to detect Parkinson's disease at an early stage. Multimed Tools Appl. 2023;82:11923-11948. doi: 10.1007/s11042-022-13759-2.
59. Huang Y, Chaturvedi K, Nayan A-A, Hesamian MH, Braytee A, Prasad M. Early Parkinson's Disease Diagnosis through Hand-Drawn Spiral and Wave Analysis Using Deep Learning Techniques. Information. 2024; 15(4):220. doi: 10.3390/info15040220.

The European Research Journal 2025;11(6):1227-1234 Original Article

DOI: https://doi.org/10.18621/eurj.1657532 *Obstetrics and Gynecology*

Prediction of placenta accreta spectrum in placenta previa surgery using systemic inflammatory markers

Ahmet Beyazıt[1], Kenan Serdar Dolapçıoğlu[1], Oya Soylu Karapınar[1], Bircan Kara[2], Hasan Cem Sevindik[1]

[1]Department of Obstetrics and Gynecology, Hatay Mustafa Kemal University, Tayfur Ata Sökmen Faculty of Medicine, Hatay, Türkiye; [2]Hatay Mustafa Kemal University, Health Training and Research Hospital, Hatay, Türkiye

ABSTRACT

Objectives: Despite advancements in medical science, existing methodologies for predicting placental invasion in patients with previa remain insufficient. Systemic inflammatory markers have emerged as potential prognostic indicators in oncology. This study aims to explore whether systemic inflammatory parameters can predict placenta accreta spectrum (PAS) in patients with previa, given the analogous mechanisms between PAS and cancer cell invasion.

Methods: This retrospective case-control study encompassed patients diagnosed with placenta previa who underwent cesarean section procedures between 2010 and 2024. The case group (PAS+) comprised patients who underwent hysterectomy, segmental resection (with histopathologically confirmed PAS), or had in situ placenta. Conversely, the control group (PAS-) included other patients. The systemic immune-inflammatory index (SII), platelet-to-lymphocyte ratio (PLR), and neutrophil-to-lymphocyte ratio (NLR) were calculated from the most recent preoperative complete blood count. A comparative analysis of demographic data, clinical findings, and systemic inflammatory markers was conducted between the groups.

Results: A total of 487 patients were analyzed, with 146 (30%) classified as PAS (+) and 341 (70%) as PAS (-). The PAS (+) group exhibited significantly higher values in terms of age, gravidity, parity, and the number of prior cesarean sections (P=0.004, P<0.001, P=0.018, and P<0.001; respectively). However, no significant correlation was identified between the two groups concerning systemic inflammatory markers.

Conclusions: Systemic inflammatory markers do not appear to be reliable for predicting placental invasion in previa patients.

Keywords: Placenta previa, placenta accreta spectrum, systemic immune-inflammatory index, inflammatory biomarkers

Obstetric hemorrhage, particularly postpartum hemorrhage, continues to be a significant global health concern, resulting in the death of one mother every minute [1]. In instances of postpartum hemorrhage, peripartum hysterectomy frequently emerges as the most efficacious life-saving surgical intervention when alternative measures, such as uterotonic agents, balloon tamponade, uterine compression sutures, and arterial ligation, prove ineffective [2].

Placenta Accreta Spectrum (PAS) and placenta

Received: March 15, 2025 **Accepted:** April 27, 2025 **Available Online:** June 4, 2025 **Published:** November 4, 2025

How to cite this article: Beyazıt A, Dolapçıoğlu KS, Soylu Karapınar O, Kara B, Sevindik HC. Prediction of placenta accreta spectrum in placenta previa surgery using systemic inflammatory markers. Eur Res J. 2025;11(6):1227-1234. doi: 10.18621/eurj.1657532

Corresponding author: Ahmet Beyazıt, MD., Assist. Prof., Phone: +90 326 229 10 00, E-mail: beyazitahmet@hotmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

previa are important causes of postpartum hemorrhage [3]. In cases of placenta previa, the incidence of hysterectomy due to postpartum hemorrhage is increased by a factor of 33.26 compared to cases with normally positioned placentas [4]. The rising rates of cesarean sections have led to a gradual increase in the incidence of placenta previa and PAS, resulting in more frequent occurrences of associated mortality and morbidities. [5]. It is important to recognize PAS cases prenatally in patients with placenta previa to provide optimal conditions for maternal and fetal care by referring patients to specialized centers, organizing the surgical-anesthesia team and to prevent possible complications. It is regrettably not always feasible to identify cases of PAS prenatally. Diagnostic methods such as ultrasound and magnetic resonance imaging (MRI) are not infallible, as they are subject to interobserver variability and can be complicated by factors such as a posterior placenta, patient obesity, and the quality of the imaging device [6]. Recent meta-analyses indicate that the diagnosis of PAS is established during labor in approximately 50% of patients. Even with the implementation of MRI screening, which is costly and not universally accessible, 25% of patients will still be diagnosed with PAS at the time of delivery [7].

Recent research has identified similarities between the invasive characteristics of trophoblast cells in patients with PAS and the metastatic properties of cancer cells [8]. Systemic inflammation markers, derived from complete blood count parameters, are widely recognized as indicators of inflammatory response and serve as prognostic factors in various cancers [9, 10]. These markers can be obtained by proportioning some parameters in the complete blood count and they are cost-effective, straightforward to interpret, and readily accessible.

This study aims to determine whether systemic inflammation markers, known for their prognostic value in cancer, can also predict invasion in patients with placenta previa, who exhibit characteristics similar to cancer invasion.

METHODS

In this retrospective case-control study, patients with placenta previa who underwent surgery at our university hospital, a referral center for high-risk pregnancies in southern Turkey, between 2010 and 2024, were included. Ethics committee approval was obtained (decision dated 20/11/2024, numbered 36). Informed consent was received from patients before surgeries. The study was conducted under the Declaration of Helsinki.

Singleton pregnant women diagnosed with placenta previa via transvaginal ultrasound and who subsequently underwent cesarean delivery at our clinic were included in the study. Exclusion criteria encompassed patients with chronic systemic diseases, active infections, anticoagulant use, or those who received blood transfusions during pregnancy. Additionally, individuals undergoing steroid therapy within the preceding three months, smokers, and those with multiple pregnancies were excluded. Pregnant women experiencing obstetric complications such as preeclampsia, intrauterine growth restriction (IUGR), or premature rupture of membranes were also excluded due to their potential impact on blood parameters. PAS was diagnosed in patients who had histopathological confirmation following surgical procedures such as hysterectomy or partial uterine resection. Additionally, patients whose placentas were left in situ due to intraoperative observation of complete placental invasion were classified as PAS (+). Both patients with histopathological confirmation and those with retained in situ placenta were included in the case group. The control group consisted of patients who underwent cesarean section for placenta previa but did not undergo hysterectomy/segmental uterine resection/placenta in situ.

Demographic data, operative reports, and hemogram results were collected from patient files and the hospital registration system. Complete blood count data from the last blood sample taken from the patient before surgery were used in the calculation of inflammation marker parameters. Systemic immune-inflammatory index (SII), platelet to lymphocyte ratio (PLR), and neutrophil to lymphocyte ratio (NLR) were analyzed. SII was calculated by the formula: platelet count × neutrophil count/lymphocyte count; NLR, neutrophil count/lymphocyte count; PLR, platelet count/lymphocyte count.

Statistical Analysis

Descriptive statistics included mean, standard deviation, median, minimum, maximum, frequency, and ratios. The distribution of variables was assessed using

the Kolmogorov-Smirnov and Shapiro-Wilk tests. Non-normally distributed quantitative data were analyzed using the Mann-Whitney U test. Qualitative independent data were evaluated using the chi-square test or Fisher's exact test, as appropriate. Statistical analyses were performed using SPSS 28.0.

Table 1. Demographic, obstetric and surgical data of patients

		Data
Age (years)		32.2±5.3 32 (19-44)
Gravidy		4.2±1.8 4 (1-14)
Parity		2.6±1.3 2 (0-9)
Prior cesarean	(-)	101 (20.7%)
	(+)	386 (79.3%)
Number of cesarean		2.4±1.0 2 (1-6)
Gestational week during labor (weeks)		36.4±2.1 37 (24-40)
Placenta Accreata Spectrum (PAS)	(-)	341 (70.0%)
	(+)	146 (30.0%)
Placental Invasion	(-)	341 (70.0%)
	Accreata	23 (4.7%)
	Increata	49 (10.1%)
	Percreata	74 (15.2%)
Surgical Intervention	(-)	341 (70.0%)
	Hysterectomy	115 (23.6%)
	Segmental resection	20 (4.1%)
	In situ placenta	2 (0.4%)
	Hysterectomy+packing	9 (1.8%)
Artery Ligation	(-)	295 (60.6%)
	Hypogastric artery	98 (20.1%)
	Uterine artery	54 (11.1%)
	Hypogastric+uterine arteries	40 (8.2%)
Complication	(-)	440 (90.3%)
	(+)	47 (9.7%)
	Ureter injury	12 (25.5%)
	Bladder injury	33 (70.2%)
	Intestinal serosal defect	2 (4.2%)
Mortality	(-)	485 (99.6%)
	(+)	2 (0.4%)

Data are shown as mean±standard deviation or median (minimum-maximum) or n (%)

RESULTS

During the study period, a total of 487 patients met the criteria and underwent surgical intervention for placenta previa at our hospital. The mean age of these patients was 32.2±5.3 years. The average number of pregnancies was 4.2±1.8, and the mean parity was 2.6±1.3. Additionally, the mean number of cesarean sections was 2.4±1, with 101 (20.7%) patients having no prior history of cesarean section.

In the study, 146 (30%) pregnant women were categorized into the PAS (+) group, while 341 (70%) were classified into the PAS (-) group. Among the hysterectomy specimens from the PAS group, 23 (4.7%) patients exhibited placenta accreta, 49 (10.1%) had placenta increta, and 74 (15.2%) had placenta percreta. Hysterectomy was performed on 115 (23.6%) patients, and segmental resection was conducted on 20 (4.1%) patients within the PAS group. In two (0.4%) cases, the placentas were left in situ, whereas 9 (1.8%) patients underwent a packing procedure in addition to hysterectomy. A total of 192 (39.4%) patients underwent arterial ligation, specifically of the hypogastric artery and/or uterine artery. Surgical complications (bladder injury, ureter injury, intestinal serosal defect) were observed in 47 (9.7%) patients, and there were 2 (0.4%) fatalities. Demographic, obstetric, and previous surgery-related data are presented in Table 1.

The variables of age, number of pregnancies, number of deliveries, and number of previous cesarean sections were significantly elevated in the PAS (+) group (P=0.004, P<0.001, P=0.018, and P<0.001; re-

Table 2. Comparison of demographic, obstetric, and surgical data between PAS (+) and PAS (-) groups

		PAS (-) (n=341)	PAS (+) (n=146)	P value
Age (years)		31.7±5.5 (32.0)	33.3±4.7 (34.0)	**0.004**[m]
Gravide		4.0±1.7 (4.0)	4.6±1.8 (4.0)	**<0.001**[m]
Parity		2.4±1.4 (2.0)	2.9±1.1 (3.0)	**0.018**[m]
Prior Cesarean	(-)	100 (29.3%)	1 (0.7%)	**<0.001**[X²]
	(+)	241 (70.7%)	145 (99.3%)	
Number of cesarean		2.2±1.0 (2.0)	2.6±1.0 (3.0)	**<0.001**[m]
Gestational week during labor		36.5±2.0 (37.0)	36.0±2.4 (37.0)	**0.003**[m]
Artery ligation				
(-)		234 (68.6%)	61 (41.8%)	**<0.001**[X²]
Hypogastric		29 (8.5%)	69 (47.3%)	
Uterine		48 (14.1%)	6 (4.1%)	
Hypogastric+uterine		30 (8.8%)	10 (6.8%)	
Complication	(-)	333 (97.7%)	107 (73.3%)	**<0.001**[X²]
	(+)	8 (2.3%)	39 (26.7%)	
Mortality	(-)	341 (100.0%)	144 (98.6%)	0.089[X²]
	(+)	0 (0.0%)	2 (1.4%)	

Data are shown as mean±standard deviation (median) or n (%). PAS= Placenta Accreata Spectrum.
[m]Mann-Whitney U test, [X²]Ki-kare test (Fischer test)

spectively). Additionally, the requirement for arterial ligation was markedly higher in the PAS (+) group compared to the non-PAS group (P<0.001). A greater incidence of complications was noted in the surgical procedures of the PAS (+) group (P<0.001). Notably, two patients succumbed, both of whom were in the PAS (+) group. Comparative obstetric, demographic, and surgical data between the groups are presented in Table 2.

Table 3 presents the hemogram parameters and systemic inflammation markers. The data indicate that inflammation markers do not have a significant impact on the determination of PAS.

DISCUSSION

Recently, the systemic inflammatory indicators SII, PLR, and NLR, derived from hemogram parameters, have garnered significant attention. Numerous studies have evaluated these markers, particularly in the context of prognostic prediction in cancer patients [11, 12]. Furthermore, within the discipline of obstetrics and gynecology, research has identified associations between these markers and the prognosis of patients experiencing ectopic pregnancy, miscarriage, and preeclampsia [13-15].

Despite differing in their developmental mechanisms, PAS and cancer exhibit analogous molecular and cellular processes that enable them to invade and proliferate within surrounding tissues. The intricate interaction of molecular signals, including growth factors, cytokines, and components of the extracellular matrix, is pivotal in regulating cell behavior. In both processes, it is essential to overcome local immunological mechanisms and activate angiogenesis [8]. The observed similarities between these two conditions indicate that systemic inflammatory markers, which are employed to predict prognosis in cancer patients, may also be applicable in predicting PAS status in patients with placenta previa.

Alterations in the maternal immune system are essential for a successful pregnancy and the accommodation of the semi-allogeneic fetus. Individuals with placenta previa exhibit an exaggerated immune response. Elevated levels of IL-1 beta, IL-6, TNF-alpha, and interferon-gamma are observed in the blood of patients with placenta previa [16]. The pathogenesis of

Table 3. Comparison of blood count parameters and inflammation markers between PAS (+) and PAS (-) groups

	PAS (-) (n=341)	PAS (+) (n=146)	P value
Leukocyte ($\times10^9$)	10.2±2.8 (10.0)	10.3±2.9 (9.7)	0.993[m]
Lymphocyte (10^3/µL)	2.0±1.0 (1.9)	1.9±0.6 (1.8)	0.740[m]
Neutrophil (10^3/µL)	7.6±2.6 (7.1)	7.6±2.6 (7.2)	0.884[m]
Platelet ($\times10^9$)	225.5±66.3 (215.0)	235.9±87.1 (225.0)	0.405[m]
PLR	129.4±57.8 (118.6)	135.7±59.3 (122.4)	0.265[m]
NLR	4.5±3.1 (3.8)	4.6±3.0 (4.0)	0.425[m]
SII	1,017.8±708.3 (815.3)	1,046.8±659.1 (874.0)	0.223[m]

Data are shown as mean±standard deviation (median) or n (%). NLR=Neutrophil/Lymphocyte ratio, PAS= Placenta Accreata Spectrum, PLR=Platelet/lymphocyte ratio, SII=Systemic Inflammatory Index

[m]Mann-Whitney U test

PAS encompasses processes of angiogenesis, inflammation, and invasion [17]. These findings suggest an elevated level of inflammation, particularly in patients with placenta previa and PAS (+). In the laboratory context, this would typically manifest as increased inflammation markers in the complete blood count. However, our study did not reveal a significant correlation in this regard.

In our study, we determined that markers of systemic inflammation may not reliably predict cases of PAS. A review of the literature reveals some studies that contradict our findings. For instance, Kele *et al*. [18] conducted a comparison involving 68 histologically confirmed PAS (+) patients and 205 control patients who underwent surgery for placenta previa without any additional procedures. Their results indicated that the parameters SII, NLR, and PLR were significantly elevated in the PAS group [18]. In our study, patients who underwent additional interventions, such as the Bakri balloon, intrauterine suture, compression suture, Lynch suture, uterine artery ligation, and hypogastric artery ligation, were included in the PAS (-) group. In contrast, Keles *et al*.'s study [18] entirely excluded patients who underwent additional procedures. These differences in patient inclusion criteria may account for the divergent results observed between the studies. In their study, Abide-Yayla *et al*. [19] investigated the findings of placental invasion via ultrasound in a cohort of 146 pregnant women. The research identified a significant NLR among 46 patients in whom placental abruption (PAS) was confirmed through histopathological analysis. Conversely, no significant association was observed for the PLR [19]. However, Abide-Yayla conducted this study on patients with placenta previa who were admitted to the hospital due to bleeding. It is possible that the complete blood count parameters were influenced by antepartum hemorrhage at the time of admission.

Firstly, while contemporary automated hematology analyzers are widely utilized, these instruments may produce inaccurate results, particularly in instances such as platelet aggregation. Furthermore, the fasting state also influences hemogram parameters. An elevation in postprandial lipids is manifested as an increase in neutrophil counts in blood values [20]. Seasonal variations and hormonal fluctuations associated with diurnal stress may influence lymphocyte levels, potentially altering inflammation parameters [21, 22].

Sometimes, even if there is placental invasion, hysterectomy or segmental resection is not performed because bleeding is controlled with methods such as intracavitary suture and compression sutures. However, these patients were considered as PAS (-) in our study because there was no histopathologic evidence of PAS. Furthermore, patients who underwent in situ placenta were classified as PAS (+) based on intraoperative observations. While these patients are unequivocally PAS, this method may result in diagnostic heterogeneity and potential misclassification. These situations may have introduced a bias in the study. Nevertheless, the substantial number of patients included in the study serves to strengthen its overall conclusions. To our knowledge, our study is the study with the largest number of patients in the literature.

Limitations

One limitation of our study is its retrospective nature. Additionally, the absence of detailed histopathological information in PAS patients constitutes another limitation. Furthermore, the lack of intraoperative data, such as perioperative blood loss, also represents a constraint.

CONCLUSION

In conclusion, while inflammatory markers present advantages due to their cost-effectiveness and accessibility, their reliability in predicting PAS is limited. They may not be dependable in patients with placenta previa, as they are prone to significant variability. A more rigorously designed study on this topic is warranted.

Ethics Approval and Consent to Participate

This study was approved by the Hatay Mustafa Kemal University Non-Interventional Clinical Research Ethics Committee (Decision no.: 36, date: 20.11.2024). All procedures performed during data collection, review of patient records, and study implementation complied with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Informed consent was waived because of the retrospective nature of the study and the analysis used anonymous clinical data.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: AB, KSD; Study Design: AB, KSD, OSK; SupervisionKSD; Funding: N/A; Materials: AB, HCS, BK; Data Collection and/or Processing: AB, HCS, BK; Statistical Analysis and/or Data Interpretation: AB, BK; Literature Review: AB, HCS; Manuscript Preparation: AB, KSD, OSK and Critical Review: AB, OSK, KSD.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Horng HC, Lai MJ, Chang WH, Wang PH. Placenta accreta spectrum (PAS) and peripartum hysterectomy. Taiwan J Obstet Gynecol. 2021;60(3):395-396. doi: 10.1016/j.tjog.2021.03.001.
2. Dogan O, Pulatoglu C, Yassa M. A new facilitating technique for postpartum hysterectomy at full dilatation: Cervical clamp. J Chin Med Assoc. 2018;81(4):366-369. doi: 10.1016/j.jcma.2017.05.010.
3. Committee on Practice Bulletins-Obstetrics. Practice Bulletin No. 183: Postpartum Hemorrhage. Obstet Gynecol. 2017;130(4):e168-e186. doi: 10.1097/AOG.0000000000002351.
4. Crane JM, Van den Hof MC, Dodds L, Armson BA, Liston R. Maternal complications with placenta previa. Am J Perinatol. 2000;17(2):101-105. doi: 10.1055/s-2000-9269.
5. Cahill AG, Beigi R, Heine RP, Silver RM, Wax JR; Society of Gynecologic Oncology; American College of Obstetricians and Gynecologists and the Society for Maternal–Fetal Medicine. Placenta Accreta Spectrum. Am J Obstet Gynecol. 2018;219(6):B2-B16. doi: 10.1016/j.ajog.2018.09.042.
6. Soyer P, Morel O, Tsatsaris V, Bourgioti C, Barat M. Placenta Accreta Spectrum: A Continuously Evolving Challenge for Radiologists. Can Assoc Radiol J. 2021;72(4):597-598. doi: 10.1177/0846537120984124.
7. Tinari S, Buca D, Cali G, et al. Risk factors, histopathology and diagnostic accuracy in posterior placenta accreta spectrum disorders: systematic review and meta-analysis. Ultrasound Obstet Gynecol. 2021;57(6):903-909. doi: 10.1002/uog.22183.
8. Rekowska AK, Obuchowska K, Bartosik M, et al. Biomolecules Involved in Both Metastasis and Placenta Accreta Spectrum-Does the Common Pathophysiological Pathway Exist? Cancers (Basel). 2023;15(9):2618. doi: 10.3390/cancers15092618.
9. Feng Z, Wen H, Bi R, et al. Preoperative Neutrophil-to-Lymphocyte Ratio as a Predictive and Prognostic Factor for High-Grade Serous Ovarian Cancer. PLoS One. 2016;11(5):e0156101. doi: 10.1371/journal.pone.0156101.
10. Akın MN, Kasap BH, Yuvacı HU. Neutrophil-to-lymphocyte ratio and platelet distribution in patients with endometrial cancer. J Obstet Gynaecol Res. 2015;41(9):1499. doi: 10.1111/jog.12723.
11. Huang Y, Chen Y, Zhu Y, et al. Postoperative Systemic Immune-Inflammation Index (SII): A Superior Prognostic Factor of Endometrial Cancer. Front Surg. 2021;8:704235. doi: 10.3389/fsurg.2021.704235.
12. Han X, Liu S, Yang G, et al. Prognostic value of systemic hemato-immunological indices in uterine cervical cancer: A systemic review, meta-analysis, and meta-regression of observational studies. Gynecol Oncol. 2021;160(1):351-360. doi: 10.1016/j.ygyno.2020.10.011.
13. Dinc K, Issın G. Novel marker to predict rupture risk in tubal ectopic pregnancies: the systemic immune-inflammation index. Ginekol Pol. 2023;94(4):320-325. doi: 10.5603/GP.a2023.0010.
14. Turgut E, Yildirim M, Sakcak B, Ayhan SG, Tekin OM, Sahin D. Predicting miscarriage using systemic immune-inflammation index. J Obstet Gynaecol Res. 2022;48(3):587-592. doi: 10.1111/jog.15156.
15. Conde-Rico ET, Naves-Sánchez J, González AP, Luna-Anguiano JLF, Paque-Bautista C, Sosa-Bustamante GP. [Inflammatory indexes and their association with the severity of preeclampsia]. Rev Med Inst Mex Seguro Soc. 2023;61(Suppl 2):S178-S184. [Article in Spanish]
16. Li L, Zhang J, Gao H, Ma Y. Nestin is highly expressed in foetal spinal cord isolated from placenta previa patients and pro-

motes inflammation by enhancing NF-κB activity. Biomarkers. 2018;23(6):597-602. doi: 10.1080/1354750X.2018.1468824.
17. Biswas R, Sawhney H, Dass R, Saran RK, Vasishta K. Histopathological study of placental bed biopsy in placenta previa. Acta Obstet Gynecol Scand. 1999 Mar;78(3):173-9.
18. Keles A, Dagdeviren G, Yucel Celik O, et al. Systemic immune-inflammation index to predict placenta accreta spectrum and its histological subtypes. J Obstet Gynaecol Res. 2022;48(7):1675-1682. doi: 10.1111/jog.15254.
19. Abide Yayla C, Ozkaya E, Tayyar A, Senol T, Senturk MB, Karateke A. Predictive value of complete blood count parameters for placental invasion anomalies. J Matern Fetal Neonatal Med. 2017;30(19):2324-2328. doi: 10.1080/14767058.2016.1247266.
20. Van Oostrom AJ, Sijmonsma TP, Rabelink TJ, Van Asbeck BS, Cabezas MC. Postprandial leukocyte increase in healthy subjects. Metabolism. 2003;52(2):199-202. doi: 10.1053/meta.2003.50037.
21. Dhabhar FS, Miller AH, Stein M, McEwen BS, Spencer RL. Diurnal and acute stress-induced changes in distribution of peripheral blood leukocyte subpopulations. Brain Behav Immun. 1994;8(1):66-79. doi: 10.1006/brbi.1994.1006.
22. Maes M, Stevens W, Scharpé S, et al. Seasonal variation in peripheral blood leukocyte subsets and in serum interleukin-6, and soluble interleukin-2 and -6 receptor concentrations in normal volunteers. Experientia. 1994;50(9):821-829. doi: 10.1007/BF01956463.

The European Research Journal 2025;11(6):1235-1243 Review

DOI: https://doi.org/10.18621/eurj.1677704 *Allied Health and Rehabilitation Science*

Artificial intelligence in assessment and intervention of speech and language disorders: A literature review

Eren Balo, Batuhan Ökte, Semra Selvi Balo

Department of Speech and Language Therapy, Anadolu University Faculty of Health Sciences Eskişehir, Türkiye

ABSTRACT

Artificial intelligence (AI) is a broad term that refers to the use of computers to replicate intelligent behavior with minimal human intervention. AI is rapidly transforming various sectors, including speech and language pathology, by offering innovative solutions to enhance therapeutic practices and client outcomes. Its application in speech and language pathology spans several domains, including medical diagnosis, therapeutic planning, and rehabilitation, utilizing tools such as machine learning and deep learning to enhance data analysis and pattern recognition. The primary aim of this study is to provide resources for speech and language pathologists on the topic of artificial intelligence by presenting research findings on the assessment and intervention of speech and language disorders using AI. Accordingly, AI studies in speech and language pathology found in the literature were included. The results of these studies were summarized, and information was provided on the use of AI in assessing and treating speech and language disorders, including swallowing disorders, voice disorders, acquired language disorders, motor and speech sound disorders, cleft palate speech, and developmental language disorder. Existing literature acknowledges and supports the growing popularity of AI and AI-based algorithms in speech and language pathology. Although the current evidence remains insufficient and concerns about ethics and implementation persist, advancing technology offers promise for applying AI in this field.

Keywords: Artificial intelligence, speech and language pathology, speech and language pathologist, voice disorders, speech disorders

Artificial intelligence (AI) is a broad term that refers to the use of computers to replicate intelligent behavior with minimal human intervention [1]. It is commonly regarded to have begun with the development of robots. Officially established as a field in 1956, artificial intelligence involves the science and engineering of creating intelligent machines. AI is applied in various areas, including medical diagnosis, medical statistics, and human biology. It is also incorporated into rehabilitation practices by various healthcare professionals, including social workers, occupational therapists, audiologists, nurses, and speech and language pathologists (SLPs) [2]. With rapid advances in technology, AI has established itself as a transformative force in various fields, including medical imaging and diagnostics. Through algorithms such as machine learning and deep learning, it has the potential to analyze complex data, identify patterns, and deliver diagnostic and prognostic insights that exceed human capabilities in terms of both speed and accuracy [3, 4].

Artificial intelligence is posited to significantly

Received: April 16, 2025 **Accepted:** June 16, 2025 **Available Online:** June 21, 2025 **Published:** November 4, 2025

How to cite this article: Balo E, Ökte B, Selvi Balo S. Artificial intelligence in assessment and intervention of speech and language disorders: A literature review. Eur Res J. 2025;11(6):1235-1243. doi: 10.18621/eurj.1677704

Corresponding author: Eren Balo, PhD-SLP., Assist. Prof., Phone: +90 (222) 335 05 80 ext. 2173, E-mail: ebalo@anadolu.edu.tr

Available Online at https://dergipark.org.tr/en/pub/eurj

augment clinical practices within speech-language pathology by developing innovative tools to enhance client health and optimize therapeutic outcomes. As technological advancements progress rapidly, SLPs must develop a comprehensive understanding of artificial intelligence clinical tools, acknowledge their inherent limitations, and employ them judiciously in clinical practice as they evolve [5].

This traditional review aims to present research on the role, significance, and applications of AI in speech and language pathology. It explores how AI is utilized in various subtypes of speech and language disorders. The review has been conducted through a comprehensive examination of current articles in the literature. The data collection process began by identifying the following keywords: artificial intelligence in speech and language disorders, artificial intelligence in speech and language therapy, artificial intelligence in speech and language assessment, advantages of artificial intelligence for speech and language therapists, disadvantages of artificial intelligence for speech and language therapists, and review. These keywords were entered into the Google search engine, and related publications were examined by reviewing American Speech-Hearing Association (ASHA), ResearchGate, Google Scholar, YÖK Academic, PubMed, and Anadolu University Library databases. National and international studies conducted between 2009 and 2025 were included in the study, whereas studies conducted before 2009 were excluded.

ARTIFICIAL INTELLIGENCE IN SPEECH AND LANGUAGE PATHOLOGY

The ASHA evaluated the advantages of AI for SLPs, categorizing the findings into four primary domains [6]. Firstly, the assessment revealed that documentation and reporting—essential components for tracking and monitoring therapeutic outcomes—can be significantly streamlined by implementing various AI algorithms and technologies. This improvement enables pathologists to allocate more time to client interactions during therapy sessions, thereby enhancing overall productivity. Secondly, the investigation underscored the utility of AI within advanced assistive technologies specifically designed to aid individuals with communicative and cognitive disorders. Thirdly, the findings indicated that SLPs have the capacity to incorporate objective and validated AI tools within clinical examinations and evaluations. For instance, a simple picture identification app on a smartphone can be used to assess remote language and speech skills. Lastly, the study highlighted that these technological advancements could enhance diagnostic and therapeutic processes, tailoring them to fit the genetic, behavioral, social, cultural, and economic profiles of patients, while also considering their individual responses to interventions.

A recent study confirmed that AI-based tools reduce the workload of SLPs and provide various conveniences to enhance their productivity [7]. Speech-to-text translation programs and automated report preparation processes have been shown to decrease administrative burdens, allowing SLPs to spend more time with their clients. Proper session preparation and planning enable the creation of therapy materials tailored to each client's needs, making therapies more effective and focused. These tools help SLPs utilize their time more efficiently and significantly support their professional processes. Furthermore, SLPs view the use of AI with a mix of optimism and caution. In a survey on the topic, pathologists indicated that they found AI tools particularly useful for diagnosis and treatment planning (50%) and rehabilitation (25%). However, only a small percentage (14.8%) believed that AI could replace professional services. Overall, they believe that AI will revolutionize the field without negatively affecting employment. Many pathologists also reported that they frequently use platforms such as ChatGPT [8].

Although SLPs have various opinions on the use of AI, it is preferred in many speech-language disorders, including swallowing disorders, voice disorders, acquired language disorders, motor and speech sound disorders, cleft palate speech, and developmental language disorder. The following subheadings include studies on the use of AI in assessment and therapy processes regarding these disorders.

AI in Swallowing Disorders

In the field of speech and language pathology, several studies have utilized AI in therapy and assessment processes, particularly concerning swallowing disorders. One study focused on the use of AI-supported video games for swallowing rehabilitation. These games exhibited a notable enhancement in swallowing

function among patients suffering from dysphagia after a stroke. Additionally, they increased oral intake and boosted patients' self-confidence by visualizing their movements during exercises. Moreover, the integration of AI and video games enhanced nutritional status by reducing complications such as weight loss and fatigue, ultimately improving quality of life. In the initial stages, rapid improvements were noted in the group using AI-supported video games. However, over the long term, both interventions resulted in similar outcomes [9]. In a similar study, surface electromyography biofeedback (sEMG-BF) combined with gaming was found to significantly enhance swallowing function in individuals with post-stroke dysphagia [10]. AI and AI-assisted tools have been effectively utilized not only in dysphagia therapy but also in its assessment. The results of a study aimed at automating the diagnosis of dysphagia through the development of an artificial intelligence-based web application, which analyzes Videofluoroscopic Swallowing Study (VFSS) data and offers clinicians more accurate evaluation opportunities, demonstrated that the YOLOv7 algorithm used in the study achieved high accuracy rates in dysphagia classification. This system can provide significant support to clinicians in patient care management [11]. A recent study aimed to evaluate the efficacy of using ultrasound as a quantitative method for assessing the kinematics of the hyoid bone during swallowing, with the goal of predicting Penetration-Aspiration Scale (PAS) scores. The researchers further explored the capabilities of a machine learning (ML) algorithm in forecasting PAS outcomes. The findings revealed that manual ultrasound measurements, when employed in isolation, did not provide a sufficiently accurate prediction of PAS scores. In contrast, the ML algorithm demonstrated a notable level of precision in predicting these results, suggesting its potential as a more reliable tool for clinical assessment in this context [12]. Sanjeevi *et al*. [13] compiled research on the use of AI techniques in VFSS analysis. They highlighted advancements in several areas, including the analysis of swallowing phases, segmentation of anatomical components, and the detection of penetration-aspiration events. However, the authors noted significant limitations in the research within this field, such as the fact that AI models often operate under specific assumptions, the lack of transparency and interpretability, and the absence of comprehensive and publicly accessible datasets [13]. Kim *et al*. [14] utilized convolutional neural networks (CNNs) to determine the presence of aspiration with high accuracy in 190 participants with dysphagia. Similarly, Iida *et al*. [15] used CNNs to detect aspiration in nearly 18,000 images, demonstrating that deep learning has the potential to detect aspiration with precision. Bandini *et al*. [16] studied specific time points in the pharyngeal phase, focusing on the moments when the bolus crosses the mandible and when the upper esophageal sphincter closes. CNN-based methods demonstrated high accuracy in detecting these moment measurements [16]. Hsiao *et al*. [17] developed a CNN-based algorithm for the automatic tracking and kinematic analysis of hyoid bone motion during swallowing. This algorithm calculates the precise positioning, direction, and velocity of the hyoid bone relative to the anatomical axis. Consequently, clinicians can conduct assessments more objectively and efficiently [17]. Nakamori *et al*. [18] demonstrated that assessing swallow sounds in patients with amyotrophic lateral sclerosis (ALS) using an electronic stethoscope and AI analysis showed a significant correlation with established swallow assessment parameters. They suggested that this method could serve as a new assessment tool suitable for home and remote medical care [18]. Despite this cutting-edge development in dysphagia management, there are also some limitations. Girardi *et al*. [19] emphasized that AI-assisted analyses have the potential to increase accuracy, speed, and efficiency in dysphagia diagnosis and treatment, but require large data sets and interdisciplinary collaborations.

AI in Voice Disorders

Another area where AI is commonly utilized in speech and language pathology is the management of voice disorders. A comprehensive review of the existing literature suggests that AI technologies are primarily used for diagnosing various voice disorders and for distinguishing between dysphonic and non-dysphonic voices. However, there are relatively few studies that classify the types of voice disorders or evaluate them using the GRBAS scale [20]. Although it was emphasized that AI technology has significant potential in detecting voice pathologies, it was indicated that clinical validation studies, data standardization, and consistent reporting methods are necessary to advance

research in this field. Kojima *et al*. [21] aimed to develop standardized methods for directly assessing pathological voice quality using one-dimensional convolutional neural network (1D-CNN) models. The findings revealed that these models were comparable in reliability to expert assessments using the GRBAS scale. In another study, a system was presented that can distinguish between healthy and pathological voices using machine learning algorithms and perform reliable voice disorder detection, operating entirely on a mobile device [22]. A similar study aimed to develop a reliable mobile health system that can intelligently classify healthy and pathological voices using machine learning algorithms and was shown to have the highest accuracy in detecting voice disorders [23]. In their study, Constantini *et al*. [24] analyzed voice properties for the early diagnosis of Parkinson's disease (PD) using ML techniques and identified voice biomarkers that could differentiate between healthy individuals, PD patients diagnosed early and not taking medication, and PD patients in the intermediate-advanced stage receiving L-Dopa treatment. In their study, Hegde *et al*. [25] examined various databases, feature extraction methods, and machine learning approaches focused on the automatic detection of voice disorders. They emphasized that these techniques can significantly enhance patients' quality of life by enabling the early identification of voice disorders [25]. Al-Hussain *et al*. [26] evaluated the effectiveness of ML algorithms in screening and diagnosing voice disorders and found that ML-based systems achieved high accuracy (93%), sensitivity (96%), and specificity (93%). In a study focused on detecting various vocal fold disorders through the recognition of pathological voice types using artificial intelligence, voice samples were collected from 189 individuals with normal voices and 552 individuals with voice disorders. These disorders included vocal atrophy, unilateral vocal fold paralysis, organic vocal fold lesions, and adductor spasmodic dysphonia. A convolutional neural network model was developed, achieving a sensitivity of 0.66, a specificity of 0.91, and an overall accuracy of 66.9% in distinguishing between normal voices and the mentioned disorders. Comparing the accuracy with the judgments of voice specialists, the overall accuracy rates were 60.1% and 56.1% for the two laryngologists and 51.4% and 43.2% for two general otolaryngologists [27]. Kim *et al*. [28] developed a model that distinguishes between voice samples of healthy individuals, patients with laryngeal cancer, and those with other laryngeal diseases using artificial intelligence. They tested various Mel-Frequency Cepstral Coefficient (MFCC) transformation methods and machine learning techniques, achieving an accuracy rate of 85% to 97% in identifying laryngeal diseases compared to healthy voices [28]. Similarly, another study examined the role of ML techniques in the diagnosis and monitoring of voice disorders and found that PD was the most frequently studied disease [29]. While the findings reveal a growing interest in ML-based voice analysis studies, the datasets used are limited and unbalanced. Research has focused on diagnosis, but insufficient attention has been paid to monitoring the disease process. The study emphasizes that these shortcomings should be addressed in future research.

AI in Acquired Language Disorders

Artificial intelligence is increasingly being utilized in neurological disorders such as aphasia and dementia. One notable study focused on an AI program that translates text into images for patients with aphasia. In this study, 189 out of 200 target texts (94.5%) were successfully visualized to convey the key concepts. However, many of the visualizations had aesthetic flaws, which could impact their effectiveness. Nouns were visualized with the highest efficiency and accuracy, followed by verbs, while visualizing complete sentences proved to be more challenging. Consequently, the ability of AI to quickly generate low-cost, high-quality images is considered a significant advancement in the assessment and treatment of aphasia [30]. Another comprehensive review study on aphasia examined different methods for an automatic speech assessment system that classifies the severity of aphasia [31]. Both AI and deep learning models were used for classification. CNN, recurrent neural networks, and hybrid models were reported to yield better results than traditional algorithms. In a systematic review of the use of AI in assessing speech and language skills to predict cognitive decline in Alzheimer's disease, promising results were reported in almost all 51 studies; however, few have been implemented in clinical research or practice [32].

Grasemann *et al*. [33] developed a neural network model called BiLex, which includes two separate phonetic maps and a common semantic map, to predict

the development of language skills in bilingual aphasia patients. The study revealed that the BiLex model was able to successfully predict the development of language skills in the treated language but had lower prediction accuracy in the untreated language and tended to underestimate cross-linguistic generalization [33]. Pustina *et al*. [34] and Kristinsson *et al*. [35] investigated the predictive power of neuroimaging data for aphasia symptoms and severity, finding that different neural networks in AI yielded successful results. Both studies showed that multimodal imaging data outperformed predictions based on a single modality and that different language abilities were best predicted with different neural predictors. These studies reveal that ML approaches can contribute to clinical decision-making processes in aphasia assessment; however, clinical validation of the models has not yet been completed. Themistocleous *et al*. [36] investigated the power of acoustic and linguistic features of connected speech to predict patients with primary progressive aphasia (PPA) and compared different AI networks. The models were generally highly accurate, especially for PPA patients with reduced verbal fluency and fluency difficulties, but less accurate in discriminating PPA patients with comprehension-related disorders. Additionally, the study compared the accuracy of the AI models with the classification performance of three less-experienced speech-language pathologists. The results showed that the AI models achieved a higher overall accuracy. These findings highlight the potential of AI to assist in clinical decision-making, indicating that it may surpass pathologists in specific situations [36]. A systematic review on this topic found that AI is primarily used for diagnosis and classification in aphasia rehabilitation. However, there is no evidence suggesting its integration into augmentative and alternative communication (AAC) devices or in direct therapy applications. Some studies have utilized AI to support aphasia therapy, while others have evaluated its effectiveness in modeling word production processes or classifying paraphasic errors [37].

AI in Motor Speech and Speech Sound Disorders

In the examination of AI applications for assessing and treating speech disorders, significant scholarly investigations have focused on both motor speech and speech sound disorders. Researchers have proposed a new language-based human-computer interaction tool, as well as a gamified AI-driven tool, specifically designed for individuals with motor speech disorders [38]. In addition, Frieg *et al*. [39] developed a digital training system for individuals with dysarthria. Another similar study evaluated the accuracy and therapeutic effect of an iPad-based speech therapy application with automatic speech recognition (ASR) software for individuals with apraxia of speech and aphasia after stroke and found that ASR agreed with expert assessment 80% of the time. Participants showed lasting improvements in word production accuracy with ASR-based feedback [40]. Ballard *et al*. [41] conducted a feasibility study of a tablet-based automated feedback tool for individuals with apraxia. In a review study, it was found that AI-based automated speech therapy tools developed for individuals with speech sound disorders offer potential benefits, but there is limited evidence of their effectiveness, and they cannot fully replace speech-language pathologists; nevertheless, it was emphasized that some time-consuming tasks of speech pathologists can be supported by artificial intelligence [42]. Another study examines how AI systems, specifically Voiceitt, can enhance AAC technologies for individuals with severe speech disorders. Voiceitt is an intelligent software application that transforms unintelligible speech into clear, understandable communication in real-time. This technology enables individuals with motor or cognitive disabilities to communicate effectively with caregivers, family members, healthcare professionals, and the broader society. Facilitating communication promotes greater participation and allows people with disabilities to live more independently [43].

AI in Cleft Palate Speech

Another area of focus concerning speech sound disorders is cleft palate speech. A systematic review by Zhang *et al*. [44] evaluated the effectiveness of AI algorithms in detecting persistent hypernasal speech that requires revision following primary repair surgery in individuals with cleft palate. The study found that these algorithms could detect hypernasality quickly and independently, achieving a high level of agreement with speech-language pathologists. These findings suggest that AI can enhance clinicians' capabilities and serve as a valuable complement to existing gold-standard practices.

AI in Developmental Language Disorder

Language disorders, particularly Developmental Language Disorder (DLD), represent significant domains in which AI is increasingly employed within speech-language pathology. A recent study focused on the early diagnosis of DLD, addressing the limitations of traditional risk factors in younger age groups and bilingual children. Researchers have developed a web-based tool called MARS, which has shown potential in distinguishing between children with and without DLD by analyzing rhythmic vocal production [45]. Studies have shown that AI-based communication devices enhance individuals' expressive speech by providing real-time feedback through natural language processing algorithms. Additionally, AI-supported interactive games and screening tools significantly contribute to language development, encourage social interaction, and facilitate the early detection of developmental disabilities. In ElHennawy's study [46], the roles of AI and ML in assessing and treating communication disorders were discussed, highlighting the importance of early and accurate diagnosis. The study demonstrated that AI can rapidly and accurately analyze large datasets to create tailored treatment plans for clients. Furthermore, it provides SLPs with various tools to enhance therapy processes [46].

Disadvantages, Limitations, and Barriers of AI in Speech and Language Pathology

Besides the advantages mentioned above, AI also presents some challenges and disadvantages in the field of speech and language pathology. Although several studies have reported positive outcomes for AI, significant limitations exist. For instance, researchers have highlighted that AI models used in swallowing disorders often operate under specific assumptions and lack transparency and interpretability. Concurrently, the absence of comprehensive and publicly available datasets has been identified as a substantial shortcoming [13]. Moreover, a study on voice disorders revealed that the diagnostic accuracy between AI and voice experts did not exceed 60.1%. This accuracy level is relatively low for critically essential conditions such as voice disorders [27]. Additionally, the researchers pointed out the existence of limited and unstable datasets in AI [29]. Another study on aphasia found that the AI program that converted text to images successfully transformed 189 out of 200 target words into images. Yet, the researchers noted that these images had aesthetic flaws that could impact their effectiveness. At the same time, while nouns and verbs can be illustrated with higher proficiency through the AI program, visualizing complete sentences remains quite challenging [30]. Although studies on the assessment of aphasia highlight the importance of AI support in clinical decision-making, research into the clinical validity of these models remains incomplete [34, 35]. Researchers in the field of motor speech disorders also emphasize the lack of sufficient evidence [42].

In a study by Suh *et al*. [7], it was indicated that SLPs face several key challenges, including keeping up with technological advancements, time constraints, resistance to technology, privacy and ethical concerns, job security concerns, and broader ethical issues such as artistic expression and intellectual property rights. In addition, Koenecke *et al*. [47] emphasized that ethical issues need to be systematically addressed in the clinical use of AI. These issues are particularly related to biases in datasets and algorithms, such as unfair representation. For instance, automatic speech recognition systems demonstrate lower accuracy for African American speakers of English compared to speakers of General American English, and this difference is particularly pronounced in children [47]. Since AI systems are often trained with real-world data, they can learn and reinforce data imbalances related to underrepresented groups; this is referred to as "algorithmic bias" and can lead to the exclusion or misclassification of individuals with age, gender, language differences, voice characteristics, or neurological disorders in SLP interventions [48]. Moreover, Kanwah *et al*. [49] reported that the operation of fully automated AI-based tools without therapist, caregiver, or parental involvement raises ethical concerns such as data bias, privacy violations, and the potential to replace speech-language pathologists, especially in speech sound disorders. In addition, concerns regarding the data privacy and security of patient information complicate research processes and hinder the integration of AI into clinical practice [50]. Furthermore, a recent study by Birol *et al*. [51], which aimed to explore the potential of ChatGPT, found that ChatGPT struggled to generate materials specific to Turkish; its responses and generation lacked sufficient depth, particularly regarding specialized terminology and culturally relevant

stimuli. Additionally, another limitation of ChatGPT is that it often requires the assistance of an experienced SLP to address its shortcomings. This study also suggests that tackling the ethical concerns, risks, and practical challenges associated with integrating AI into clinical practice is crucial. AI systems require large amounts of patient data, which raises issues related to data storage, sharing, protection, and patient privacy. Moreover, these systems can produce misdiagnoses, particularly in cases that fall outside their training data. Overreliance on AI tools may also diminish clinicians' critical thinking skills. The high costs of developing and maintaining AI could limit accessibility for some healthcare providers. To address these concerns, proper training for clinicians is essential, along with establishing clear ethical guidelines and standards to govern the use of AI in healthcare [51].

CONCLUSION

The increasing popularity of AI and AI-based algorithms in speech and language pathology has been acknowledged and backed by existing literature. AI is now being used not only for assessments but also in therapies, monitoring, and reporting. Although the current evidence remains insufficient and there are ongoing concerns about ethics and implementation, the advancement of technology offers promise for the application of AI in speech and language pathology. Future studies will enhance speech-language pathologists' understanding and awareness of AI, contribute to a more substantial evidence base, and lead to improvements in clinical practice.

Ethics Approval and Consent to Participate

Ethical approval is not required for this study. There are no human or animal elements in the study. This review was carried out by a brief literature screening.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: EB, BÖ; Study Design: BÖ, EB; Supervision: EB, BÖ, SSB; Funding: N/A; Materials: N/A; Data Collection and/or Processing: BÖ; Statistical Analysis and/or Data Interpretation: EB, BÖ, SSB; Literature Review: EB, BÖ; Manuscript Preparation: EB, BÖ, SSB; and Critical Review: EB, BÖ, SSB.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

Generative Artificial Intelligence Statement
The author(s) declare that artificial intelligence tools were used in accordance with academic ethical standards during the preparation of this manuscript. They received support from Grammarly AI for some language edits. The all content of the study was produced by the author(s) in accor dance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the author(s) and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Morandín-Ahuerma F. What is Artificial Intelligence? International Journal of Research Publication and Reviews. 2022;3(12):1947-1951. doi: 10.55248/gengpi.2022.31261.
2. Villamil V, Deloria R, Wolbring G. Artificial intelligence and machine learning: What is the role of social workers, occupational therapists, audiologists, nurses and speech language pathologists according to academic literature and Canadian newspaper coverage? In: ACM International Conference Proceeding Series.

Association for Computing Machinery; 2019:83-86. doi: 10.1145/3364138.3364158.
3. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88. doi: 10.1016/j.media.2017.07.005.
4. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: Past, present and future. Stroke Vasc Neurol. 2017;2(4):230-243. doi: 10.1136/svn-2017-000101.
5. Duffy JR. Motor Speech Disorders: Where Will We Be in 10 Years? Semin Speech Lang. 2016;37(3):219-224. doi: 10.1055/s-0036-1584154.
6. ASHA. How Will Artificial Intelligence Reshape Speech-Language Pathology Services and Practice in the Future? 2020. (access date: 13.04.2024) https://academy.pubs.asha.org/2020/08/how-will-artificial-intelligence-reshape-speech-language-pathology-services-and-practice-in-the-future.
7. Suh H, Dangol A, Meadan H, Miller CA, Kientz JA. Opportunities and Challenges for AI-Based Support for Speech-Language Pathologists. In: ACM International Conference Proceeding Series. Association for Computing Machinery. 2024. doi: 10.1145/3663384.3663387.
8. Aggarwal K, Ravi R, Yerraguntla K. Artificial Intelligence and Patient Care: Perspectives of Audiologists and Speech-Language Pathologists. Intell Based Med. 2025;11:100214. doi: 10.1016/j.ibmed.2025.100214.
9. Zhang B, Wong KP, Liu M, et al. Effect of artificial intelligence-based video-game system on dysphagia in patients with stroke: A randomized controlled trial. Clin Nutr. 2025;45:81-90. doi: 10.1016/j.clnu.2024.12.022.
10. Hou M, Zhao Y, Zhao L, Yuan X, Liu Z, Li H. Efficacy of game training combined with surface electromyography biofeedback on post-stroke dysphagia. Geriatr Nurs (Minneap). 2024;55:255-262. doi: 10.1016/j.gerinurse.2023.11.019.
11. Jeong CW, Lee CS, Lim DW, et al. The Development of an Artificial Intelligence Video Analysis-Based Web Application to Diagnose Oropharyngeal Dysphagia: A Pilot Study. Brain Sci. 2024;14(6):546. doi: 10.3390/brainsci14060546.
12. Barron K, Blaivas M, Blaivas L, Sadler J, Deal I. Bedside Ultrasound to Identify and Predict Severity of Dysphagia Following Ischemic Stroke: Human Versus Artificial Intelligence. Ultrasound Med Biol. 2024;50(1):99-104. doi: 10.1016/j.ultrasmedbio.2023.09.008.
13. Sanjeevi G, Gopalakrishnan U, Pathinarupothi RK, Iyer KS. Artificial Intelligence in Videofluoroscopy Swallow Study Analysis: A Comprehensive Review. Dysphagia. 2025 Feb 17. doi: 10.1007/s00455-025-10812-8.
14. Kim JK, Choo YJ, Choi GS, Shin H, Chang MC, Park D. Deep Learning Analysis to Automatically Detect the Presence of Penetration or Aspiration in Videofluoroscopic Swallowing Study. J Korean Med Sci. 2022;37(6):e42. doi: 10.3346/jkms.2022.37.e42.
15. Iida Y, Näppi J, Kitano T, Hironaka T, Katsumata A, Yoshida H. Detection of aspiration from images of a videofluoroscopic swallowing study adopting deep learning. Oral Radiol. 2023;39(3):553-562. doi: 10.1007/s11282-023-00669-8.
16. Bandini A, Steele CM. The effect of time on the automated detection of the pharyngeal phase in videofluoroscopic swallowing studies. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. Institute of Electrical and Electronics Engineers Inc.; 2021:3435-3438. doi: 10.1109/EMBC46164.2021.9629562.
17. Hsiao MY, Weng CH, Wang YC, et al. Deep Learning for Automatic Hyoid Tracking in Videofluoroscopic Swallow Studies. Dysphagia. 2023;38(1):171-180. doi: 10.1007/s00455-022-10438-0.
18. Nakamori M, Ishikawa R, Watanabe T, et al. Swallowing sound evaluation using an electronic stethoscope and artificial intelligence analysis for patients with amyotrophic lateral sclerosis. Front Neurol. 2023;14:1212024. doi: 10.3389/fneur.2023.1212024.
19. Girardi AM, Cardell EA, Bird SP. Artificial Intelligence in the Interpretation of Videofluoroscopic Swallow Studies: Implications and Advances for Speech-Language Pathologists. Big Data Cog Comput. 2023;7(4):178. doi: 10.3390/bdcc7040178.
20. Liu GS, Jovanovic N, Sung CK, Doyle PC. A Scoping Review of Artificial Intelligence Detection of Voice Pathology: Challenges and Opportunities. Otolaryngol Head Neck Surg. 2024;171(3):658-666. doi: 10.1002/ohn.809.
21. Fujimura S, Kojima T, Okanoue Y, et al. Classification of Voice Disorders Using a One-Dimensional Convolutional Neural Network. J Voice. 2022;36(1):15-20. doi: 10.1016/j.jvoice.2020.02.009.
22. Verde L, De Pietro G, Alrashoud M, Ghoneim A, Al-Mutib KN, Sannino G. Leveraging Artificial Intelligence to Improve Voice Disorder Identification Through the Use of a Reliable Mobile App. IEEE Access. 2019;7:124048-124054. doi: 10.1109/ACCESS.2019.2938265.
23. Ur Rehman M, Shafique A, Azhar QUA, Jamal SS, Gheraibia Y, Usman AB. Voice disorder detection using machine learning algorithms: An application in speech and language pathology. Eng Appl Artif Intell. 2024;133:108047. doi: 10.1016/j.engappai.2024.108047.
24. Costantini G, Cesarini V, Di Leo P, et al. Artificial Intelligence-Based Voice Assessment of Patients with Parkinson's Disease Off and On Treatment: Machine vs. Deep-Learning Comparison. Sensors (Basel). 2023;23(4):2293. doi: 10.3390/s23042293.
25. Hegde S, Shetty S, Rai S, Dodderi T. A Survey on Machine Learning Approaches for Automatic Detection of Voice Disorders. J Voice. 2019;33(6):947.e11-33. doi: 10.1016/j.jvoice.2018.07.014.
26. Al-Hussain G, Shuweihdi F, Alali H, Househ M, Abd-Alrazaq A. The Effectiveness of Supervised Machine Learning in Screening and Diagnosing Voice Disorders: Systematic Review and Meta-analysis. J Med Internet Res. 2022;24(10):e38472. doi: 10.2196/38472.
27. Hu HC, Chang SY, Wang CH, et al. Deep learning application for vocal fold disease prediction through voice recognition: Preliminary development study. J Med Internet Res. 2021;23(6):e25247. doi: 10.2196/25247.
28. Kim HB, Song J, Park S, Lee YO. Classification of laryngeal diseases including laryngeal cancer, benign mucosal disease, and vocal cord paralysis by artificial intelligence using voice analysis. Sci Rep. 2024;14(1):9297. doi: 10.1038/s41598-024-58817-x.
29. Idrisoglu A, Dallora AL, Anderberg P, Berglund JS. Applied Machine Learning Techniques to Diagnose Voice-Affecting Conditions and Disorders: Systematic Literature Review. J Med In-

ternet Res. 2023;25:46105. doi: 10.2196/46105.
30. Cordella C, Marte MJ, Liu H, Kiran S. An Introduction to Machine Learning for Speech-Language Pathologists: Concepts, Terminology, and Emerging Applications. Perspect ASHA Spec Interest Groups. 2024;10(2):432-450. doi: 10.1044/2024_PERSP-24-00037.
31. Jothi KR, Sivaraju SS, Yawalkar PJ. AI based Speech Language Therapy using Speech Quality Parameters for Aphasia Person: A Comprehensive Review. In: Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2020. Institute of Electrical and Electronics Engineers Inc.; 2020:1263-1271. doi: 10.1109/ICECA49313.2020.9297591.
32. de la Fuente Garcia S, Ritchie CW, Luz S. Artificial Intelligence, Speech, and Language Processing Approaches to Monitoring Alzheimer's Disease: A Systematic Review. J Alzheimers Dis. 2020;78(4):1547-1574. doi: 10.3233/JAD-200888.
33. Grasemann U, Peñaloza C, Dekhtyar M, Miikkulainen R, Kiran S. Predicting language treatment response in bilingual aphasia using neural network-based patient models. Sci Rep. 2021;11(1):10497. doi: 10.1038/s41598-021-89443-6.
34. Pustina D, Coslett HB, Ungar L, et al. Enhanced estimations of post-stroke aphasia severity using stacked multimodal predictions. Hum Brain Mapp. 2017;38(11):5603-5615. doi: 10.1002/hbm.23752.
35. Kristinsson S, Zhang W, Rorden C, et al. Machine learning-based multimodal prediction of language outcomes in chronic aphasia. Hum Brain Mapp. 2021;42(6):1682-1698. doi: 10.1002/hbm.25321.
36. Themistocleous C, Ficek B, Webster K, den Ouden DB, Hillis AE, Tsapkini K. Automatic Subtyping of Individuals with Primary Progressive Aphasia. J Alzheimers Dis. 2021;79(3):1185-1194. doi: 10.3233/JAD-201101.
37. Azevedo N, Kehayia E, Jarema G, Le Dorze G, Beaujard C, Yvon M. How artificial intelligence (AI) is used in aphasia rehabilitation: A scoping review. Aphasiology. 2024;38(2):305-336. doi: 10.1080/02687038.2023.2189513.
38. Bilkova Z, Novozamsky A, Bartos M, et al. Human computer interface based on tongue and lips movements and its application for speech therapy system. In: Proceeding of IS and T International Symposium on Electronic Imaging: Visualization and Data Analysis. 2020: pp. 389.1-5. doi: 10.2352/ISSN.2470-1173.2020.1.VDA-389.
39. Frieg H, Muehlhaus J, Ritterfeld U, Bilda K. ISi-Speech: A Digital Training System for Acquired Dysarthria. Stud Health Technol Inform. 2017;242:330-334. doi: 10.3233/978-1-61499-798-6-330.
40. Saz O, Yin SC, Lleida E, Rose R, Vaquero C, Rodríguez WR. Tools and Technologies for Computer-Aided Speech and Language Therapy. Speech Commun. 2009;51(10):948-967. doi: 10.1016/j.specom.2009.04.006.
41. Ballard KJ, Etter NM, Shen S, Monroe P, Tien Tan C. Feasibility of Automatic Speech Recognition for Providing Feedback During Tablet-Based Treatment for Apraxia of Speech Plus Aphasia. Am J Speech Lang Pathol. 2019;28(2S):818-834. doi: 10.1044/2018_AJSLP-MSC18-18-0109.
42. Deka C, Shrivastava A, Abraham AK, Nautiyal S, Chauhan P. AI-based automated speech therapy tools for persons with speech sound disorder: a systematic literature review. Speech Lang Hear. 2025;28(1):2359274. doi: 10.1080/2050571X.2024.2359274.
43. Murero M, Vita S, Mennitto A, D'Ancona G. Artificial Intelligence for Severe Speech Impairment: Innovative approaches to AAC and Communication. PSYCHOBIT 2020. Proceedings of the Second Symposium of Psychology-Based Technologies. Naples, Italy, September 28-29, 2020. CEUR Workshop Proceedings. 2020;2730:31. https://ceur-ws.org/Vol-2730/paper31.pdf. (access date: 13.04.2024)
44. Zhang A, Pyon RE, Chen K, Lin AY. Speech Analysis of Patients with Cleft Palate Using Artificial Intelligence Techniques: A Systematic Review. Face. 2023;4(3):327-337. doi: 10.1177/27325016231187985.
45. Beccaluva EA, Catania F, Arosio F, Garzotto F. Predicting developmental language disorders using artificial intelligence and a speech data analysis tool. Hum Comput Interact. 2024;39(1-2):8-42. doi: 10.1080/07370024.2023.2242837.
46. El-Hennawy M. The Impact of Artificial Intelligence (AI) in the Assessment and Treatment of Communication Disorders (A Review of Literature). Egypt J Lang Eng. 2024;11(2):36-45. doi: 10.21608/EJLE.2024.303151.1069.
47. Koenecke A, Nam A, Lake E, Nudell J, Quartey M, Mengesha Z, Toups C, Rickford JR, Jurafsky D, Goel S. Racial disparities in automated speech recognition. Proc Natl Acad Sci U S A. 2020;117(14):7684-7689. doi: 10.1073/pnas.1915768117.
48. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. J Glob Health. 2019;9(2):010318. doi: 10.7189/jogh.09.020318.
49. Kanwal K, Ibrahim Z, Siddiqui AA. The Convergence of Artificial Intelligence, Digital Therapeutics, and Biomedical Electronics in the field of Speech-Language Therapy. Allied Med Res J. 2024;2(2):277-292. doi: 10.59564/amrj/02.02/030.
50. Zhang J, Zhang ZM. Ethics and governance of trustworthy medical artificial intelligence. BMC Med Inform Decis Mak. 2023;23(1):7. doi: 10.1186/s12911-023-02103-9.
51. Birol NY, Çiftci HB, Yılmaz A, Çağlayan A, Alkan F. Is there any room for ChatGPT AI bot in speech-language pathology? Eur Arch Otorhinolaryngol. 2025;282(6):3267-3280. doi: 10.1007/s00405-025-09295-y.

The European Research Journal 2025;11(6):1244-1248 Case Report

DOI: https://doi.org/10.18621/eurj.1663642 *Cardiology*

A case of acute pancreatitis complicated by acute coronary syndrome

Alkame Akgümüş, Ali Duygu

Department of Cardiology, Bandırma Onyedi Eylul University, Faculty of Medicine, Balıkesir, Türkiye

ABSTRACT

An inflammatory condition affecting the pancreatic parenchyma is called acute pancreatitis. Alcohol consumption and gallstones are the most frequent etiological factors. Patients with acute pancreatitis can appear clinically in a variety of ways. While individuals with acute pancreatitis may experience ECG abnormalities that resemble acute myocardial infarction, it is uncommon for acute pancreatitis to coexist with actual acute myocardial infarction. Our goal was to present a case of acute pancreatitis exacerbated by acute coronary syndrome. The patient arrived at the emergency room with 60 minutes of epigastric abdominal pain that was prominent there and radiating to the back.

Keywords: Acute pancreatitis, acute coronary syndrome, anjioplasty

Acute pancreatitis is an inflammatory disease of the pancreas clinically characterised by abdominal pain accompanied by elevated serum levels of pancreatic enzymes (amylase and lipase). Pain is a prominent and distinctive feature of acute pancreatitis. It is localised in the epigastric region in more than 60% of patients [1]. Multiple systems are linked to a number of acute pancreatitis problems. Acute pancreatitis may result in systemic problems such as pulmonary, cardiovascular, haematological, renal, metabolic, and central nervous system abnormalities, as well as local complications such as pancreatic necrosis, abscess, or pseudocyst. Cardiovascular complications of acute pancreatitis include especially myocardial infarction (MI), shock, hypovolemia and pericardial effusion [2]. Here, we aimed to emphasise the effect of early coronary stenting on possible mortality and complications in a case of acute pancreatitis complicated with acute coronary syndrome (ACS).

CASE PRESENTATION

A 57-year-old woman was admitted to the emergency department with complaints of abdominal pain in the epigastric region for approximately 60 minutes. During hospitalisation, the temperature was 36.7 °C, blood pressure 110/70 mmHg and pulse rate 67/min. Physical examination was unremarkable except for tenderness in the epigastric region. The patient had a history of hypertension. Additionally, she had a 20-year history of smoking. The medical history obtained revealed no history of malignancy, autoimmune disease, hepatobiliary or gastrointestinal disease. Body mass index (BMI) was 33.2 (height: 1.55 m, weight: 80 kg). A prior myocardial infarction history was discovered. In the inferior leads, the ECG had abnormal Q waves. (Fig. 1).

Laboratory examination: leukocyte: 8.29/mm3, haemoglobin: 12.6 g/dL, hematocrit: 40.1%, platelet:

Received: March 23, 2025 **Accepted:** April 28, 2025 **Available Online:** June 11, 2025 **Published:** November 4, 2025

How to cite this article: Akgümüş A, Duygu A. A case of acute pancreatitis complicated by acute coronary syndrome. Eur Res J. 2025;11(6):1244-1248. doi: 10.18621/eurj.1663642

Corresponding author: Ali Duygu, MD., Assist. Prof., Phone: +90 226 717 01 17, E-mail: ali.duygu1989@gmail.com, aduygu@bandirma.edu.tr

Available Online at https://dergipark.org.tr/en/pub/eurj

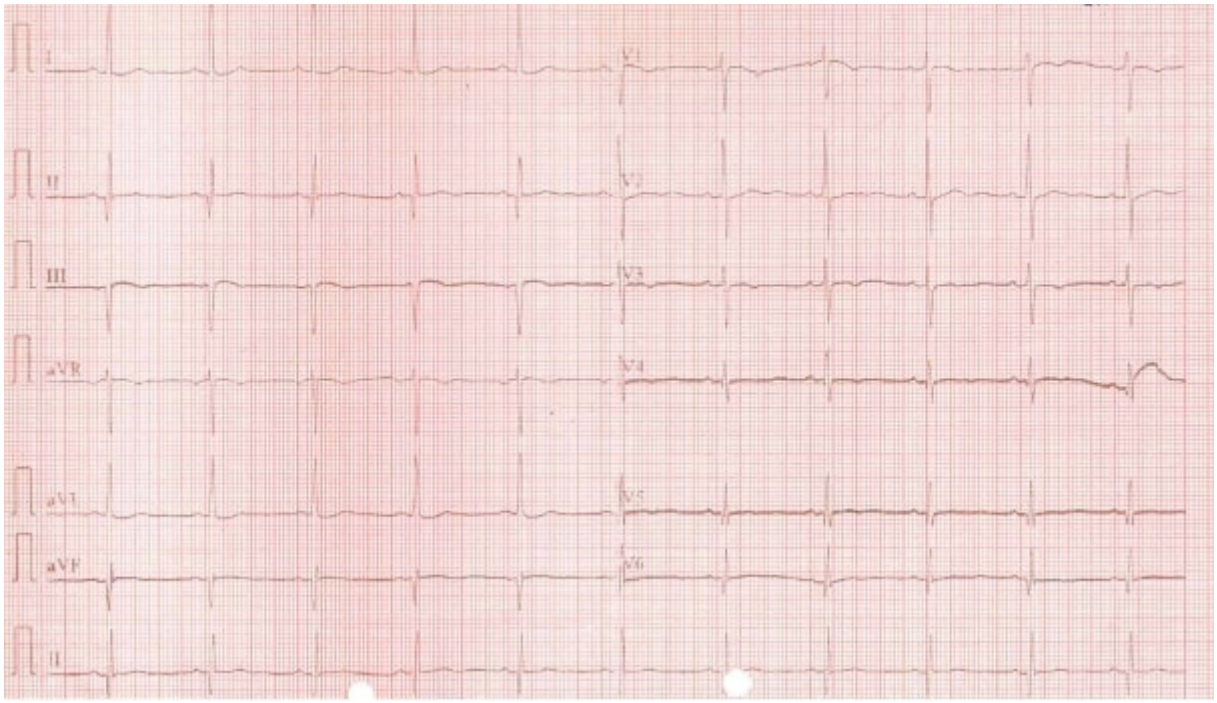

Fig. 1. ECG at the time of admission.

283.000/mm3, glucose: 144 mg/dl, urea: 30 mg/dl, creatinine: 0.98 mg/dl, alanine aminotransferase: 38 IU/L, aspartate aminotransferase: 142 IU/L, sodium: 141 mmol/L, potassium: 3.9 mmol/L, alkaline phosphatase 176 U/L and troponin 300 pg/mL (0.1-15.6). The patient was diagnosed as ACS and emergency coronary angiography (CAG) was performed. Plaque in the left anterior descending coronary artery (LAD) and circumflex coronary artery and thrombus image in the mid right coronary artery (RCA) were detected and stent placement and complete patency was achieved (Fig. 2). During coronary intensive care unit follow-up, the patient's complaints and troponin values increased and control CAG was performed and the stent in the RCA was found to be patent (Fig. 3).

Noncardiac pathologies were thought to be the cause of the patient's complaints. Abdominal examination revealed tenderness in the epigastric region.

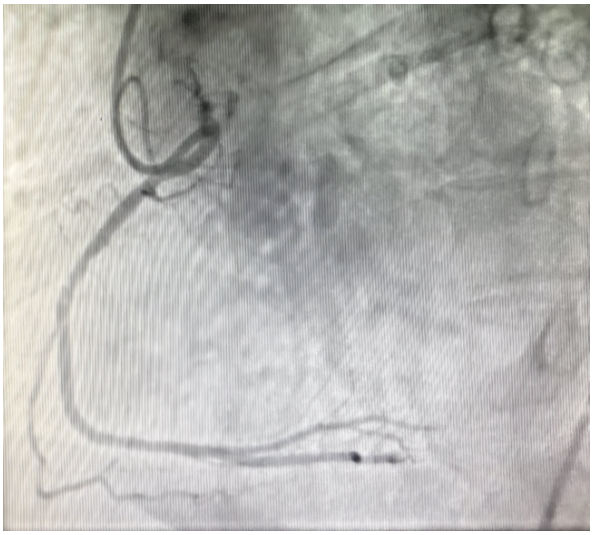

Fig. 2. RCA before stent implantation.

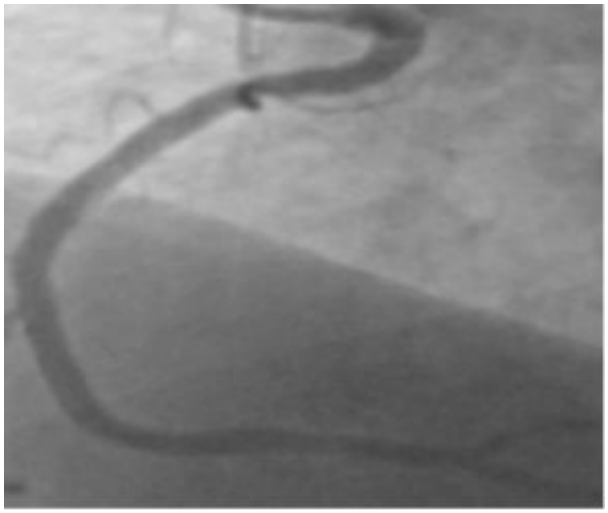

Fig. 3. Image of patent RCA after control CAG and implanted stent.

Laboratory tests were re-studied more comprehensively. Amylase value was found to be 1,218 U/L. Abdominal tomography showed heterogeneous and oedematous pancreatic head corpus and bile sludge in the gallbladder (Fig. 4). When assessed according to Ranson criteria, the diagnosis was made with a score of 3, indicating severe pancreatitis. After 48 hours of follow-up, the hematocrit was measured at 32.1 (decrease >%10), and due to fluid deficit, the score was 5, with a mortality risk of 40%.

In the coronary intensive care unit, acute pancreatitis treatment was organised simultaneously with ACS treatment. Oral intake was stopped and IV hydration treatment was started. In daily follow-up, amylase values normalised and complaints regressed. Vital values were stable, medical treatment was organised and the patient was discharged with recommendations.

DISCUSSION

Acute pancreatitis is a clinical picture with high mortality and morbidity that develops as a result of activation of inactive enzymes in the pancreas due to various reasons and autodigestion of the pancreas, manifested by severe abdominal pain and may lead to local and systemic complications. Although many factors play a role in the etiology of acute pancreatitis, biliary gallstones and excessive alcohol consumption are responsible for 70% of cases. Drugs, hyperlipidaemia, pregnancy, endoscopic retrograde cholangiopancreatography (ERCP) and trauma are among other causes. Mortality rate ranging from 1% to 9% is influenced by the severity of the disease and several prognostic factors [3, 4].

Many complications of acute pancreatitis are related to multiple systems. Local complications such as pancreatic necrosis, abscess or pseudocyst and systemic complications such as pulmonary, cardiovascular, haematological, renal, metabolic and central nervous system abnormalities may occur in the course of acute pancreatitis. While ST segment elevation is rare among cardiovascular events, other ECG findings including arrhythmia, conduction anomalies and du-

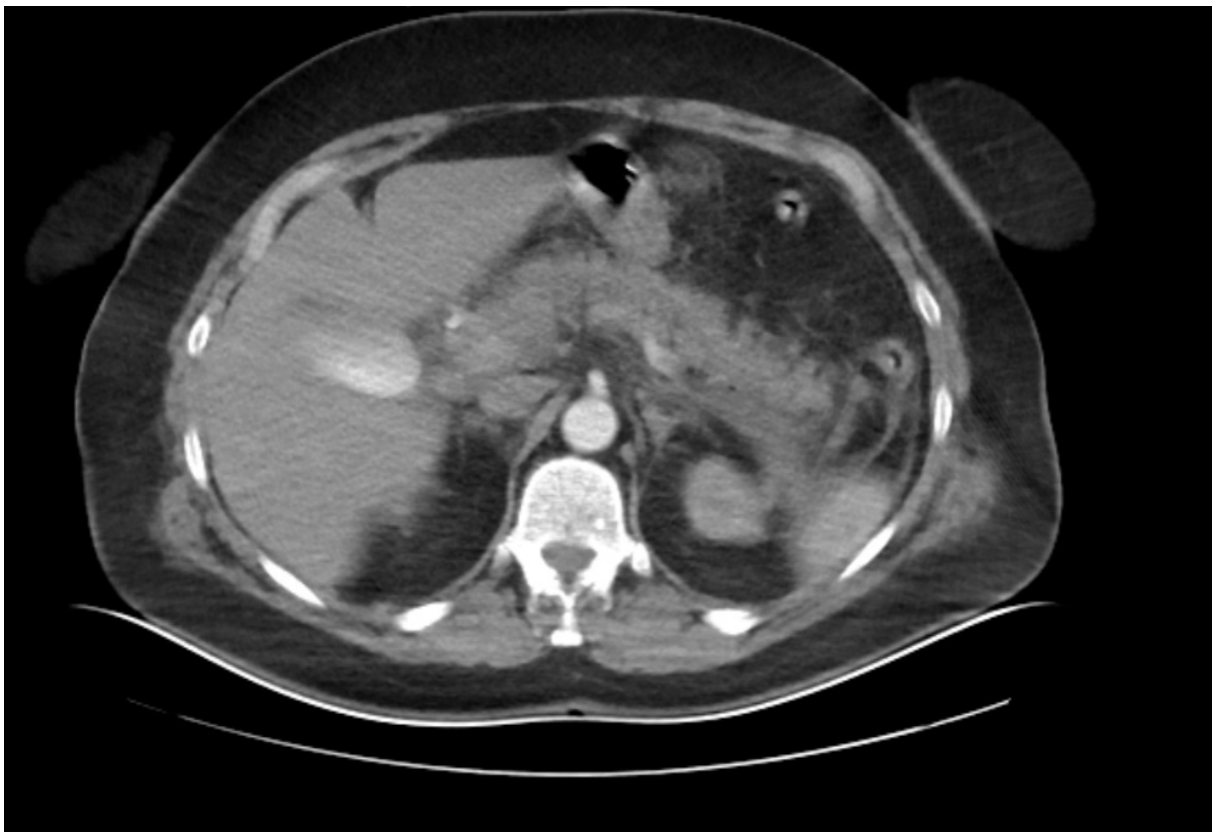

Fig. 4. Computed tomography image of acute pancreatitis.

ration changes in T wave or QT period are more common [2]. Although patients with acute pancreatitis may exhibit ECG abnormalities that resemble acute myocardial infarction, it is extremely uncommon for acute pancreatitis to coexist with acute myocardial infarction [5]. Cases of acute infarction with elevated troponin level detected with chest pain despite a normal ECG have also been described [6]. The mechanism of myocardial involvement observed in the course of acute pancreatitis is not very clear. According to clinical research, intravenous injections of pancreatic proteolytic enzymes into a rabbit model have been demonstrated to induce acute myocardial necrosis, which resolves after two weeks [7]. The existence of a cardiobiliary reflex is another mechanism that might harm the heart by changing the flow of blood to the coronary arteries. Patients with underlying coronary artery disease should pay particular attention to this mechanism [8]. Furthermore, a third fluid gap is known to be caused by acute pancreatitis, and this intravascular loss might have played a role in coronary hypoperfusion. More research is still needed in this area. In our study, we tried to contribute to this aim by presenting this rare case to the literature.

There are many hypotheses proposed between myocardial disease and ECG abnormalities seen in acute pancreatitis. These include vagally mediated reflexes (cardiobiliary reflex), metabolic and electrolyte abnormalities, toxic effects of pancreatic enzymes on the myocardium, coronary artery spasm, haemodynamic dysregulation and/or systemic inflammatory response and prothrombotic conditions. It has been reported that trypsin, which has an important role in the pathophysiology of pancreatitis, may alter platelet adhesion, affect the coagulation system and lead to coronary thrombosis [9]. Under appropriate clinical conditions, ST-segment elevation on ECG is a prominent feature of acute myocardial infarction and urgent reperfusion therapy is recommended. However, other conditions in which ST-segment elevation on ECG may occur for reasons other than acute myocardial ischaemia should be well known [9]. Acute pancreatitis is both an emergency and a clinical entity that must be taken into consideration.

Acute pancreatitis is a challenging disease that can lead to many complications. Its association with ACS is rare and our case is an extreme example. Previously published studies have reported cases of ACS as a complication of acute pancreatitis. However, cases of acute pancreatitis secondary to ACS are extremely rare and there are few examples in the literature. Despite all these studies, the association of acute pancreatitis and ACS is still awaiting clarification. In our case, we think that multidisciplinary evaluation of the patients can broaden the physician's perspective and it is extremely important in terms of patient survival by making the diagnosis of diseases with similar symptoms quickly.

CONCLUSION

In conclusion, it should be kept in mind that patients presenting with ACS may have concurrent cardiac and/or noncardiac pathologies. It should be kept in mind that various ECG changes may also occur in noncardiac patients such as acute pancreatitis. The diagnosis of true myocardial infarction and the application of appropriate treatment approaches are extremely important in terms of morbidity and mortality of patients.

Ethics Approval and Consent to Participate

The authors confirm that written consent for the submission and publication of this case, including images, was obtained from the patient in line with COPE guidance. Patient was informed about the purpose of the case report, and informed consent was obtained from the patient for this publication.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: AD; Study Design: AA; Supervision: AA; Funding: AA; Materials: AD; Data Collection and/or Processing: AA; Statistical Analysis and/or Data Interpretation: AD; Literature Review: AA; Manuscript Preparation: AD and Critical Review: AA.

Conflict of interest

The authors disclosed no conflict of interest during

the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used during the preparation process of this manuscript. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the author(s) and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Murphy JO, Mehigan BJ, Keane FB. Acute pancreatitis. Hosp Med. 2002;63(8):487-492. doi: 12968/hosp.2002.63.8.1970.
2. Yegneswaran B, Kostis JB, Pitchumoni CS. Cardiovascular manifestations of acute pancreatitis. J Crit Care. 2011;26(2):225.e11-8. doi: 10.1016/j.jcrc.2010.10.013.
3. Grendell J H. Acute pancreatitis. In Current Diagnosis and Treatment in gastroenterology. 2nd ed., New York: Lange Medical Boks/Mc Graw-Hill; 2003, pp. 489-495.
4. Avunduk C. Manual of Gastroenterology. Philedelphia: Lippincott Williams and Wilkins.; 2002.
5. Korantzopoulos P, Pappa E, Dimitroula V, et al. ST-segment elevation pattern and myocardial injury induced by acute pancreatitis. Cardiology 2005;103(3):128-130. doi: 10.1159/000083438.
6. Aundhakar S, Mahajan S, Agarwal A, Mhaskar D. Acute pancreatitis associated with elevated troponin levels: whether to thrombolyse or not? Ann Med Health Sci Res. 2013;3(Suppl 1):S50-2. doi: 10.4103/2141-9248.121225.
7. Kellner A, Robertson T. Selective necrosis of cardiac and skeletal muscle induced experimentally by means of proteolytic enzyme solutions given intravenously. J Exp Med. 1954;99(4):387-404. doi: 10.1084/jem.99.4.387.
8. Lieberman Js, Taylor A, Wright Is. The effect of intravenous trypsin administration on the electrocardiogram of the rabbit. Circulation. 1954;10(3):338-42. doi: 10.1161/01.cir.10.3.338.
9. Wang K, Asinger RW, Marriott HJ. ST-segment elevation in conditions other than acute myocardial infarction. N Engl J Med. 2003;349(22):2128-2135. doi: 10.1056/NEJMra022580.

The European Research Journal 2025;11(6):1249-1253 Case Report

DOI: https://doi.org/10.18621/eurj.1670211 *Medical Pathology*

Cytological misdiagnosis of high-grade medullary thyroid carcinoma with papillary-like nuclear features: A case report

Ali Mızrak

Department of Medical Pathology, Sultangazi Haseki Training and Research Hospital, İstanbul Türkiye

ABSTRACT

Medullary thyroid carcinoma is an uncommon form of thyroid cancer that arises from the parafollicular C cells. In this case report, a high-grade medullary thyroid carcinoma is discussed, which was initially diagnosed as papillary carcinoma due to the presence of papillary-like nuclear features in the cytological evaluation. A 65-year-old female presented to the ear, nose, and throat department with a palpable neck mass. Based on the cytological features of the fine needle aspiration (FNA), a diagnosis of 'malignant, papillary carcinoma', (Bethesda category 6) was made. After surgery, histological and immunohistochemical results led to a diagnosis of medullary thyroid carcinoma. Papillary carcinoma-like nuclear features may also be observed in medullary thyroid carcinoma and some cases could be mistakenly interpreted as papillary carcinoma in cytological examination.

Keywords: Cytology, intranuclear pseudo inclusions, medullary thyroid carcinoma, rearrangement during transfection (RET) gene mutation

Medullary thyroid carcinoma is an uncommon form of thyroid cancer that arises from the parafollicular C cells. Clinical findings, ultrasonographic features of the lesion, aspiration needle washout fluid, and serum biochemical markers can be helpful in establishing a diagnosis [1-3]. Immunohistochemical studies, combined with characteristic cytomorphological features observed in FNA samples, can be used to reach a definitive diagnosis [4-7]. However, medullary thyroid carcinoma may show various cytological and histological features that can be confused with many thyroid neoplasms [8, 9]. This case report discusses a high-grade medullary thyroid carcinoma with a rearrangement during transfection (RET) proto-oncogene mutation, initially misdiagnosed as papillary carcinoma in FNA cytology due to papillary-like nuclear features.

CASE PRESENTATION

A 65-year-old female presented to the ear, nose, and throat department with a palpable neck mass. Ultrasonographic examination revealed a solid mass lesion measuring 4·3 cm with irregular borders in the right lobe of the thyroid. A FNA was performed on the identified lesion. The samples were subsequently stained using Papanicolaou (PAP) and Giemsa stains. Although the cell block contained insufficient cells, the smears were cellular. Cytological examination re-

Received: April 4, 2025 **Accepted:** July 21, 2025 **Available Online:** July 22, 2025 **Published:** November 4, 2025

How to cite this article: Mızrak A. Cytological misdiagnosis of high-grade medullary thyroid carcinoma with papillary-like nuclear features: A case report. Eur Res J. 2025;11(6):1249-1253. doi: 10.18621/eurj.1670211

Corresponding author: Ali Mızrak, MD., Phone: +90 212 453 20 00, E-mail: dralimizrak@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

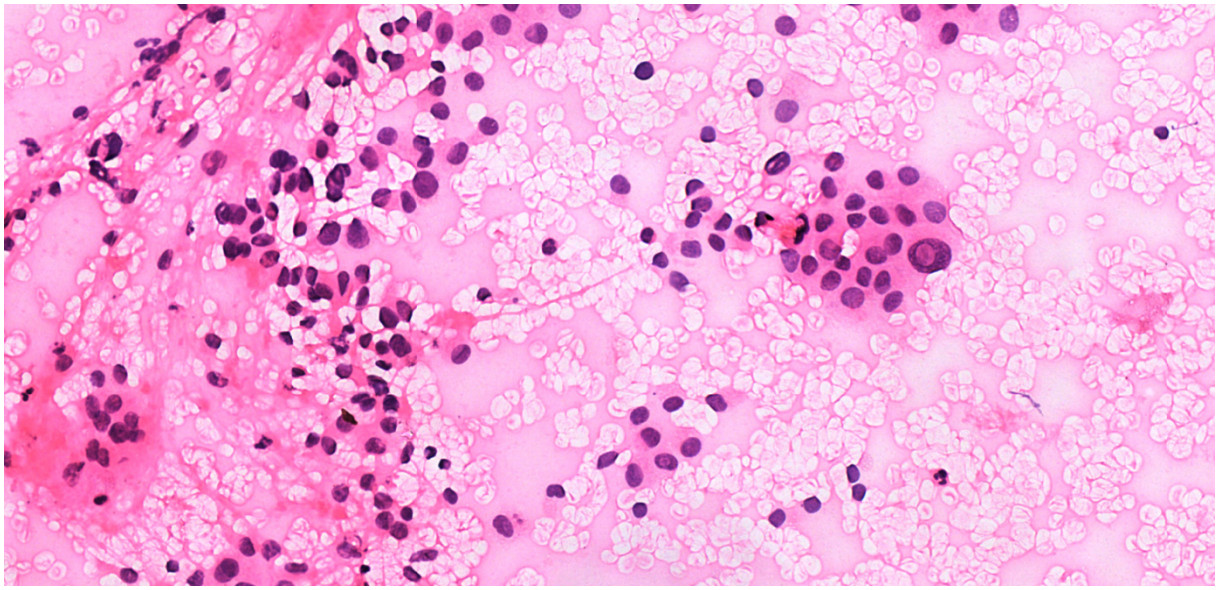

Fig. 1. **Cytological appearance of neoplastic cells (PAP, ×400).**

vealed nuclear enlargement, irregular borders, nuclear clearing, grooves, and intranuclear pseudo inclusions. The cytomorphological features resembled those of papillary thyroid carcinoma, leading to an initial misdiagnosis. The cytological diagnosis was reported as 'malignant, papillary thyroid carcinoma' based on the Bethesda classification (Fig. 1). The surgeon reviewed the radiological and cytological findings and decided to proceed with a total thyroidectomy. During surgery, it was noted that the lesion was adherent to the trachea, preventing the performance of a total thyroidectomy. The surgery was completed with a right thyroidectomy. Histological examination revealed a neoplastic lesion invading the surrounding muscle tissue. Acantholytic pseudopapillary structures were observed in the majority of the lesion. Other areas showed focal solid, follicular, and trabecular arrangements (Figs. 2 and 3). The cells exhibited nuclear clearing, grooves, and intranuclear pseudoinclusions. Numerous mitoses (15 in 2 mm^2) were observed under high magnification. Im-

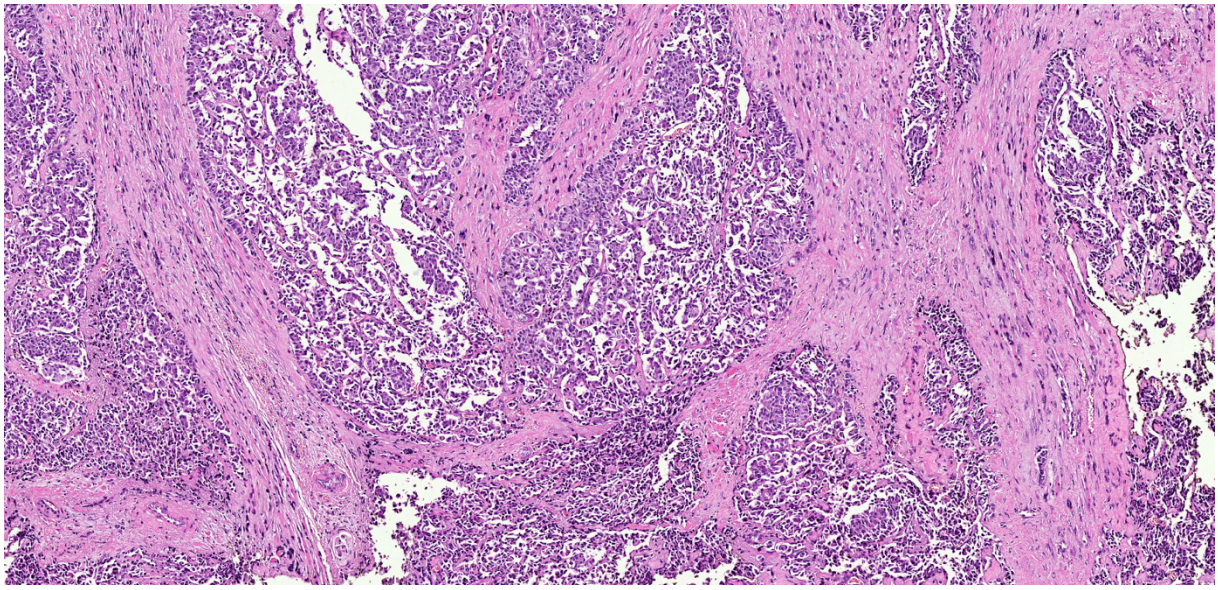

Fig. 2. **Thyroid medullary carcinoma (H&E, ×40).**

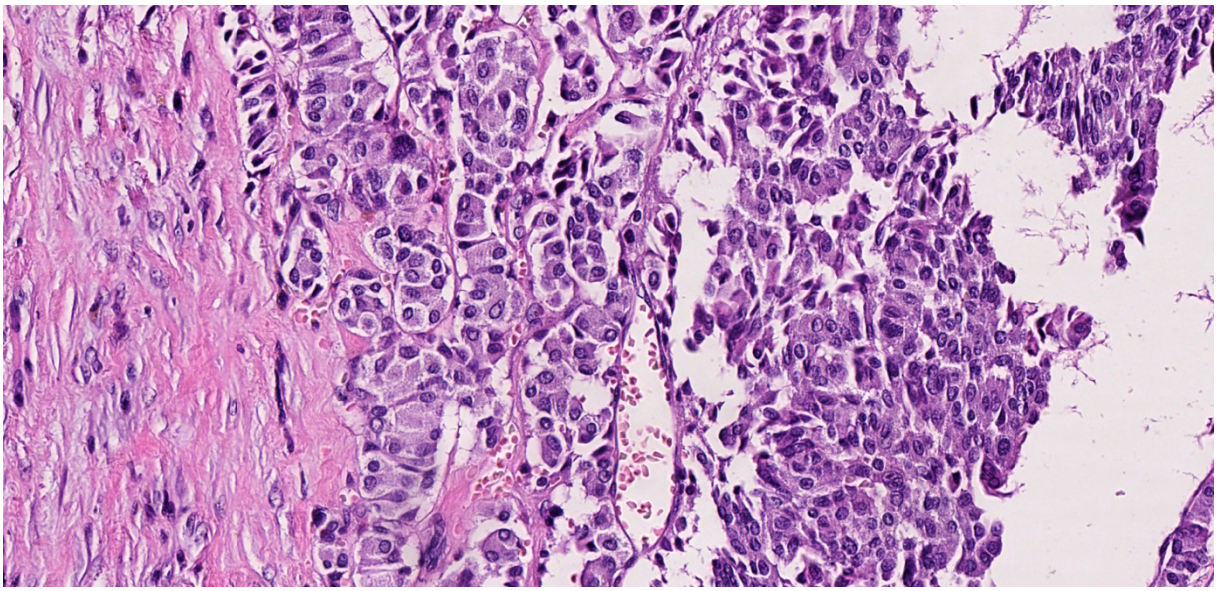

Fig. 3. Thyroid medullary carcinoma (H&E, ×400).

munohistochemical staining showed positivity for synaptophysin, chromogranin, thyroid transcription factor 1 (TTF-1), monoclonal carcinoembryonic antigen (m CEA), and calcitonin. The Ki-67 proliferation index was 50% in focal areas and averaged 20%. After surgery, histological and immunohistochemical results led to a diagnosis of medullary thyroid carcinoma. After the surgery, the patient's serum calcitonin level was found to be high (1395 ng/L). Molecular testing was conducted at the oncologist's request. Next-generation sequencing revealed a RET mutation in exon 16 (M918T) and exon 14 (V804L) of chromosome 10.

DISCUSSION

Medullary thyroid carcinoma represents a small proportion of all thyroid cancers. While not always achievable, preoperative diagnosis plays a crucial role in developing appropriate treatment strategies for patients. Clinicians or pathologists evaluating cytological samples should consider the possibility of medullary carcinoma, which could help reduce the risk of misdiagnosis. A major cause of misdiagnosis in pathological examination is the histological spectrum of thyroid medullary carcinoma, which can mimic nearly all other thyroid neoplasms [8, 9]. The presence of rare variants of thyroid medullary carcinoma adds another layer of complexity [10-12]. Azurophilic cytoplasmic granules, the presence of amyloid, plasmacytoid or spindle cells, and salt-and-pepper chromatin are the key cytologic features of medullary thyroid carcinoma. Papillary carcinoma-like nuclear features may also be observed. Cases showing all the features of 'Indented membrane, and lobulation, ground-glass chromatin pattern and nuclear grooves' in cytological examination were accepted as 'medullary thyroid carcinomas with papillary-like nuclear features' by Yamao, and in the same study it was observed that some cases could be mistakenly interpreted as papillary carcinoma in cytological examination [13].

Numerous studies have explored the prognostic features of medullary thyroid carcinoma [14-17]. Recently, pathological grading has been advocated for medullary thyroid carcinoma. Different systems have been proposed for grading [18-20]. Mitosis, necrosis, and Ki-67 proliferation index are the key parameters used in these grading systems. In recent publications, two-grade (high grade-low grade) and three-grade (low-intermediate-high grade) classification systems have been suggested for grading thyroid medullary carcinoma. Based on these classification systems, our case falls into the high-grade category, with a Ki-67 proliferation index of 20% and 15 mitoses per 2 mm². The occurrence of sporadic medullary thyroid carcinoma is linked to somatic mutations in the RET gene [21]. In our case, two mutations were identified in the RET gene: exon 16 (M918T) and exon 14 (V804L).

CONCLUSION

Although not always indicative of high-grade malignancy, thyroid medullary carcinoma exhibiting papillary-like nuclear features should be considered in the differential diagnosis of cytological samples. Immunohistochemical studies can be conducted on both cell blocks and smears, yielding reliable results.

Ethics Approval and Consent to Participate

Since this is a case report presentation, no ethics committee approval was required; however, informed consent of writing and publication was obtained from the patient prior to the preparation of the manuscript. Patient was informed about the purpose of the case report, and written informed consent was obtained from the patient for publication of this case and any accompanying pictures or data.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: AMzk; Study Design: AMzk; Supervision: AMzk; Funding: N/A; Materials: AMzk; Data Collection and/or Processing: AMzk; Statistical Analysis and/or Data Interpretation: AMzk; Literature Review: AMzk; Manuscript Preparation: AMzk; and Critical Review: AMzk.

Conflict of interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Trimboli P, Crescenzi A. Diagnosing medullary thyroid carcinoma is facilitated by measuring calcitonin in FNA washout fluids: Alea iacta est. Cancer Cytopathol. 2024;132(3):139. doi: 10.1002/cncy.22797.
2. Ogmen BE, Ince N, Aksoy Altınboga A, et al. An old friend, a new insight: Calcitonin measurement in serum and aspiration needle washout fluids significantly increases the early and accurate detection of medullary thyroid cancer. Cancer Cytopathol. 2024 Mar;132(3):161-168. doi: 10.1002/cncy.22779.
3. Liu S, Zhao H, Li X. Serum Biochemical Markers for Medullary Thyroid Carcinoma: An Update. Cancer Manag Res. 2024;16:299-310. doi: 10.2147/CMAR.S440477.
4. Nikas IP, Kazamias G, Vrontaki M, Rapti AS, Mastorakis E. Medullary thyroid carcinoma diagnosed with liquid-based cytology and immunocytochemistry. J Immunoassay Immunochem. 2022;43(5):502-515. doi: 10.1080/15321819.2022.2070025.
5. Liu CY, Chen CC, Bychkov A, et al. Constitutive Cytomorphologic Features of Medullary Thyroid Carcinoma Using Different Staining Methods. Diagnostics (Basel). 2021;11(8):1396. doi: 10.3390/diagnostics11081396.
6. Ahmed M, Abi-Raad R, Fu L, Holt EH, Adeniran AJ, Cai G. Performing Calcitonin Immunocytochemistry on an Additional ThinPrep Slide in Fine-Needle Aspiration Diagnosis of Medullary Thyroid Carcinoma. Am J Clin Pathol. 2022;157(3):426-433. doi: 10.1093/ajcp/aqab141.
7. Trimboli P, Treglia G, Guidobaldi L, et al. Detection rate of FNA cytology in medullary thyroid carcinoma: a meta-analysis. Clin Endocrinol (Oxf). 2015;82(2):280-285. doi: 10.1111/cen.12563.
8. Kang M, Kim NR, Seok JY. Non-papillary thyroid carcinoma diagnoses in The Bethesda System for Reporting Thyroid Cytopathology categories V and VI: An institutional experience. Ann Diagn Pathol. 2024;71:152263. doi: 10.1016/j.anndiagpath.2023.152263.
9. Agarwal C, Raychaudhuri S, Batra A, Pujani M, Dhingra S. Medullary carcinoma of thyroid mimicking Hurthle cell neoplasm on cytology: a diagnostic dilemma. Diagn Cytopathol. 2019;47(9):943-947. doi: 10.1002/dc.24220.

10. Kaushal S, Iyer VK, Mathur SR, Ray R. Fine needle aspiration cytology of medullary carcinoma of the thyroid with a focus on rare variants: a review of 78 cases. Cytopathology. 2011;22(2):95-105. doi: 10.1111/j.1365-2303.2010.00747.x.
11. Schreiner AM, Yang GC. Medullary thyroid carcinoma presenting as rectangular cell type on fine-needle aspiration. Diagn Cytopathol. 2009;37(3):213-216. doi: 10.1002/dc.21041.
12. Trivedi D, George L. An Aggressive Form of Medullary Thyroid Carcinoma-Melanocytic Subtype: A Case Report. Cureus. 2023;15(12):e50310. doi: 10.7759/cureus.50310.
13. Yamao N, Hirokawa M, Suzuki A, et al. High-grade medullary thyroid carcinoma with papillary-like nuclear features: A report of five cases. Diagn Cytopathol. 2024;52(1):58-64. doi: 10.1002/dc.25243.
14. Aksoy YA, Xu B, Viswanathan K, et al. Novel prognostic nomogram for predicting recurrence-free survival in medullary thyroid carcinoma. Histopathology. 2024;84(6):947-959. doi: 10.1111/his.15141.
15. Wang Z, Fan X, Zha X, et al. A Proposed Modified Staging System for Medullary Thyroid Cancer: A SEER Analysis With Multicenter Validation. Oncologist. 2024;29(1):e59-e67. doi: 10.1093/oncolo/oyad165.
16. Najdawi F, Ahmadi S, Capelletti M, Dong F, Chau NG, Barletta JA. Evaluation of grade in a genotyped cohort of sporadic medullary thyroid carcinomas. Histopathology. 2021;79(3):427-436. doi: 10.1111/his.14370.
17. Toraih E, Hussein M, Anker A, et al. Survival Outcomes of Medullary Thyroid Cancer With and Without Amyloid Deposition. Endocr Pract. 2024;30(4):311-318. doi: 10.1016/j.eprac.2024.01.001.
18. Fuchs TL, Nassour AJ, Glover A, et al. A Proposed Grading Scheme for Medullary Thyroid Carcinoma Based on Proliferative Activity (Ki-67 and Mitotic Count) and Coagulative Necrosis. Am J Surg Pathol. 2020;44(10):1419-1428. doi: 10.1097/PAS.0000000000001505.
19. Alzumaili B, Xu B, Spanheimer PM, et al. Grading of medullary thyroid carcinoma on the basis of tumor necrosis and high mitotic rate is an independent predictor of poor outcome. Mod Pathol. 2020;33(9):1690-1701. doi: 10.1038/s41379-020-0532-1.
20. Xu B, Fuchs TL, Ahmadi S, et al. International Medullary Thyroid Carcinoma Grading System: A Validated Grading System for Medullary Thyroid Carcinoma. J Clin Oncol. 2022;40(1):96-104. doi: 10.1200/JCO.21.01329.
21. Dvoráková S, Václavíková E, Sýkorová V, et al. New multiple somatic mutations in the RET proto-oncogene associated with a sporadic medullary thyroid carcinoma. Thyroid. 2006;16(3):311-316. doi: 10.1089/thy.2006.16.311.

The European Research Journal 2025;11(6):1254-1257 Case Report

DOI: https://doi.org/10.18621/eurj.1713202 *Nephrology*

Review of hematuria case-based perspective with emphasis on vascular cause: Nutcracker syndrome

Sara İleri

Department of Nephrology, University of Health Sciences, Gülhane Training and Research Hospital, Ankara, Türkiye

ABSTRACT

Hematuria, the presence of erythrocytes in urine, is classified into two main types: macroscopic, visible to the naked eye, and microscopic hematuria not visible without a microscope. A careful medical history is the best guide for evaluation. Hematuria accompanied by proteinuria or hypertension is particularly significant in differential diagnosis. Treatment options range from regular follow-up to nephrectomy; thus, an accurate and early diagnosis is essential for managing patients with hematuria. A 52-year-old Caucasian white female presented with painless hematuria. Cystoscopy was performed following the initial ultrasound evaluation. Urine cytology and pathological findings were benign. Renal function test results were within normal limits, and there was no proteinuria. Urinary sediment analysis revealed 10–15 isomorphic erythrocytes per high-power field. Glomerulonephritis was excluded in the diagnosis. Her medical history revealed a Whipple procedure performed three years earlier. So, Doppler ultrasound was performed to assess the possibility of renal vein compression. Doppler imaging revealed that the diameter of the left renal vein was 9.5 mm before the superior mesenteric artery and 1.8 mm after it, findings consistent with Nutcracker Syndrome. As her renal function was normal and she was normotensive, clinical follow-up was recommended at six-month intervals. Nutcracker Syndrome is a rare vascular compression disorder that typically involves entrapment of the left renal vein between the aorta and superior mesenteric artery. Given the risk of unnecessary interventions, this rare condition should be considered and excluded in the differential diagnosis of hematuria.

Keywords: Nutcracker syndrome, hematuria, renal vein

Hematuria, defined as the presence of blood in urine, is a common clinical finding that can indicate a wide spectrum of underlying conditions, ranging from benign causes to serious malignancies. Nutcracker syndrome (NCS) is one of the important causes of hematuria [1-4]. It is broadly classified into gross hematuria (visible to the naked eye) and microscopic hematuria (non-visible, detected only through laboratory analysis, typically with more than three red blood cells per high-power field on urine microscopy). It is further distinguished by the morphology of red blood cells and presence or absence of associated proteinuria. Systematic evaluation is essential to identify the underlying cause and guide appropriate management [5].

Current literature emphasizes that all cases of hematuria warrant investigation to identify the underlying etiology, with a particular focus on ruling out ur-

Received: June 3, 2025 **Accepted:** July 20, 2025 **Available Online:** July 28, 2025 **Published:** November 4, 2025

How to cite this article: İleri S. Review of hematuria case-based perspective with emphasis on vascular cause: Nutcracker syndrome. Eur Res J. 2025;11(6):1254-1257. doi: doi: 10.18621/eurj.1713202

Corresponding author: Sara İleri, MD., Phone: +90 312 304 20 00, E-mail: sarayavuz79@gmail.com

Available Online at https://dergipark.org.tr/en/pub/eurj

ological malignancies. The risk of malignancy is higher in patients with gross hematuria, older age, male sex, and history of smoking.

CASE PRESENTATION

A 52-year-old Caucasian white female patient was admitted to the nephrology outpatient clinic with a complaint of persistent hematuria that could not be diagnosed. Her past medical history was notable for Whipple surgery (3 years ago). She underwent total thyroidectomy for multinodular goiter one year ago, and did not use any chronic medication other than L-thyroxine. As persistent hematuria was detected during the polyclinic follow-up, cystoscopy and urine cytology were performed. Although renal function tests and ultrasound findings were within normal limits, she had mild anemia. There were 10-15 isomorphic erythrocytes in the urinary sediment at each magnification field, but proteinuria and hemoglobinuria were negative. A renal vein Doppler ultrasound (RVDUS) revealed that the left renal vein (LRV) diameter was 9.5 mm before the superior mesenteric artery (SMA) (red arrow), and it was 1.8 mm (yellow arrow) after the artery (Fig. 1).

The patient was diagnosed with NCS of the renal vein. As the patient was asymptomatic, normotensive, and had normal renal function with no proteinuria, follow-up at 6-month intervals was recommended. The decision is mainly made according to the clinical situation and the severity of LRV hypertension.

DISCUSSION

NCS, caused by extrinsic compression of the LRV, most commonly between SMA and aorta (anterior NCS), presents with isolated, painless microscopic hematuria, often isomorphic, distinguishing it from glomerular causes [1, 2]. This case is unique due to LRV compression following pancreatic surgery (Whipple procedure), a rarely reported etiology in the literature, where postoperative anatomical changes

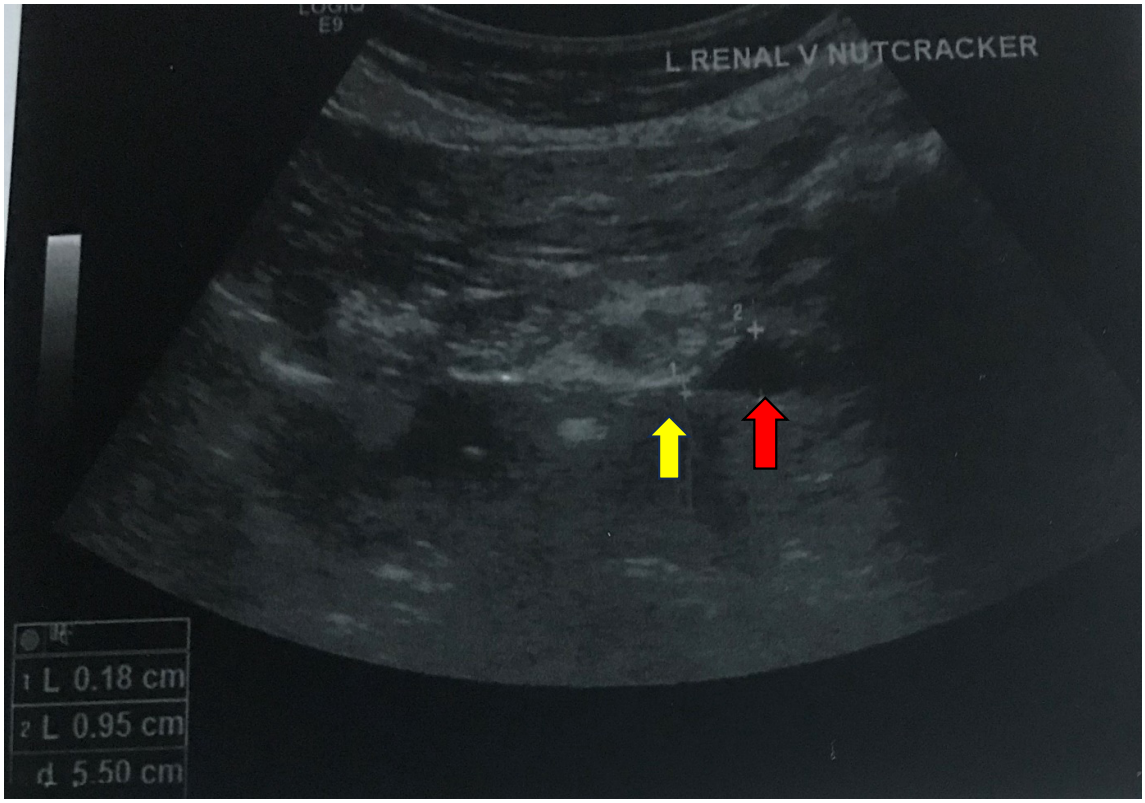

Fig. 1. Superior mesenteric artery located area. (1) Renal vein after superior mesenteric artery (yellow arrow), (2) Renal vein before superior mesenteric artery (red arrow).

predispose to NCS [3].

The clinical presentation of NCS varies widely, ranging from asymptomatic microscopic hematuria to flank pain, orthostatic proteinuria, and pelvic congestion symptoms such as varicocele and fatigue [2, 4]. Our patient was normotensive and asymptomatic aside from hematuria, consistent with mild NCS cases managed conservatively [5, 6].

Doppler ultrasound (DUS) remains a valuable, noninvasive screening tool for NCS, with diagnostic criteria including a peak systolic velocity ratio >4.7 between the aortomesenteric segment and hilar portion of the LRV or a compression ratio >5 [2]. In this case, DUS demonstrated marked LRV diameter reduction after SMA, confirming the diagnosis. However, equivocal or complex postoperative anatomy cases may require cross-sectional imaging, such as computed tomographic urography or magnetic resonance urography. Retrograde venography remains the gold standard, defined by a renocaval pressure gradient greater than 3 mmHg [2, 7].

Most literature on secondary NCS focuses on compression by tumors, lymphadenopathy, or vascular anomalies [3, 8]. Postoperative NCS is scarcely documented, making this case a notable contribution that highlights the importance of a detailed surgical history in evaluating hematuria. This aligns with recommendations from the American Urological Association guidelines advocating a risk-stratified approach incorporating patient history, imaging, and cystoscopy [9]. Management of NCS depends on symptom severity and renal function impact. Conservative management with observation and symptom control is appropriate for mild cases, especially in younger patients where spontaneous resolution occurs [5]. Surgical options, including LRV transposition or endovascular stenting, are reserved for refractory or severe symptoms but carry risks such as stent migration and thrombosis [10, 11]. Our patient's asymptomatic status justified conservative monitoring.

This case highlights the diagnostic challenge of NCS due to nonspecific symptoms and rare postoperative etiology. It reinforces the need to consider NCS in the differential diagnosis of hematuria when common causes are excluded, particularly in patients with prior abdominal surgery altering retroperitoneal anatomy. Early recognition using DUS and appropriate imaging facilitates timely diagnosis and management. But sometimes altered postoperative anatomy may limit the sensitivity of conventional imaging and DUS alone could miss or underestimate compression severity. So the cross-sectional imaging or retrograde venography might be necessary in equivocal cases.

This case differs from typical presentations in the literature mainly due to its rare postoperative etiology. While most NCS cases arise from extrinsic compression of the LRV between the SMA and aorta, this case involved LRV compression caused byanatomical changes following pancreatic surgery, which is scarcely reported in the literature [1, 6]. The case's postoperative anatomical cause of LRV compression after pancreatic surgery is rare and not well covered in standard NCS management algorithms, which typically focus on congenital or idiopathic vascular compressions or secondary causes like tumors or lymphadenopathy. This suggests a need for clinicians to consider surgical history and altered anatomy when evaluating hematuria and NCS, potentially expanding diagnostic vigilance beyond typical risk factors.

CONCLUSION

This case emphasizes the importance of considering rare vascular causes, such as Nutcracker Syndrome, in the differential diagnosis of hematuria, especially in patients with altered abdominal anatomy due to prior surgeries. Early recognition can prevent unnecessary interventions such as nephrectomy and guide appropriate long-term management.

Ethics Approval and Consent to Participate

Since this is a case report presentation, no ethics committee approval was required; however, informed consent of writing and publication was obtained from the patient prior to the preparation of the manuscript. Patient was informed about the purpose of the case report, and written informed consent was obtained from the patient prior to the case report preparation.

Data Availability

All data generated or analyzed during this study are included in this published article. The data that support the findings of this study are available on re-

quest from the corresponding author, upon reasonable request.

Authors' Contribution

Study Conception: Sİ; Study Design: Sİ; Supervision: Sİ; Funding: N/A; Materials: Sİ; Data Collection and/or Processing: Sİ; Statistical Analysis and/or Data Interpretation: Sİ; Literature Review: Sİ; Manuscript Preparation: Sİ; and Critical Review: Sİ.

Conflict of interest

The author(s) disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The author(s) disclosed that they did not receive any grant during the conduction or writing of this study.

Acknowledgments

The authors have no acknowledgments to declare.

Generative Artificial Intelligence Statement

The author(s) declare that no artificial intelligence-based tools or applications were used. The all content of the study was produced by the author(s) in accordance with scientific research methods and academic ethical principles.

Editor's note

All statements made in this article are solely those of the authors and do not represent the views of their affiliates or the publisher, editors, or reviewers. Any claims made by any product or manufacturer that may be evaluated in this article are not guaranteed or endorsed by the publisher.

REFERENCES

1. Fu WJ, Hong BF, Xiao YY, et al. Diagnosis of the nutcracker phenomenon by multislice helical computed tomography angiography. Chin Med J (Engl). 2004;117(12):1873-1875.
2. Park SJ, Shin JI. Renal Doppler ultrasonography in the diagnosis of nutcracker syndrome. Eur J Pediatr. 2013;172(1):135-136. doi: 10.1007/s00431-012-1859-5.
3. Kurklinsky AK, Rooke TW. Nutcracker phenomenon and nutcracker syndrome. Mayo Clin Proc. 2010; 85(6):552-559. doi: 10.4065/mcp.2009.0586.
4. Shokeir AA, el-Diasty TA, Ghoneim MA. The nutcracker syndrome: new methods of diagnosis and treatment. Br J Urol. 1994;74(2):139-143. doi: 10.1111/j.1464-410x.1994.tb16574. x.
5. Rista E, Llambro M, Lazaj J, Aliu D, Saliaj K. Endovascular treatment of nutcracker syndrome. Kidney Int. 2022;102(4):942. doi: 10.1016/j.kint.2022.04.038.
6. Nastasi DR, Fraser AR, Williams AB, Bhamidi V. A systematic review on nutcracker syndrome and proposed diagnostic algorithm. J Vasc Surg Venous Lymphat Disord. 2022; 10(6):1410-1416. doi: 10.1016/j.jvsv.2022.08.003.
7. Takebayashi S, Ueki T, Ikeda N, Fujikawa A. Diagnosis of the nutcracker syndrome with color Doppler sonography: correlation with flow patterns on retrograde left renal venography. AJR Am J Roentgenol. 1999;172(1):39-43. doi: 10.2214/ajr.172.1.9888735.
8. Ananthan K, Onida S, Davies AH. Nutcracker Syndrome: An Update on Current Diagnostic Criteria and Management Guidelines. Eur J Vasc Endovasc Surg. 2017;53(6):886-894. doi: 10.1016/j.ejvs.2017.02.015.
9. Mohammadi A, Ghasemi-Rad M, Mladkova N, Masudi S. Varicocele and nutcracker syndrome: sonographic findings. J Ultrasound Med. 2010;29(8):1153-1160. doi: 10.7863/jum.2010.29.8.1153.
10. Hartung O, Grisoli D, Boufi M, et al. Endovascular stenting in the treatment of pelvic vein congestion caused by nutcracker syndrome: lessons learned from the first five cases. J Vasc Surg. 2005;42(2):275-280. doi: 10.1016/j.jvs.2005.03.052.
11. Rudloff U, Holmes RJ, Prem JT, Faust GR, Moldwin R, Siegel D. Mesoaortic compression of the left renal vein (nutcracker syndrome): case reports and review of the literature. Ann Vasc Surg. 2006;20(1):120-129. doi: 10.1007/s10016-005-5016-8.