YALOVA TIP DERGİSİ

Cilt: 1 - Sayı: 1 - Ocak 2026

YALOVA MEDICAL JOURNAL

Volume: 1 - Issue: 1 - January 2026

YALOVA ÜNİVERSİTESİ

yamj.yalova.edu.tr

YALOVA ÜNİVERSİTESİ
TIP FAKÜLTESİ

YALOVA ÜNİVERSİTESİ REKTÖRÜNDEN
FROM THE RECTOR OF YALOVA UNIVERSITY

Değerli Bilim İnsanları ve Okuyucular,

Üniversiteler, yalnızca bilginin aktarıldığı eğitim kurumları değil, aynı zamanda evrensel bilimin üretildiği, tartışıldığı ve insanlığın hizmetine sunulduğu öncü merkezlerdir. Genç ve dinamik yapısıyla öne çıkan Yalova Üniversitesi olarak bilimsel düşüncenin yaygınlaşması ve akademik mirasın gelecek kuşaklara aktarılması konusunda üzerimize düşen sorumluluğun bilincindeyiz.

Tıp bilimi, doğası gereği dinamik, sürekli yenilenen ve teknolojik gelişmelerle yakından ilişkili bir alandır. İnsan sağlığını koruma ve geliştirme hedefine hizmet eden tıp alanındaki bilimsel çalışmalar, yalnızca akademik literatürü zenginleştirmekle kalmayıp doğrudan yaşamın kendisine dokunmaktadır. Bu bağlamda yayın hayatına başlayan Yalova Tıp Dergisi, üniversitemizin araştırma kültürünü ve bilimsel birikimini ulusal ve uluslararası arenaya taşıma kararlılığımızın somut bir göstergesidir.

Büyük bir emek ve özverinin ürünü olan dergimizin yayına hazırlık sürecini yöneten Editörler Kurulumuzu kutluyor, araştırmalarıyla bilimsel literatüre katkı sağlayan, bilgiyi paylaşarak çoğaltan tüm yazarlarımıza teşekkür ediyorum.

Dergimizin tıp literatürüne nitelikli bir ivme kazandırmasını, özgün çalışmalarıyla alana yön vermesini ve yayın hayatının uzun yıllar başarıyla sürmesini temenni ederim.

Yalova Üniversitesi Rektörü
Prof. Dr. Mehmet BAHÇEKAPILI

Distinguished Members of the Scientific Community and Readers,

Universities are not merely educational institutions where knowledge is imparted; they are also pioneering centers where universal science is developed, discussed, and presented to serve humanity. As Yalova University, distinguished by its young and dynamic structure, we are conscious of our responsibility to promote scientific thinking and pass on the academic legacy to future generations.

The medical sciences are, by their very nature, dynamic, constantly evolving, and closely related to advancements in technology. Scientific studies in the field of medicine, which serve the goal of protecting and improving human health, not only enrich academic literature but also directly touch life itself. In this context, the launch of the Yalova Medical Journal is a concrete manifestation of our determination to carry our university's research culture and scientific accumulation to the national and international arena.

I congratulate our Editorial Board, which managed the preparation process of our journal, a product of great effort and dedication. I also thank all our authors who contribute to the scientific literature with their research and expand knowledge by sharing it.

I hope that our journal will give a qualitative momentum to the medical literature, guide the field with its original studies, and continue its publication life successfully for many years.

The Rector of Yalova University
Prof. Dr. Mehmet BAHÇEKAPILI

YALOVA TIP DERGİSİ

YALOVA MEDICAL JOURNAL

YALOVA , TÜRKİYE | YALOVA, TURKEY

Year / Yıl: **2026** | Volume / Cilt: **1** | Issue / Sayı: **1**

Dergi Bilgileri | Journal Information

- **Web Sayfası / Web Page**:
  yamj.yalova.edu.tr
- **Dil / Language**:
  İngilizce / Türkçe | English / Turkish
- **Yayın Aralığı / Frequency**:
  Yılda üç kere yayınlanır / Published three times in a year
- **Yayın Türü / Type of Publication**:
  Süreli Yayın / Periodical
- **Telefon / Telephone**:
  +90- 226 815 68 03
- **E-posta / E-mail:**
  yamj@yalova.edu.tr
- **Yazışma Adresi / Correspondence Adress:** Yalova Üniversitesi Tıp Fakültesi, Yalova Üniversitesi Merkez Yerleşkesi, Çınarcık Yolu Üzeri 77200 Yalova, Turkiye
  Yalova University Faculty of Medicine Yalova University Central Campus, On the Çınarcık Road 77200 Yalova, Turkey
- **Journal Cover Design / Dergi Kapak Tasarımı:**
  Seda Erdokur, Bursa İletişim Ajansı.

BAŞ EDİTÖRDEN
FROM THE EDITOR IN CHIEF

Yalova Tıp Dergisi, tıp alanındaki güncel bilimsel üretimi nitelikli bir editoryal süreç içinde buluşturmayı amaçlayan bağımsız ve akademik bir platform olarak yayın hayatına başlamaktadır. Bu ilk sayı, dergimizin bilimsel standartlara bağlı, yöntemsel dikkat ve disiplinler arası etkileşimi önemseyen yaklaşımının somut bir yansımasıdır. Amacımız yalnızca yayın yapmak değil; araştırma kültürünü, eleştirel düşünceyi ve bilimsel iletişimi destekleyen sürdürülebilir bir ortam oluşturmaktır.

Bu sayıda klinik uygulamalara dayanan olgu sunumları, belirli konu başlıklarını derinlemesine ele alan derleme çalışmaları ve güncel yöntemlerin tıp alanındaki kullanımına ilişkin özgün araştırmalar yer almaktadır. Farklı uzmanlık alanlarından gelen bu katkılar, tıbbın genişleyen perspektifini ve çeşitliliğini yansıtmaktadır. Sunulan çalışmaların ortak noktası, klinik pratiğe ve bilimsel anlayışa anlamlı bir katkı sunma çabasıdır. Dergimize katkı sağlayan tüm yazarlarımıza, değerlendirme süreçlerinde görev alan akademisyenlere, yayın ve editör kurullarımıza, yardımcı editörümüze ve bölüm editörlerimize teşekkür ederim. Hazırlık sürecinde gösterdikleri dikkat ve akademik duyarlılık, bu ilk sayının niteliğine doğrudan yansımıştır.

Yalova Tıp Dergisi, ilerleyen sayılarında da genç araştırmacıları destekleyen, disiplinler arası iş birliğini teşvik eden ve tıp bilimlerine katkı sunmayı önceleyen çizgisini sürdürecektir. Bu ilk sayının, dergimizin kapsamlı ve uzun soluklu akademik yolculuğuna güçlü bir başlangıç oluşturmasını diliyor; tüm okuyucularımıza verimli bir okuma deneyimi sunmasını temenni ediyorum.

Baş Editör ve Yalova Üniversitesi Tıp Fakültesi Dekanı
Prof. Dr. Nurşah BAŞOL

Yalova Medical Journal is launching its publication life as an independent and academic platform aiming to bring together current scientific production in the field of medicine within a rigorous editorial process. This inaugural issue represents a concrete reflection of our journal's commitment to scientific standards, methodological rigor, and the value of interdisciplinary interaction. Our goal is not merely to publish articles, but to create a sustainable environment that supports research culture, critical thinking, and scientific communication.

This issue includes case reports based on clinical practice, review articles that examine specific topics in depth, and original research addressing the application of contemporary methods in medicine. Contributions from diverse fields of expertise reflect the expanding perspective and diversity of medical science. The common denominator of these studies is their effort to make a meaningful contribution to clinical practice and scientific understanding.

I would like to express my sincere gratitude to all our authors who contributed to the journal, the academicians who took part in the peer review processes, our publication and editorial boards, our associate editor, and our section editors. The care and academic sensitivity they demonstrated during the preparation process are directly reflected in the quality of this first issue.

In its forthcoming issues, Yalova Medical Journal will continue to support young researchers, encourage interdisciplinary collaboration, and prioritize contributing to the medical sciences. I hope that this inaugural issue will mark a strong beginning for the journal's comprehensive and long-term academic journey, and I wish all our readers a productive and enriching reading experience.

Editor-in-Chief and Dean of the Faculty of Medicine, Yalova University
Prof. Dr. Nurşah Başol

İÇİNDEKİLER | CONTENTS

YALOVA TIP DERGİSİ
YALOVA MEDICAL JOURNAL
ÖZGÜN MAKALE | ORIGINAL ARTICLE

Performance of Large Language Models in Ophthalmology Questions: Do Language Differences Matter?

Elvin Halili Celenk[1], Nursel Melda Yenerel[2], Peykan Turkcuoglu[3]

[1] Yalova University, Faculty of Medicine, Department of Ophthalmology, Yalova, Turkey.
[2] Medipol Acibadem Regional Hospital, Istanbul, Turkey.
[3] Veni Vidi Eye Hospital Atasehir, Istanbul, Turkey.

Abstract

Aim: To evaluate the effectiveness of ChatGPT-4.0, Copilot AI, Gemini AI, and Claude AI chatbots in answering the fundamental principles of ophthalmology.

Materials and Methods: All forty questions in the study questions section of the American Academy of Ophthalmology (AAO) 2024-2025 Basic and Clinical Science Course (BCSC) Fundamentals and Principles of Ophthalmology book were asked to ChatGPT-4.0, Copilot AI, Gemini AI, and Claude AI in both English and Turkish. The questions were asked only once and at different times of the day for each language. The responses provided by the chatbots were compared with the official answer key provided at the end of the book and classified as correct or incorrect.

Results: For English questions, Copilot AI (95%) and Claude AI (92.5%) showed higher accuracy than ChatGPT-4.0 (75%) and Gemini AI (87.5%). In Turkish questions, the accuracy rates of all models were found to be close to each other (85–92.5%), and no significant superiority was detected. Across languages, all AI applications showed medium-high consistency. The highest level of consistency was detected in Copilot AI and Claude AI.

Conclusion: Although Copilot AI and Claude AI outperformed the other two bots in English questions, they did not demonstrate a meaningful superiority over each other in Turkish questions.

Keywords: ChatGPT- 4.0, Copilot, Gemini, Claude, Ophthalmology Education

Corresponding Author: Assistant Professor, Elvin Halili Celenk

Address: *Yalova University, Faculty of Medicine, Department of Ophthalmology, Yalova, Turkey.*
E-mail: *elvin.celenk@yalova.edu.tr*

Introduction

Today, AI is used in many fields, and significant developments have been made in the field of medicine. Deep learning and broad-based language models can perform human-like language-based tasks such as text comprehension, question answering, and text generation [1,2]. These technologies also show high success in image analysis and clinical decision support processes. Ophthalmology is one of the disciplines that benefits most from AI applications, as it is a field that makes intensive use of visual data. Algorithms developed for retinal image analysis, early diagnosis of diseases, risk classification, and patient follow-up have begun to transform clinical applications.

The fundamental principles of ophthalmology are among the most important topics that must be learned during ophthalmology training. It is important for the quality of education that tools are available to help find answers to the questions we use during ophthalmology training and while preparing for exams, and that these tools ensure standardization in the answers.

The reliability of artificial intelligence chatbots in answering questions on different topics in ophthalmology is also a subject of research. Research has revealed both the benefits and shortcomings of artificial intelligence chatbots [3-7]. A study conducted to prepare for the "Fellowship of the European Board of Ophthalmology" and "Royal College of Ophthalmologists Fellowship" exams reported that chatbots were effective [3-7].

In November 2023, Microsoft renamed the new Bing to Copilot. Google renamed it Bard Gemini in 2024.
The aim of our study is to evaluate the effectiveness of ChatGPT-4.0 (OpenAI; San Francisco, USA), Copilot AI (Microsoft; Redmond, USA), Gemini AI (Google; Mountain View, USA), and Claude AI (Anthropic, USA), recently referred to as "ethical artificial intelligence," in terms of the basic principles of ophthalmology. To the best of our knowledge, there is no publication in the literature comparing the performance of these chatbots in answering basic ophthalmology questions in different languages.

Materials and Methods

All forty questions from the study questions section of the American Academy of Ophthalmology (AAO) 2024-2025 Basic and Clinical Science Course (BCSC) Fundamentals and Principles of Ophthalmology book were included in the study, and their Turkish translations were performed by a certified translator [8]. This book contains approximately the same number of questions as the other volumes in the AAO BCSC series. The accuracy and appropriateness of the translations were evaluated by an expert ophthalmologist. We did not deem it appropriate to add to these questions, adhering to the AAO basic and clinical sciences book.

Questions were asked in both English and Turkish to ChatGPT-4.0, Copilot AI, Gemini AI, and Claude AI on January 30, 2025.All artificial intelligence models included in this study were accessed via their official web-based interfaces using standard personal using accounts. Mobile applications were not utilized at any stage of the study. Regarding model settings, all AI platforms were used with their default configurations Before each question, the command "I will ask you a multiple-choice question; please provide the answer as an option" was given. After the bot responded, the session was closed, and the same command was reused for each question. Questions were asked only once and at different times of the day for each language. The responses provided by the chatbots were compared with the official answer key provided at the end of the book and classified as correct or incorrect. The evaluatıon of chatbot responces was conducted by a single experienced ophthalmologist. Since our study did not involve human or animal data, no ethics committee approval was obtained.

ChatGPT 4.0 is the latest version of GPT, released in March 2023. It is a Large Language Model (LLM)-based artificial intelligence chatbot that has access to a vast data network, is trainable, and can generate responses similar to human intelligence [8]. Copilot AI is an LMM-based AI bot used in many fields thanks to its natural language processing capability integrated with the GPT-4 AI system [9]. Gemini AI, on the other hand, uses the LaMDA language family to produce realistic language in natural language processing and strives to provide realistic responses. The 1.5 pro version was used in our study [10]. Claude AI is a generative artificial intelligence (AI) chatbot and large language model (LLM) family developed by the research company Anthropic. The Claude 3.5 sonnet version was used in our study.

Results

All 40 questions in the basic information and principles section of the ophthalmology textbook were asked in English to artificial intelligence chatbots. ChatGPT 4.0 answered 30 of these questions (75%) correctly and 10 (25%) incorrectly. Copilot AI answered 38 (95%) questions correctly and 2 (5%) incorrectly. Gemini AI answered 35 (87.5%) questions correctly and 5 (12.5%) questions incorrectly. Claude AI answered 37 (92.5%) questions correctly and 3 (7.5%) questions incorrectly (Figure 1). Copilot AI and Claude AI outperformed the other two chatbots (Chat GPT 4.0, Gemini AI) in correctly answering questions in English. A moderate level of agreement was observed between the responses of the two AI applications (κ=0.362). This agreement was found to be statistically significant (p=0.019) (Table 1).

The Turkish versions of the same questions were applied to the AI chatbots. Chat GPT 4.0 answered 35 of these questions correctly (87.5%) and 5 incorrectly (12.5%). Copilot AI answered 37 questions correctly (92.5%) and 3 incorrectly (7.5%). Gemini AI answered 34 questions correctly (85%) and 6 questions incorrectly (15%). Claude AI answered 34 questions correctly (85%) and 6 questions incorrectly (15%) (Figure 1), (Table 1). No superiority was detected among the four artificial intelligences in answering Turkish questions.

ChatGPT-4.0 provided identical answers to 33 questions of English and Turkish questions (82.5%) and different answers to 7 questions(17.5%). Of the questions with different answers, 6 were answered correctly (15%) when asked in Turkish, while 1 was answered incorrectly (2.5%) when asked in Turkish. There was no statistically significant difference between ChatGPT 4.0's success rates in answering English and Turkish questions (p=0.125). The Cohen's Kappa coefficient between ChatGPT- 4.0 (English) and ChatGPT-4.0 (Turkish) was calculated as κ=0.440, indicating that the artificial intelligence application provided consistent answers in different languages (Table 1).

Copilot AI answered 39 questions of the English and Turkish questions (97.5%) identically and 1 question(2.5%) differently. The differently answered question was incorrect when asked in Turkish. There was no statistically significant difference between Copilot AI's success rates in answering English and Turkish questions (p=0.997). The Cohen's Kappa coefficient between Copilot AI-English and Copilot AI-Turkish was calculated as κ=0.497, indicating that the AI application provided consistent answers in different languages (Table 1).

Gemini AI provided identical answers to 37 questions of the English and Turkish questions (92.5%) and different answers to 3 questions(7.5%). Of the questions with different answers, 2 were answered incorrectly (5%) when asked in Turkish, and 1 was answered correctly (2.5%) when asked in English.

Table 1. Artificial intelligence chatbots' responses to questions related to the American Academy of Ophthalmology and Ophthalmology Fundamentals and Principles of Ophthalmology and their changes.

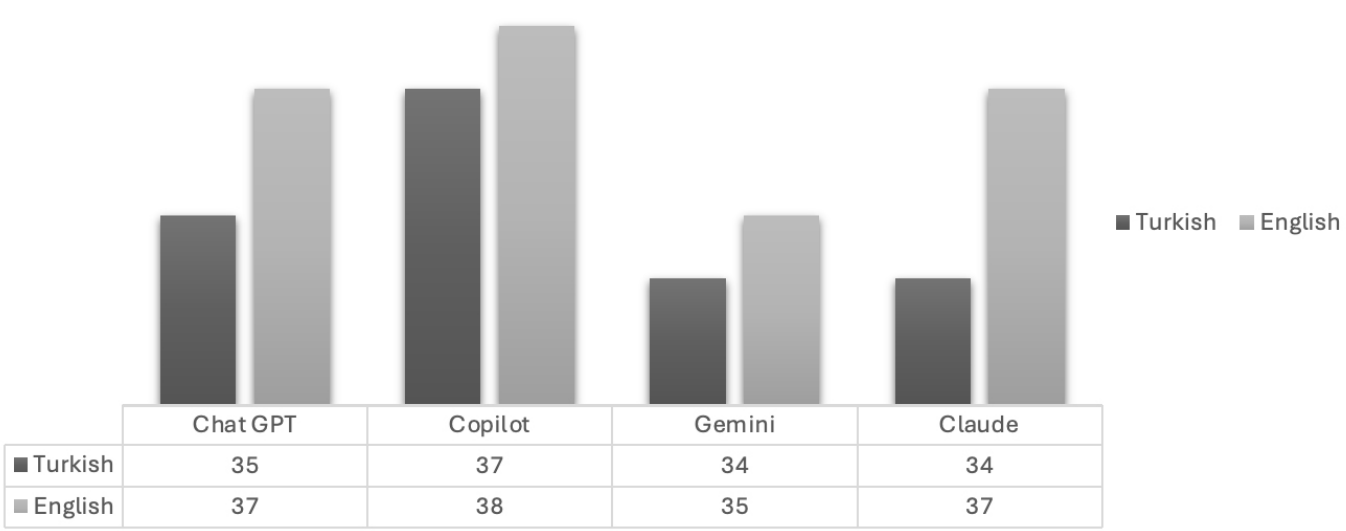
Answers	Chat GPT (English)	Chat GPT (Turkish)	Copilot (English)	Copilot (Turkish)	Gemini (English)	Gemini (Turkish)	Claude (English)	Claude (Turkish)
True	30	35	38	37	35	34	34	37
False	10	5	2	3	5	6	6	3
Giving the same answer	33 (82.5%)		39 (97.5%)		37 (92.5%)		37(92.5%)	
Giving a different answer	7 (17.5%)		1 (2.5%)		3 (7.5%)		3 (7.5%)	
True-false change	1 (2.5%)		1 (2.5%)		2 (5%)		3 (7.5%)	
False-true change	6 (15%)		0 (0%)		1 (2%)		0 (0%)	
[a]p	0.125		0.997		0.995		0.250	
κ	0.440 p=0.002**		0.497 p=0.001**		0.684 p=0.001**		0.630 p=0.001**	

*[a]Mc Nemar Testi, **p< 0,01. *p< 0.05, Cohen's Kappa katsayısı ; κ*

There was no statistically significant difference between the success rates for English and Turkish (p=0.995). The Cohen's Kappa coefficient between Gemini AI-English and Gemini AI-Turkish was calculated as κ=0.684, indicating that the AI application showed a high level of consistency in both languages (Table 1). Claude AI answered 37 questions of the same questions in English and Turkish correctly(92.5%), and 3 questions incorrectly(7.5%). Of the questions answered differently, 3 (7.5%) were answered incorrectly when asked in Turkish.

There was no statistically significant difference between Claude AI's success rates in answering English and Turkish questions (p=0.250). The Cohen's Kappa coefficient between Claude AI-English and Claude AI-Turkish was calculated as κ=0.630, indicating that the AI application provided consistent answers in different languages (Table 1).

Figure 1. Levels of Correct Answer for Turkish and English



Discussion

In the field of ophthalmology, AI- based systems are increasingly being used to assess students' knowledge levels and improve educational processes [6,7,11,12]. Various studies in the literature show that these systems perform at different levels in exams and knowledge assessment processes. Raimondi et al. achieved the highest accuracy rate of 82.9% with Bing Chat in chatbots used in the Royal College of Ophthalmologists' specialty exams in the UK, without any guidance or instruction adjustments [7]. Kung H. T. et al., in their study evaluating ChatGPT's performance on the United States Medical Licensing Examination (USMLE), showed that it achieved success close to or at the 60% accuracy threshold and could therefore assist human learners as a preliminary step toward future integration into clinical decision-making processes in medical education settings [12]. Antaki F. et al. reported that ChatGPT demonstrated promising performance in a simulated ophthalmology knowledge assessment exam and that domain-specific pre-training may be necessary to enhance performance with customized large language models (LLMs) [13]. Panthier et al., in their study examining ChatGPT's successful completion of the French version of the European Board of Ophthalmology (EBO) exam, reported a 91% success rate, indicating a high level of proficiency in ophthalmology knowledge and practice [6]. Şensoy E. et al., in their study where questions on Ophthalmic Pathology and Intraocular Tumors were asked to AI chatbots, reported that ChatGPT, Bing, and Bard answered correctly at rates of 58.6%, 63.9%, and 69.4%, respectively; However, they reported that the differences in accuracy rates between the three programs were not statistically significant ($p>0.05$) [14].

In our study, the accuracy rates of correct answers to English questions were similar across the four chatbots, with Copilot AI providing more correct answers than the others. It was determined that there was a good level of agreement between the responses given by Claude AI and Gemini to English questions (Cohen's Kappa: 0.448) and that this was statistically significant ($p=0.003$). When comparing Claude AI and Copilot, a moderate level of agreement was found (Cohen's Kappa: 0.362), which was also statistically significant ($p=0.019$). When the four chatbots were evaluated together, it was observed that Gemini AI, Copilot AI, and Claude AI, in particular, provided similar responses to questions regarding the fundamentals and principles of eye diseases.

Mihalache A. et al. evaluated the performance of Gemini and Bard in different countries via the "EyeQuiz" platform and reported that these chatbots performed acceptably in answering exam questions, with no statistically significant differences between country versions [3]. In our study, the accuracy rates of the four chatbots were similar for Turkish questions, with Copilot providing more correct answers than the others. Claude AI and ChatGPT showed a high level of agreement for the Turkish questions (Cohen's $\kappa = 0.684$, $p = 0.001$). In addition, a good level of agreement was observed between Claude AI–Turkish and Gemini AI–Turkish (Cohen's $\kappa = 0.608$, $p = 0.001$), as well as between ChatGPT–Turkish and Copilot AI–Turkish (Cohen's $\kappa = 0.448$, $p = 0.003$).

Şensoy E. et al., in their study evaluating the effects of language differences on multiple-choice questions about eye inflammation and uveitis using ChatGPT-3.5, Copilot AI, and Gemini AI, reported accuracy rates of 63.9%, 63.9%, and 50% for English questions, respectively; and 52.8%, 52.8%, and 66.7% accuracy rates for Turkish questions, respectively. The researchers noted that the AI programs had different accuracy rates when answering English and Turkish questions, but there was no statistically significant difference between their performance ($p>0.05$) [15]. In our study, four chatbots also provided correct answers at different rates when answering English and Turkish questions, but no statistically significant difference was found since the p-value was equal to the traditional threshold of 0.05.

The limitations of this study include the use of a limited number of questions and the lack of comparative analysis on specific subtopics within the field. It would be beneficial to conduct assessments in different languages using a larger number of questions.

In conclusion, this study evaluated not only the performance of AI chatbots but also the existence of performance differences between the English and Turkish versions of the same questions. While Copilot AI and Claude AI outperformed the other two bots on English questions, they did not demonstrate a significant advantage over each other on Turkish questions. We believe that the small number of questions may have affected the statistical significance of the results. Our findings show that AI-based chatbots can be used as auxiliary tools in ophthalmology education, particularly facilitating students' quick access to information during the learning of basic principles and exam preparation processes. Clinically, it can be said that these systems are not yet mature enough to be used as decision support mechanisms, but they have the potential to serve as additional tools that will reduce the burden on clinicians in the future. Therefore, both the development of customized models that will provide greater accuracy and a critical approach by users are of great importance for safe integration in education and clinical practice. We believe that this topic requires further evaluation through comprehensive studies that address additional questions.

Acknowledgments and Funding Statement: Not applicable

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

Author Contributions: Concept: E. H. C.; Design: E. H. C.; Supervision: N. M. Y., P. T.; Materials: E. H. C.; Data Collection and/or Processing: E. H. C.; Analysis and/or Interpretation: E. H. C., N. M. Y.; Literature Review: E. H. C., P. T; Writing: E. H. C.; Critical Review: E. H. C.

References

1. Flanagin A, Bibbins-Domingo K, Berkwits M, Christiansen SL. Nonhuman "authors" and implications for the integrity of scientific publication and medical knowledge. JAMA. 2023;329(8):637–9. doi:10.1001/jama.2023.1344.

2. Rahimy E. Deep learning applications in ophthalmology. Curr Opin Ophthalmol. 2018;29(3):254–60. doi:10.1097/ICU.0000000000000470.

3. Mihalache A, Popovic MM, Muni RH. Performance of an artificial intelligence chatbot in ophthalmic knowledge assessment. JAMA Ophthalmol. 2023;141(6):589–97. doi:10.1001/jamaophthalmol.2023.1144.

4. Canleblebici M, Dal A, Erdağ M. Evaluation of the performance of large language models (ChatGPT-3.5, ChatGPT-4, Bing and Bard) in Turkish ophthalmology chief-assistant exams: a comparative study. Turk Klin J Ophthalmol. 2024;33(3):163–70. doi:10.5336/ophthal.2024-102632.

5. Tan TF, Thirunavukarasu AJ, Campbell JP, Keane PA, Pasquale LR, Abramoff MD, et al. Generative artificial intelligence through ChatGPT and other large language models in ophthalmology: clinical applications and challenges. Ophthalmol Sci. 2023;3(4):100394. doi:10.1016/j.xops.2023.100394.

6. Panthier C, Gatinel D. Success of ChatGPT, an artificial intelligence language model, on the French version of the European Board of Ophthalmology examination: a new approach to medical knowledge assessment. J Fr Ophtalmol. 2023;46(7):706–11. doi:10.1016/j.jfo.2023.05.006.

7. Raimondi R, Tzoumas N, Salisbury T, Di Simplicio S, Romano MR; North East Trainee Research in Ophthalmology Network (NETRiON). Comparative analysis of major language models in Royal College of Ophthalmologists fellowship examinations. Eye (Lond). 2023;37(17):3530–3. doi:10.1038/s41433-023-02563-3.

8. Waisberg E, Ong J, Masalkhi M, Zaman N, Sarker P, Lee AG, et al. Google's AI chatbot "Bard": a side-by-side comparison with ChatGPT and its utilization in ophthalmology. Eye (Lond). 2024;38(4):642–5. doi:10.1038/s41433-023-02760-0.

9. Microsoft. Bing Chat | Microsoft Edge. Accessed July 4, 2024. Available from: https://www.microsoft.com/en-us/edge/features/bing-chat.

10. Google. Google AI updates: Bard and new AI features in Search. Accessed July 4, 2024. Available from: https://blog.google/technology/ai/bard-google-ai-search-updates.

11. Kleinig O, Gao C, Kovoor JG, Gupta AK, Bacchi S, Chan WO. How to use large language models in ophthalmology: from prompt engineering to privacy protection. Eye (Lond). 2024;38(4):649–53. doi:10.1038/s41433-023-02772-w.

12. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on the USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health. 2023;2(2):e0000198. doi:10.1371/journal.pdig.0000198.

13. Antaki F, Touma S, Milad D, El-Khoury R, Duval R. Evaluating the performance of ChatGPT in ophthalmology: an analysis of its successes and shortcomings. Ophthalmol Sci. 2023;3(4):100324. doi:10.1016/j.xops.2023.100324.

14. Sensoy E, Citirik M. A comparative study on the knowledge levels of artificial intelligence programs in diagnosing ophthalmic pathologies and intraocular tumors. Int Ophthalmol. 2023;43(12):4905–9. doi:10.1007/s10792-023-02893-x.

15. Şensoy E, Çıtırık M. Performance analysis of ChatGPT-3.5, Copilot and Gemini in multiple choice questions on ocular inflammation and uveitis: the effect of language differences. Turk Klin J Ophthalmol. 2025;34(1):12–6.

VAKA TAKDİMİ | CASE REPORT

Low Dose Modafinil-Associated Visual Hallucinations: A Case Report

Musa SAHPOLAT[1]

[1] *Yalova University, Faculty of Medicine, Department of Mental Health and Diseases, Yalova, Turkey.*

Abstract: Modafinil, a wake-promoting agent, is generally considered safe. We report a 32-year-old male with narcolepsy who developed visual hallucinations (VHs) after one month on a low dose (100 mg/day) of modafinil. This suggests modafinil may cause VHs even at low therapeutic doses.

Objective: To present a case of visual hallucinations (VHs) associated with low-dose modafinil use (100 mg/day) in a patient with narcolepsy, and to emphasize the importance of clinical awareness regarding this potential adverse effect.

Case: A 32-year-old male with narcolepsy developed daily VHs and aggressive behavior after one month of treatment with 100 mg/day modafinil. The VHs resolved completely after modafinil was discontinued and risperidone treatment was initiated.

Conclusion: Modafinil use, even at low therapeutic doses, may be associated with the emergence of visual hallucinations. Clinicians should be aware of this potential adverse drug reaction.

Keywords: Adverse Reactions, Drug-Related Side Effects, Hallucination, Modafinil

Corresponding Author: Assistant Professor, Musa Sahpolat M.D.

Address: *Yalova University School of Medicine, Central Campus, On Çınarcık Road, 77200, Yalova, Türkiye*
E-mail: *musa.sahpolat@yalova.edu.tr*
Contact Number: *+905077517606, ORCID ID: 0000-0002-0022-2389*

Abstract

Modafinil is a new non-sympathomimetic drug with neurochemical and behavourial effects, distinct from those of amphetamine. The Food and Drug Agency (FDA) of United States of America has approved to provide wakefulness in patients who were treated for narcolepsy, or obstructing sleep apnea or excessive sleepiness due to shift-based working. However, potential psychotic symptoms due to adverse effects of modafinil have not been investigated. We present a 32-year-old male who had been followed up with a diagnosis of narcolepsy, whom was on 100 mg/day low dose modafinil for a month and presented with visual hallucinations (VHs). We suggest that modafinil may have hallucinations side effects also in low therapeutic doses which clinicians should be aware of.

Introduction

Modafinil is a new non-sympathomimetic drug with neurochemical and behavourial effects, distinct from those of amphetamine. The Food and Drug Agency (FDA) of United States of America has approved to provide wakefulness in patients who were treated for narcolepsy, or obstructing sleep apnea or excessive sleepiness due to shift-based working [1,2]. There are evidences indicating that modafinil is effective in treatments of some psychiatric diseases such as especially attention deficit and hyperactivity disorder, depression, schizophrenia, also to excessive sleepiness during daytime and narcolepsy [3,4]. Even though the exact mechanism of modafinil is still unknown. Its pharmacological profile is different from conventional psychostimulants. It has fewer side effects (anxiety, motor hyperactivity, irritability) than conventional stimulants. Daily recommended dose is 200-400mg/day as one or in two divided doses [4].

We present a 32-year-old male who had been followed up with a diagnosis narcolepsy, whom was on 100 mg/day low dose modafinil for a month and presented with visual hallucinations (VHs). We suggest that modafinil may have hallucinations side effects also in low therapeutic doses.

Case

A 32-year-old male who had been followed up with a diagnosis of narcolepsy suffering from visual hallucinations (VHs) and aggressive behavior was admitted to the Department of Psychiatry. He had no history of substance use disorder. In his medical history, he received low doses modafinil (100mg/day) for narcolepsy. He reported that VHs occurred almost every day while taking modafinil. There was no history of use of additional medication (psychotropic and other drugs) in combination with modafinil. In his routine laboratory tests, fasting blood glucose, electrolytes, whole blood count, renal and liver and thyroid function tests were normal with normal vitamin B12 and folic acid levels. In the psychiatric examination, visual signs were slightly predominant VHs, and he did not have any signs of mania, delusion and auditory hallucination. He had no other comorbid psychiatric disease. Neurological system examination was normal, and he had no chronic systemic disease. The ophthalmic consult was recommended to exclude organic causes for presenting symptoms, his ophthalmic examination was normal. No cause which could explain the etiopathogenesis of VHs could be found and it was thought that this was probably related with use of modafinil. After discharge, modafinil was discontinued and risperidone was initiated. The dose of risperidone was gradually increased up to 3 mg/day. VHs symptoms were not observed at monthly follow-up visits for six months after initiation of risperidone. His psychiatric and neurological examinations were found to be normal after risperidone treatment. The patient is currently being followed up with regular follow-up visits.

Discussion

The mechanism of action of modafinil, in addition to its potential mechanism of inducing psychosis, is still unknown. This report describes a case of VHs associated with modafinil use. The VHs appeared 3 to 4 weeks after taking modafinil, lasted for the duration of the dose, and completely disappeared 5 to 6 days after stopping the medication. These points support that the VHs were actually associated with the use of modafinil.

In the literature, psychosis and mania cases were noticed to take modafinil at a dose of 200–400 mg/day or higher [5,6]. It was reported that exacerbation of psychotic symptoms was observed in some schizophrenic patients who have taken modafinil due to sedative side effects of antipsychotic drugs [7-9]. Our case did not have any physical or psychiatric illness, he did not use any drug or substance during modafinil treatment, and he was not a shift worker. After cessation of modafinil and initiation of antipsychotic treatment, his symptoms were completely resolved. No psychiatric symptom was seen thereafter.

Its potential mechanism of inducing psychosis is still unknown. It has been stated that modafinil may restrain the release of y-aminobutyric acid (GABA) resulting in a loss of inhibition of the excitatory cholinergic and glutaminergic pathways resulting in psychosis [10,11]. It has also been recommended that modafinil increases dopamine levels via inhibition of GABA release and owing to weak inhibition of dopamine reuptake thus strengthening psychosis [9,11]. Further large-scale studies are required to conducted regarding the mechanisms of action of modafinil and its potential to induce psychosis particularly hallucination. According to the adverse drug reaction (ADR) assessment test performed in this patient, the possibility that the advers drug reaction occurred in relation with paliperidone was in the "probable" category by Naranjo's ADR probability score.

In conclusion, we suggest that modafinil might be associated the emergence of VHs. Even though the pathophysiology of modafinil-associated VHs remains still unclear, we suggest that modafinil may have hallucinations side effects also in low therapeutic doses which clinicians should be aware of.

Acknowledgements and funding statement
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest
The authors declare that there is no conflict of interest.

Author Contributions
Plan, data collection, data analysis and comments writing and corrections: MS

Ethics and consent to participate
This study was conducted in accordance with ethical standards, and written informed consent was obtained from the patient.

References

1. Ozsoy S, Aydin D, Ekici F. Effects of modafinil on pentylenetetrazol-induced convulsive epilepsy. Bratisl Lek Listy 2015; 116 (3): 162-166.

2. Swapnajeet S, Subodh B, Gourav G. Modafinil dependence and hypersexuality: a case report and review of the evidence. Clin Psychopharmacol Neurosci. 2016; 14: 402-404.

3. Ballon JS, Feifel D. A systematic review of modafinil: potential clinical uses and mechanisms of action. J Clin Psychiatry 2006; 67: 554-566.

4. Schwartz JR. Modafinil: new indications for wake promotion. Expert Opin Pharmacother 2005; 6: 115-129.

5. Wu P, Jones S, Ryan CJ, Michail D, Robinson TD. Modafinil-induced psychosis. Intern Med J. 2008; 38: 677-678.

6. Mariani JJ, Hart CL. Psychosis associated with modafinil and shift work. Am J Psychiatry. 2005; 162: 1983.

7. Özer S, Demir B. Hypomania in a Schizophrenic Patient Treated with Modafinil for Clozapine-Induced Sedation. Arch Neuropsychiatr. 2010; 47: 171-173.

8. Aggarwal A, Garg A, Jiloha RC. Psychotic relapse in a patient with schizophrenia associated with modafinil therapy. Primary Psychiatry. 2009; 16: 26-27.

9. Narendran R, Young CM, Valenti AM, Nickolova MK, Pristach CA. Is psychosis exacerbated by modafinil? Arch Gen Psychiatry. 2002; 59: 292-293.

10. McClellan KJ, Spencer CM. Modafinil: A review of its pharmacology and clinical efficacy in the management of narcolepsy. CNS Drugs. 1998; 9(4): 311-324.

11. Wolf J, Fiedler U, Anghelescu I, Schwertfeger N. Manic switch in a patient with treatment-resistant bipolar depression treated with modafinil. Journal of Clinical Psychiatry. 2006; 67(11): 1817.

VAKA TAKDİMİ | CASE REPORT

Ruptured Non-Coronary Sinus of Valsalva Aneurysm Fistulizing into the Left Atrium: A Rare Case Presentation

Mehmet Hasan Ozdil[1], Ahmet Ferhat Kaya[1], Ismail Ungan[2], Cemalettin Yilmaz[2]

[1] Mus State Hospital, Department of Cardiology, Mus, Turkey.
[2] Yalova University, School of Medicine, Department of Cardiology, Yalova, Turkey.

Abstract

Objective: We report a rare case of a ruptured sinus of valsalva aneurysm originating from the non-coronary sinus and fistulizing into the left atrium, presenting with acute heart failure symptoms.

Case: A 54-year-old male patient was admitted with signs and symptoms of heart failure. Transthoracic echocardiography revealed turbulent color flow from the non-coronary sinus (NCS) to the left atrium. An aneurysmatic appearance with contour irregularity of the sinus of Valsalva within the NCS was observed in the parasternal long-axis view. Sinus of Valsalva aneurysms (SVAs) can result from congenital or acquired factors. Aneurysmatic tissue may develop due to dysplasia of the sinus of Valsalva during embryonic development, and rupture can occur following infective endocarditis, intense physical activity, atherosclerosis, or trauma exposure.

Conclusion: The non-coronary sinus is the second most common site for the development of aneurysms after the right coronary sinus, and SVA ruptures frequently fistulize into the right heart cavities. Hence, our patient's diagnosis of a non-coronary SVA ruptured into the left atrium (LA) represents an even rarer occurrence, which warrants reporting.

Keywords: Rupture of Sinus of Valsalva, Fistulization into the Left Atrium, Heart Failure

Introduction

The sinus of Valsalva is a pouch-like structure that is part of the aorta, and it serves as a junction between the aorta and the heart chambers. When it ruptures, it can lead to serious cardiovascular problems. This report describes a case of a 54-year-old male patient who presented with symptoms of heart failure due to the rupture of the sinus of Valsalva.

Case

A 54-year-old male presented to our clinic with dyspnea, orthopnea and fatigue for 15 days. Physical examination revealed tachypnea, blood pressure of 100/65 mmHg, heart rate of 121 bpm, and SpO2 of 92%. A cardiac exam showed a continuous murmur at the left parasternal border. The electrocardiogram (ECG) showed sinus tachycardia, and the NT-proBNP level was elevated at 1240 ng/L. Transthoracic echocardiography showed an enlarged left atrium (LA) (5.2 cm) and left ventricle (diastolic diameter: 6.8 cm), and left ventricle ejection fraction of 45%. Turbulent color flow from the non-coronary sinus (NCS) to the LA was observed. An aneurysmatic appearance with contour irregularity of the sinus of Valsalva (SV) in the NCS was evident (Figure 1A). A systolic-diastolic flow pattern extending from the NCS to the LA was visualized in various views (Figure 1B, 1C, 1D). Contrast-enhanced computed tomography revealed calcification at the SV level, SV wall irregularity, and fistula extending from NCS to the LA (Figure 2). Due to rupture of the sinus of Valsalva aneurysm (SVA) with fistulization into the LA, the patient underwent cardiovascular surgery.

Corresponding Author: Assistant Professor, Cemalettin Yilmaz

Address: *Yalova University, School of Medicine, Department of Cardiology, Yalova, Turkey.*
E-mail: *cemalettin.yilmaz@yalova.edu.tr*

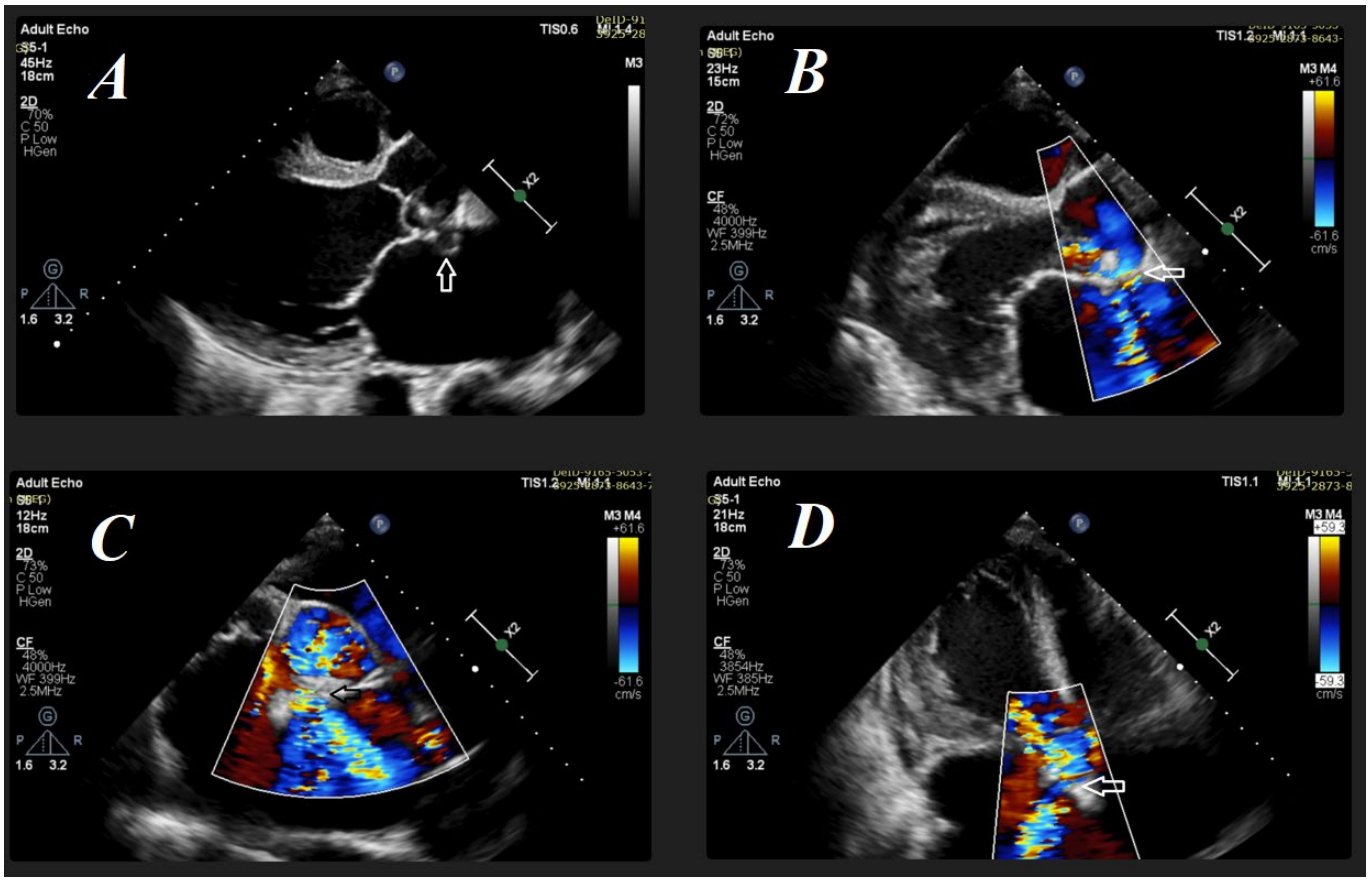


Figure 1. *(A) Echocardiography revealing sinus of Valsalva dilatation and irregularity of the sinus of Valsalva wall in the parasternal long axis view. The arrow indicates the site of the ruptured sinus of Valsalva aneurysm (SVA). (B, C, D) Echocardiography with color Doppler demonstrating a high-velocity aliasing mosaic of blood flow from the non-coronary sinus (NCS) to the left atrium (LA). The arrow indicates the color flow from the ruptured SVA to the LA.*

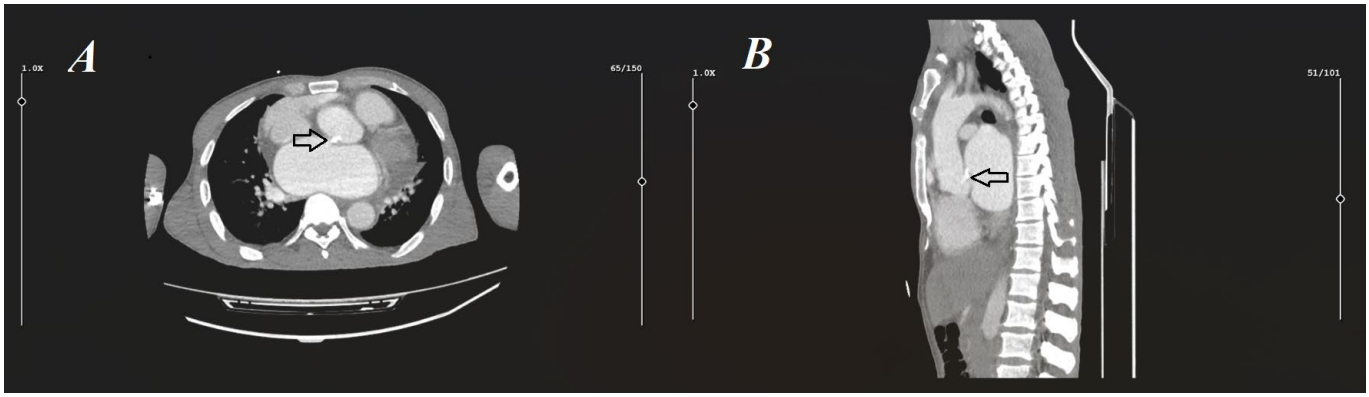


Figure 2. *Contrast-enhanced computed tomography revealing calcification at the level of the sinus of Valsalva, irregularity of the sinus of Valsalva wall, and a fistula extending from the aorta to the left atrium.*

Discussion

SVA may result from congenital or acquired causes [1]. Dysplasia of SV tissue during embryonic development may lead to aneurysm rupture due to infective endocarditis, intense physical activity, atherosclerosis, or trauma exposure [2]. Untreated ruptured aneurysms may lead to heart failure and increased mortality [3]. In this case, blood cultures showed no growth, and there was no history of trauma, suggesting a congenital origin.

In this specific case, transthoracic echocardiography was used to diagnose the condition. The imaging showed turbulent color flow from the non-coronary sinus to the left atrium, indicating a problem with the SV. Additionally, an aneurysmatic appearance and contour irregularity of the SV were observed in the NCS in the parasternal long-axis view.

What makes this case even rarer is that the patient had a rupture of the NCS into the LA. Typically, ruptures of the SV often fistulize into the right heart cavities. Therefore, this case is noteworthy due to the unusual site of rupture.

Conclusion

In summary, this report highlights the diagnosis and presentation of a patient with a non-coronary SVA that had ruptured into the LA. This case underscores the importance of early diagnosis and intervention in such cases to prevent potentially life-threatening complications and serves as a valuable contribution to the medical literature.

Acknowledgements and Funding Statement

The authors received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors would like to thank all healthcare personnel involved in the diagnosis, treatment, and follow-up of the patient.

Conflicts of Interest

The authors declare that they have no conflicts of interest relevant to this manuscript.

Author Contribution

All authors contributed to the study conception, data collection, analysis, and interpretation. All authors participated in drafting or revising the manuscript, approved the final version, and agree to be accountable for all aspects of the work.

Ethics and Consent to Participate

Written informed consent was obtained from the patient for publication of this case report and accompanying images.

References

1. Duval JL, Ramsingh RA, Rahaman NC, Rampersad RD, Angelini GD, Teodori G. Rupture of sinus of Valsalva aneurysm: case report and review of contemporary literature. Perfusion. 2021;36(8):883-886.doi:10.1177/0267659120966915.

2. Wingo M, de Angelis P, Worku BM, Leonard JR, Khan FM, Hameed I, et al. Sinus of Valsalva aneurysm repairs: Operative technique and lessons learned. J Card Surg. 2019;34(6):400-403. doi:10.1111/jocs.14041

3. Moustafa S, Mookadam F, Cooper L, Adam G, Zehr K, Stulak J, et al. Sinus of Valsalva aneurysms--47 years of a single center experience and systematic overview of published reports. Am J Cardiol. 2007;99(8):1159-1164. doi:10.1016/j.amjcard.2006.11.047

YALOVA TIP DERGİSİ | YALOVA MEDICAL JOURNAL

Yalova Tıp Dergisi 1(1):19-23, 2026 | Yalova Medical Journal 1(1):19-23, 2026

Cilt: 1 - Sayı: 1 - Ocak 2026 | Volume: 1 - Issue: 1 - January 2026

DERLEME | REVIEW

Güncel Bilgiler Eşliğinde Trombolitik Tedavi

Nurşah BAŞOL[1]

[1] *Yalova Üniversitesi, Tıp Fakültesi, Acil Tıp AD, Yalova, Türkiye*

Özet

Trombolitik tedavi, damar içi trombüsün farmakolojik olarak eritilmesini sağlayan ve özellikle hayatı tehdit eden tromboembolik durumlarda kullanılan önemli bir tedavi yöntemidir. Fibrinolitik tedavi olarak da bilinen bu yaklaşım, plazminojenin plazmine dönüştürülmesi yoluyla fibrinin parçalanmasını sağlar ve kardiyovasküler acillerde erken reperfüzyonun sağlanmasında kritik bir rol oynar. Günümüzde trombolitik uygulamalar; pulmoner emboli (PE), akut miyokard infarktüsü (AMI) ve akut iskemik inme (Aİİ) başta olmak üzere çeşitli arteriyel ve venöz tıkanıklıklarda kullanılmaktadır. Bu tedavide zamanın belirleyici olması nedeniyle erken tanı, endikasyonların doğru değerlendirilmesi ve kontrendikasyonların hızlıca dışlanması özellikle acil serviste büyük önem taşır. Bu derlemede, son yıllarda yayınlanan güncel kılavuzlar ışığında PE, AMI ve Aİİ'de trombolitik tedavi önerileri, ajan seçimi, uygulama stratejileri, zaman penceresi ve özel durumlara yönelik yeni yaklaşımlar ayrıntılı olarak ele alınmaktadır.

Anahtar Kelimeler: Akut iskemik inme, Fibrinolitik ajanlar, Trombolitik tedavi

Abstract

Thrombolytic therapy is a pharmacological method used to dissolve intravascular thrombi and plays a critical role in the management of life-threatening thromboembolic conditions. Also known as fibrinolytic therapy, it promotes the conversion of plasminogen to plasmin, leading to fibrin degradation and restoration of perfusion. Today, thrombolytic agents are widely used in several major clinical emergencies, particularly pulmonary embolism (PE), acute myocardial infarction (AMI), and acute ischemic stroke (AIS). Because time is crucial in achieving optimal outcomes, early diagnosis, rapid determination of indications, and swift exclusion of contraindications are essential—especially in emergency departments. This review summarizes the most recent guideline recommendations on thrombolytic therapy in PE, AMI, and AIS, focusing on indications, contraindications, agent selection, dosage strategies, timing of administration, and new imaging-based approaches for extended therapeutic windows. Advances in clinical evidence and ongoing large-scale studies continue to refine the role of thrombolysis in modern acute care.

Keywords: Acute ischemic stroke, Fibrinolytic agents, Thrombolytic therapy

Sorumlu Yazar: Prof. Dr. Nurşah Başol, M.D.

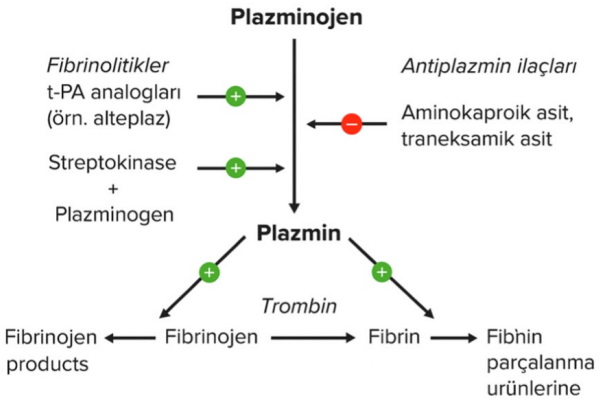
Addres: *Yalova Üniversitesi, Tıp Fakültesi, Çınarcık Yolu, 77200, Yalova, Türkiye*
E-mail: *nursah.basol@yalova.edu.tr*
ORCID ID: *0000-0002-0883-6553*

Giriş

Trombolitik tedavi; trombolitik bir ajanla mevcut trombüs yani pıhtının eritilmesi tedavisine denir ve fibrinolitik tedavi olarak da bilinir [1]. Venöz veya arteriyel tromboembolik şikayetlerin neden olduğu her türlü klinik durumda uygulanabilir. Tromboz normalde herhangi bir damar yaralanması sonrası kanamayı sınırlamak adına vücudun sergilediği bir fizyolojik yanıttır. Yaralanmanın olmadığı durumlarda gerçekleşen tromboz ise pek çok hayatı tehdit eden klinik duruma neden olabilir. Bunların başında venöz tromboemboliler, myokard infarktüsü ve akut iskemik inme gelmektedir.

Trombolitik ajanlar fibrin spesifik olanlar ve fibrin spesifik olmayanlar olarak iki gruba ayrılabilir. Fibrin spesifik olanlar doku plazminojen aktivatörleridir ve klinik kullanımda t-PA olarak bilinir. Bunlar; alteplaz, reteplaz ve tenekteplazdır. Fibrin spesifik olmayan ajanlar ise streptokinaz, anistreplaz, ürokinaz ve proürokinazdır [2]. Trombolitik ajanlar, plazminojeni plazmine çevirerek oluşan fibrinin parçalanmasını sağlar. (Şekil-1) Oluşan plazmin, fibrin, fibrinojen, protrombin ve Faktör 5 ve 8'i sindiren proteolitik bir ajana dönüşür. Taze fibrin parçaları, fibrin yıkım ürünlerinin ortaya çıkması ile parçalanır. Ürokinaz peptit bağı ile plazminojen molekülüne iki farklı taraftan bağlanarak onu aktive eder. Yarı ömrü 16 dakikadır. Streptokinaz ise indirek aktivatördür. Başlangıçta bir aktivatör kompleksi oluşturmak için plazminojen ile birleşir. Bu kompleks daha sonra serbest plazminojeni plazmine dönüştürür. Streptokinazın bu mekanizmalarla biri 18 dakika, diğeri 83 dakikaklık iki yarı ömrü mevcuttur [3].

Trombolitik ajanlar; periferal bir damar içi uygulama ile sistemik olarak dolaşıma verilebileceği gibi bir kateter yardımıyla direkt olarak pıhtının üzerine de uygulanabilir. İlk olarak 1946'da Tillet ve Sherry'nin streptokinazın pıhtılar üzerindeki etkilerini incelemesi ile uygulanmaya başlanmış ve ilk akut myokard infarktüsünde (AMI) hastalar üzerinde kullanılmıştır [4]. Günümüzde klinikte trombolitik uygulanan durumlar; pulmoner emboli (PE), akut myokard infarktüsü (AMI), akut iskemik inme (Aİİ), derin ven trombozu (DVT), akut periferal arter tıkanıklığı, intrakardiyak thrombus formasyonu ve kateter tıkanıklıklarıdır. Bu yazıda son kılavuzlar eşliğinde PE, AMI ve Aİİ'de trombolitik tedavi uygulamalarından bahsedilecektir.



Şekil-1 *Fibrin Oluşum Yolu*

Pulmoner Embolide Trombolitik Tedavi

Sistemik trombolitik tedavi PE'nin rezolüsyonunu hızlı bir şekilde sağlar ve pulmoner arter basıncını düşür, arteriyel oksijenizasyonu arttırır ve perfüzyon hasarını düzeltir. Bunun yanında kanama riskini arttırması hasta seçiminin iyi yapılmasını beraberinde getirir. 2021 yılında Venöz Tromboemboli'de (VTE) trombolitik tedaviler üzerine 2016 yılındaki yayınlanan son kılavuzun üzerine güncellemeler yayınlanmıştır [5]. Buna göre maddeler halinde PE tedavisi hakkında kılavuzdaki mevcut değişikliklerden bahsedilecektir.

1- Hipotansiyon ile ilişkili akut PE hastalarında eğer yüksek kanama riski yoksa, hiç tedavi uygulamamak yerine sistemik antitrombotik tedavi uygulanması önerilmektedir (zayıf öneri, düşük kanıt düzeyi).

2- Hipotansiyon ile ilişkili olmayan akut PE hastalarında, sistemik trombolitik tedavinin uygulanmasının karşısında durulmaktadır (güçlü öneri, düşük kanıt düzeyi). Bu öneri bir önceki kılavuzda zayıf öneri iken bu kılavuzda güçlü öneri olarak güncellenmiştir. Yazarlar, antitrombotik tedavinin bu hasta grubunda faydalarının değişken olup potansiyel zararlarının ön planda olması nedeniyle bu değişikliğe gerek görmüştür.

3- Anti-koagülan tedavi sonrası kötüleşme görülen ama henüz hipotansiyon gelişmeyen hastalarda kanama riski kabul edilebilir düzeyde ise sistemik antitrombotik tedavi önerilmektedir (zayıf öneri, düşük kanıt düzeyi). Klinik kötüleşme kriterleri olarak; sistolik Kan Basıncında (SKB) düşme, kalp hızında azalma, gaz değişim bozukluğu, yetersiz perfüzyon bozukluğu, sağ ventriküler fonksiyon bozukluğu veya kardiyak biyomarkerlarda yükseklik olarak belirtilmiştir. Bu kötüleşme kriterleri şokun bariz bulguları olmasa bile göz önünde bulundurulmalıdır.

Uzamış hipotansiyon ile seyreden akut PE hastalarında kontrendikasyon yoksa trombolitik uygulamak önerilmektedir. Hipotansiyonun olmadığı akut PE hastalarında tromboliz ile kardiyovasküler kollaps riskinde düşme olsa da majör kanama riskinin artmış olması, kâr zarar oranına bakıldığında trombolitik tedavi lehine bir karar aldırmamaktadır. 2020 Nice kılavuzunda hemodinamik instabilite halinde sistemik trombolitik tedavi önerilir. Hemodinamik olarak stabil olan hastalarda ise sağ ventriküler disfonksiyon bulguları gelişmedikçe sistemik trombolitik uygulaması önerilmez [6].

Kateter destekli trombolitik uygulamaları hakkındaki birkaç öneri şöyledir;

1- Trombolitik ajan ile tedavi edilen akut PE'li hastalarda kateter destekli tedavi yerine sistemik trombolitik tedavi önerilir (zayıf öneri, düşük kanıt düzeyi). Hipotansiyonu olan akut PE hastalarında a) yüksek kanama riski b) sistemik trombolizisin başarısızlığı c) sistemik tromboliz etki gösterene kadar ölüm olasılığı yüksek olan şoktaki hasta gruplarında eğer yeterli deneyim ve kaynaklar mevcutsa kateter destekli trombolitik önerilir (zayıf öneri, düşük kanıt düzeyi).

2- 2019 yılındaki ESC kılavuzunda kateter destekli trombolitik tedavisi sistemik

trombolizisin etkisiz olduğu veya kontrendike olduğu yüksek riskli PE hastalarında önerilmektedir. 2020 NICE kılavuzu ise sadece sistemik trombolitik tedaviyi önerir [7].

PE'de kullanılan trombolitik ajanların uygulama şekli sistemik intravenöz infüzyon şeklindedir. En sık kullanılan ilaç alteplaz (t-PA) olup standart doz 100 mg/2 saat i.v infüzyon şeklindedir. Kateter ile lokal tromboliz uygulanacaksa, ilaç daha düşük dozlarda (genellikle 10–20 mg bolus ardından düşük doz infüzyon) doğrudan pulmoner arter içine uygulanır. İlacın uygulanmasından önce ve sonra hastanın hemodinamik durumu, kanama risk faktörleri ve koagülasyon parametreleri yakından takip edilmelidir.

Akut Myokard İnfarktüsünde Trombolitik Tedavi

Fibrinolitik tedavi; ST elevasyonlu AMI hastalarında primer Perkutan Koroner Girişim'in (PKG) zamanında yapılamayacağı durumlarda düşünülmesi gereken önemli bir reperfüzyon stratejisidir. Fibrinolitik tedavi, semptom başlangıcının ilk 12 saatinde olan ST elevasyonlu AMI hastalarında eğer 120 dakika içerisinde primer PKG uygulanamayacaksa önerilmektedir. Fibrinolitik tedavi uygulandıktan sonra hastanın primer PKG yapılan bir merkeze nakli sağlanır. Burada da merkeze ulaşma süresine göre ve reperfüzyon kriterlerini karşılayıp karşılamamasına göre ya kurtarıcı PKG ya da rutin PKG yapılır. Fibrinolitik tedavi için kontrendikasyonlar açısından mutlaka hasta değerlendirilmeli ve ona göre karar verilmelidir. Bir çalışmaya göre, semptom başlangıcından itibaren ilk 6 saatte başvuran ST elevasyonlu AMI hastalarında her 1000 hastada 30 erken ölümü önlemiştir. Bu tedaviden en çok fayda sağlayan hasta grubu semptom başlangıcından itibaren ilk 2 saatte başvuran hastalardır [8]. Bu da fibrinolitik tedavide zamanın oldukça önemli olduğunu göstermektedir.

Zamanın bu derece önemli oluşu hastane öncesi dönemde fibrinolitik uygulamayı gündeme getirmiştir. Altı randomize klinik çalışmanın meta analizinin yapıldığı bir çalışmada; özellikle semptomlarının başlangıcının ilk iki saatinde, hastane öncesi fibrinolitik tedavi uygulanmasının hastane içi fibrinolitik uygulanmasına göre erken mortaliteyi %17 oranında azalttığı bildirilmiştir [9]. EKG'yi analiz edebilecek yetişmiş personelin olması halinde ST elevasyonlu AMI tespit edilen hastalarda fibrinolitik tedavinin ilk 10 dakikada başlatılması önerilmektedir. Semptomların başlangıcından itibaren başvuru zamanı uzadıkça (özellikle ilk üç saati geçtiyse) önceliğin hastayı primer PKG yapılacak merkeze transferi olması gerekmektedir.

Uygulanan fibrinolitik tedavinin başarısız olması durumunda (Fibrinolitik tedavi sonrası 60-90 dakikada ST segment elevasyondaki düzelmenin %50'den az oluşu gibi) kurtarıcı PKG yapılması gerekmektedir. Hemodinamik veya elektriksel instabilite, kalıcı göğüs ağrısı, iskeminin kötüleşmesi gibi durumlarda da kurtarıcı PKG yapılmalıdır. Başarılı fibrinolizis sonrası ise 2-24 saat içerisinde rutin erken PKG uygulanması lazımdır [10].

2017 ESC Kılavuzuna göre, fibrinolitik tedavide fibrin spesifik ajanların tercih edilmesi önerilir (Klas 1, A). Tek doz verilen ve kiloya göre ayarlanan Tenekteplaz (TNK-t-PA), hızlandırılmış TNK-t-PA ile kıyaslandığında 30 günlük mortalitede aynı etkilere sahip olmakla birlikte, beyin harici gelişen kanamaları ve transfüzyon gerekliliğini azaltması bakımından ondan daha üstündür. Hastane öncesi uygulamada da kullanımı daha kolaydır [11].

Kılavuzda önerilen ajanlar ve dozları tablo 1'de gösterilmiştir.

İlaç	Doz
Streptokinaz	1,5 milyon ünite 30-60 dakikada i.v
Alteplaz (t-PA)	15 mg. i.v bolus 0,75 mg/kg i.v (30 dak.dan fazla) (50 mg.a kadar) Sonra 0,5 mg/kg i.v (60 dak.dan fazla) (35 mg.a kadar)
Reteplaz (r-PA)	10 ünite + 10 ünite i.v bolus 30 dak. aralar ile
Tenekteplaz (TNK- t-PA)	Tek i.v bolus: 30 mg (6000IU) <60 kg. 35 mg (7000IU) 60- <70 kg 40 mg (8000IU) 70- <80 kg 45 mg (9000IU) 80- <90 kg 50 mg (10000IU) >90 kg 75 ve üzeri yaş grubunda yarı doz önerilir.

Tablo 1. *Akut Myokard İnfarktüsünde fibrinolitik tedavide kullanılan ilaçlar ve dozları*

Akut İskemik İnmede Trombolitik Tedavi

Akut iskemik inmede onaylanmış tek sistemik reperfüzyon tedavisi i.v trombolitik uygulanmasıdır. Diğer hastalıklarda olduğu gibi iskemik inmede de zaman tedaviye karar verme aşamasında oldukça önemlidir. Semptom başlangıcının ilk 4,5 saatinde başvuran ve herhangi bir kontrendikasyon bulunmayan tüm hastalarda trombolitik uygulanması önerilir.

Avrupa İnme Organizasyonu (ESO) 2021 yılı aralık ayında Aİİ'de trombolitik tedavi uygulamaları ile ilgili yeni bir kılavuz yayınladı [12]. Bu kılavuza göre trombolitik tedavi önerileri bu kısımda paylaşılacaktır.

İlk 4,5 saat içerisinde başvuran tüm Aİİ hastalarında kesin bir kontrendikasyon olmadığı müddetçe alteplaz ile trombolitik tedavi önerilmektedir (güçlü öneri, yüksek kanıt düzeyi). Bu bağlamda erken başvuran hastalar için bu öneri de herhangi bir değişiklik bulunmamaktadır. Bu hasta grubunda intrakraniyal hemoraji riski artmış olsa bile 90 günlük ölüm oranlarında bir artış yoktur. Kılavuz; laküner infarktlarda ya da görünür geniş arter tıkanıklıklarının olmadığı Aİİ tablolarında trombolitik uygulanmasına yönelik yeni bir öneri sunmamaktadır. ESO, bu kılavuzda özellikle 4,5 saat sonra başvuran hasta grubuna yönelik yeni öneriler getirmiştir. Öncesinde, başlangıcı bilinen 4,5-9 saatlik Aİİ hasta grubunda sadece Beyin BT ile değerlendirildiyse yüksek kanıt düzeyi ile trombolitik uygulanması önerilmemekteydi. Bu kılavuzda ise kanıt düzeyi orta olarak değiştirilmiştir.

Yine 4,5-9 saatlik sürede saptanan Aİİ hastalarında çekilen beyin BT (BT, perfüzyon BT) ile MRI (difüzyon veya perfüzyon ağırlıklı) arasında uyuşmama varsa bu penumbral serebral dokuyu tanımlayabileceğinden dolayı 4,5 saati geçen vakalarda faydalı olabilir. Bu grupta mekanik trombektomi planlanmıyorsa veya endike değilse alteplaz ile trombolitik tedavi önerilmektedir (düşük kanıt düzeyi, güçlü öneri).

Uykudan uyanmakla birlikte semptomları başlayan Aİİ hastalarında, eğer en son 4,5 saatten önce iyi halde görüldüyse ve MRI difüzyon- flair kesitlerinde uyumsuzluk varsa (mekanik trombektomi planlanmıyorsa veya endike değilse) trombolitik uygulanabilir (yüksek kanıt düzeyi, güçlü öneri).

Uykudan uyanmakla semptomları başlamış olan ve uykunun ortasından itibaren 9 saat içindeki Aİİ hastalarında BT veya MRI difüzyon/perfüzyon uyumsuzluğu varsa (mekanik trombektomi planlanmıyorsa veya endike değilse) trombolitik uygulanabilir (orta kanıt düzeyi, güçlü öneri).

Tüm bu yaklaşımlar 4,5 saat üzerinde başvuran ya da uykudan uyanma ile başladığı için tam başlangıç zamanı tahmin edilemeyen hasta grubunda görüntüleme yöntemlerinin öne çıkarılarak trombolitik tedavi başlanmasını amaçlamaktadır. Görüntüleme yöntemleri ile Aİİ'nin erken dönemde ise tanınması mümkün olabilmektedir. Erken tedavinin iyi sonuçlar ve az yan etkiyi beraberinde getirdiği düşünülürse bu alanda yapılan çalışmalar oldukça değerlidir.

Trombolitik ajanların seçimi konusunda kılavuz alteplazı öne çıkarmıştır. Dozajlama konusunda da bir değişiklik yapılmamıştır. Sadece erken dönemdeki Aİİ hastalarında eğer mekanik trombektomi planlanmış ve işlem öncesi trombolitik uygulaması yapılacaksa alteplaz yerine tenekteplaz önerilmektedir. Aİİ'de sistemik trombolitik tedavide ilk tercih edilen ajan alteplaz (t-PA) olup uygulama dozu ve şekli standarttır: 0,9 mg/kg (maksimum 90 mg) olacak şekilde uygulanır.

Toplam dozun %10'u bolus olarak 1 dakikada, kalan kısmı ise 60 dakikalık i.v infüzyon şeklinde verilir. Alternatif olarak tenekteplaz kullanılacaksa 0,25 mg/kg tek bolus halinde uygulanır. Tedavi öncesi kontrendikasyonlar dikkatle değerlendirilir ve uygulama sırasında kan basıncı, nörolojik durum ve olası hemoraji gelişimi açısından hasta yakından izlenir.

Sonuç

Yıllar içerisinde yararlılığını kanıtlayan ve kullanım alanı giderek genişleyen trombolitik tedaviler, olası yan etkileri, kontrendike olduğu durumların iyi ayırt edilmesi ve erken başlangıç ile daha iyi sonuçlar vermesi ile özellikle Acil Servis'te dikkatli hasta seçimini beraberinde getirmektedir. Aynı zamanda uygun ajanın seçilmesi de önemlidir. Trombolitik tedaviler ile ilgili halihazırda devam eden geniş çaplı araştırmalar ile şekillenen kılavuzları takip etmek trombolitik tedavi gereken hastaların yönetimi açısından faydalı olacaktır.

Teşekkür ve Finansman
Bu araştırma, kamu, ticari veya kâr amacı gütmeyen sektörlerdeki fon sağlayan kuruluşlardan herhangi bir özel hibe almamıştır.

Çıkar Çatışması
Yazar herhangi bir çıkar çatışması olmadığını beyan eder.

Yazar Katkıları
Planlama, yorum yazma ve düzeltmeler: NB

Etik ve Katılım Onayı
Bu çalışma etik standartlara uygun olarak yürütülmüştür.

Referanslar

1. Marder VJ. Thrombolytic therapy: 2001. Blood Rev. 2001 Sep;15(3):143-57. doi: 10.1054/blre.2001.0161.

2. Kluft C, Sidelmann JJ, Gram JB. Assessing Safety of Thrombolytic Therapy. Semin Thromb Hemost. 2017 Apr;43(3):300-310. doi: 10.1055/s-0036-1584130.

3. Nicolaides A, Hull RD, Fareed J. Thrombolytic therapy. Cardiovascular Disease Educational and Research Trust; European Venous Forum, North American Thrombosis Forum; International Union of Angiology and Union Internationale du Phlebologie.Clin Appl Thromb Hemost. 2013 Mar-Apr;19(2):198-204. doi: 10.1177/1076029612474840o.

4. Zuo Z, Yue J, Dong BR, Wu T, Liu GJ, Hao Q. Thrombolytic therapy for pulmonary embolism. Cochrane Database Syst Rev. 2021 Apr 15;4(4):CD004437. doi: 10.1002/14651858.CD004437.pub6.

5. Stevens SM, Woller SC, Baumann Kreuziger L, Bounameaux H, Doerschug K, Geersing GJ, Huisman MV, Kearon C, King CS, Knighton AJ, Lake E, Murin S, Vintch JRE, Wells PS, Moores LK. Executive Summary: Antithrombotic Therapy for VTE Disease: Second Update of the CHEST Guideline and Expert Panel Report. Chest. 2021 Dec;160(6):2247-2259. doi: 10.1016/j.chest.2021.07.056.

6. No authors listed. Venous thromboembolic diseases: diagnosis, management and thrombophilia testing.
London: National Institute for Health and Care Excellence (NICE); 2020 Mar 26.PMID: 32374563 Free Books & Documents. Review.

7. Konstantinides SV, Meyer G, Becattini C, Bueno H, Geersing GJ, Harjola VP, Huisman MV, Humbert M, Jennings CS, Jiménez D, Kucher N, Lang IM, Lankeit M, Lorusso R, Mazzolai L, Meneveau N, Ní Áinle F, Prandoni P, Pruszczyk P, Righini M, Torbicki A, Van Belle E, Zamorano JL; ESC Scientific Document Group. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). Eur Heart J. 2020 Jan 21;41(4):543-603. doi: 10.1093/eurheartj/ehz405.

8. Ibanez B, James S, Agewall S, Antunes MJ, Bucciarelli-Ducci C, Bueno H, Caforio ALP, Crea F, Goudevenos JA, Halvorsen S, Hindricks G, Kastrati A, Lenzen MJ, Prescott E, Roffi M, Valgimigli M, Varenhorst C, Vranckx P, Widimský P; ESC Scientific Document Group. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). .Eur Heart J. 2018 Jan 7;39(2):119-177. doi: 10.1093/eurheartj/ehx393.

9. Morrison LJ, Verbeck PR, McDonald AC, Sawadsky BV, Cook DJ. Mortality and prehospital thrombolysis for acute myocardial infarction: a meta-analysis. JAMA 2000;283(20):2686–2692.

10. Armstrong PW, Gershlick AH, Goldstein P, Wilcox R, Danays T, Lambert Y, Sulimov V, Rosell Ortiz F, Ostojic M, Welsh RC, Carvalho AC, Nanas J, Arntz HR, Halvorsen S, Huber K, Grajek S, Fresco C, Bluhmki E, Regelin A, Vandenberghe K, Bogaerts K, Van de Werf F, STREAM Investigative Team. Fibrinolysis or primary PCI in ST-segment elevation myocardial infarction. N Engl J Med 2013;368(15):1379–1387

11. Task Force on the management of ST-segment elevationsacute myocardial infarction of the European Society of Cardiology (ESC), Steg PG, James SK, Atar D, Badano LP, Blomstrom-Lundqvist C, Borger MA, Di Mario C, Dickstein K, Ducrocq G, Fernandez-Aviles F, Gershlick AH, Giannuzzi P, Halvorsen S, Huber K, Juni P, Kastrati A, Knuuti J, Lenzen MJ, Mahaffey KW, Valgimigli M, van 't Hof A, Widimsky P, Zahger D. ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. Eur Heart J 2012;33(20):2569–2619.

12. Berge E, Whiteley W, Audebert H, De Marchis GM, Fonseca AC, Padiglioni C, de la Ossa NP, Strbian D, Tsivgoulis G, Turc G. European Stroke Organisation (ESO) guidelines on intravenous thrombolysis for acute ischaemic stroke. Eur Stroke J. 2021 Mar;6(1):I-LXII. doi: 10.1177/2396987321989865. Epub 2021 Feb 19.

ÖZGÜN MAKALE | ORIGINAL ARTICLE

Quality and Clinical Safety of ChatGPT-Generated Discharge Summaries in Orthopaedic Surgery: A Cross-Sectional Evaluation

Mustafa Yalin[1], Gozde Kirgin[1], Ferid Samedov[1], Mustafa Bakir[1], Abbas Tokyay[1]

[1] *Yalova University, Faculty of Medicine, Department of Orthopedics and Traumatology, Yalova, Turkey.*

Abstract

Aim: To evaluate the clarity, completeness, and clinical safety of discharge summaries written by ChatGPT for ten common orthopaedic surgical procedures. Materials and Methods: Ten representative orthopaedic procedures were selected: total hip arthroplasty (THA), total knee arthroplasty (TKA), meniscus repair, distal radius fracture fixation, proximal femoral nailing, Achilles tendon repair, rotator cuff repair, lumbar discectomy, carpal tunnel release, and anterior cruciate ligament (ACL) reconstruction. ChatGPT (GPT-4, March 2025 version) was prompted to generate standardized patient discharge summaries for each procedure. Three orthopaedic surgeons independently rated each summary on three domains—clarity, completeness, and clinical safety—using 5-point Likert scales. Readability was assessed using the Flesch–Kincaid Grade Level (FKGL) and Reading Ease (FKRE). Inter-rater reliability was calculated via intraclass correlation coefficient (ICC).

Results: The median total quality score was 13.4 (IQR 12.9–14.0) out of 15, with high ratings for clarity (4.5) and safety (4.6). Inter-rater reliability was excellent (ICC = 0.92). Arthroplasty-related summaries (THA, TKA) received the highest scores, while soft-tissue procedures scored lowest. Mean FKGL was 9.1 ± 0.4, and mean FKRE was 56.6 ± 2.1, indicating moderately difficult but comprehensible language. ICC for total quality was 0.92 (95% CI 0.85–0.97). Minor omissions (e.g., missing DVT prophylaxis details) were noted in two summaries; however, none posed safety risks.

Conclusion: ChatGPT can produce discharge summaries that are clear, accurate, and clinically safe across various orthopaedic procedures. While the generated notes approximate expert-level quality, professional oversight remains essential to ensure completeness and individualized care recommendations.

Keywords: ChatGPT, Artificial Intelligence, Orthopaedic Surgery, Discharge Summary, Patient Safety, Readability

Introduction

Effective discharge documentation is a cornerstone of postoperative patient safety, ensuring continuity of care and reducing the risk of complications or hospital readmission [1]. Discharge summaries provide patients with critical information on wound care, medications, physical activity, and warning signs, while also serving as a communication bridge between surgeons, rehabilitation teams, and primary care providers [2]. However, the preparation of clear, comprehensive, and patient-friendly discharge notes is time-consuming and prone to variability among clinicians.

Recent advances in artificial intelligence (AI), particularly the emergence of large language models (LLMs) such as ChatGPT (OpenAI), have created new opportunities to assist clinicians with medical documentation [3,4]. ChatGPT is capable of generating coherent and grammatically accurate text in response to natural-language prompts, and has demonstrated potential in drafting clinical correspondence, academic abstracts, and patient education materials [5,6]. Early studies have explored its performance in summarizing radiology reports, writing operative notes, and answering medical questions [7,8], yet its role in generating structured discharge summaries—especially within orthopaedics—remains under-investigated.

Corresponding Author: Assistant Professor, Mustafa Yalin, M.D.

Address: *Yalova University Faculty of Medicine, Department of Orthopedics and Traumatology, 77200 Yalova*
E-mail: *mustafa.yalin@yalova.edu.tr*
Phone: *+905546977423*

Accurate and well-structured discharge summaries play a crucial role in preventing postoperative complications. Clear, timely, and patient-oriented discharge information improves medication adherence, reduces miscommunication during transitions of care, and enables early recognition of complications. Orthopaedic discharge notes must balance medical accuracy with simplicity, addressing postoperative restrictions, rehabilitation instructions, and warning signs in clear language that patients can easily understand. An AI-assisted approach could reduce clinicians' workload and improve documentation consistency, but may also introduce risks if information is incomplete or inaccurate. Therefore, before integration into practice, the quality and safety of AI-generated discharge summaries should be systematically evaluated. The present study aimed to assess the clarity, completeness, and clinical safety of discharge summaries generated by ChatGPT for ten of the most common orthopaedic surgical procedures. Secondary objectives included evaluating readability levels, identifying procedural differences in performance, and examining inter-rater reliability among orthopaedic reviewers. By providing structured evidence on ChatGPT's capacity to produce clinically sound postoperative documentation, this study seeks to inform the potential role of LLMs in surgical communication and patient education.

Materials and Methods

This cross-sectional study evaluated the quality of discharge summaries generated by ChatGPT for ten common orthopaedic surgical procedures. Since no patient data were used and all texts were synthetically produced by an artificial intelligence model, ethics committee approval was not required. The study was conducted between March and April 2025 using the GPT-4 model (OpenAI, San Francisco, CA).

The ten procedures were selected to represent a broad range of orthopaedic practice areas, including arthroplasty, trauma, sports, upper limb, and spine surgery. The selected procedures were: total hip arthroplasty (THA), total knee arthroplasty (TKA), arthroscopic meniscus repair, distal radius fracture fixation, proximal femoral nailing, Achilles tendon repair, arthroscopic rotator cuff repair, lumbar disc herniation surgery, carpal tunnel release, and anterior cruciate ligament (ACL) reconstruction. Each procedure was described to ChatGPT through a standardized prompt designed to simulate a realistic clinical request: "Prepare a detailed patient discharge summary for a postoperative case of [procedure], including diagnosis, operation performed, medications, wound care instructions, physical activity or rehabilitation plan, warning signs, and follow-up recommendations."

All prompts were entered within a single ChatGPT session to avoid version drift or memory variation. The model outputs were copied verbatim without manual editing or correction. Each discharge summary was written in English, with an average length of approximately 100-150 words (Supplementary index). All generated discharge summaries are provided in Supplementary File.

Three orthopaedic surgeons, each with at least five years of independent clinical experience, independently reviewed and scored every discharge summary. Evaluation was based on three domains: clarity, completeness, and clinical safety. Clarity referred to readability, organization, and overall linguistic coherence. Completeness was defined as inclusion of the essential elements expected in a standard discharge note—diagnosis, operation details, wound care, activity restrictions, medication, and follow-up. Clinical safety focused on the presence or absence of misleading, incorrect, or potentially harmful statements. Each domain was rated on a 5-point Likert scale (1 = very poor, 5 = excellent), and the total quality score for each summary was calculated as the sum of three domains (maximum 15 points).

In addition to expert scoring, textual readability was assessed using the Flesch-Kincaid Grade Level (FKGL) and Flesch Reading Ease (FKRE) indices, computed through an automated readability tool (Hemingway Editor, version 3.0). FKGL estimates the U.S. educational grade required to understand the text, while FKRE represents overall ease of reading, where higher scores indicate simpler language.

Statistical analysis was performed using IBM SPSS Statistics version 25 (IBM Corp., Armonk, NY, USA) and Python (pandas and scipy libraries). Continuous variables were summarized as medians and interquartile ranges (IQR) due to the small sample size. Group differences among the ten procedures were analyzed using the Kruskal-Wallis test, and pairwise post-hoc comparisons were adjusted with the Dunn correction. Correlations between readability and expert scores were examined with Spearman's rank correlation coefficient (ρ). The inter-rater reliability among the three orthopaedic reviewers was determined by the two-way random effects intraclass correlation coefficient (ICC [3,1]), interpreted according to Koo and Li's classification (excellent ≥ 0.90). Statistical significance was set at $p < 0.05$.

The primary outcome of the study was the total quality score of each ChatGPT-generated discharge summary. Secondary outcomes included domain-specific scores (clarity, completeness, safety), readability indices (FKGL, FKRE), and inter-rater agreement.

Results

Ten ChatGPT-generated discharge summaries corresponding to common orthopaedic surgical procedures were evaluated by three independent orthopaedic surgeons. All generated texts were successfully analyzed and included in the study. The procedures assessed were THA, TKA, arthroscopic meniscus repair, distal radius fracture fixation, proximal femoral nailing, Achilles tendon repair, rotator cuff repair, lumbar disc herniation surgery, carpal tunnel release, and ACL reconstruction.

Overall performance

The median total quality score across all summaries was 13.4 (IQR 12.9–14.0) out of a maximum of 15, reflecting overall high performance. Among subdomains, clarity achieved the highest median value of 4.5 (IQR 4.3–4.6), followed by clinical safety at 4.6 (IQR 4.4–4.8) and completeness at 4.3 (IQR 4.1–4.6). The average FKGL was 9.1 ± 0.4, corresponding approximately to a high school reading level, while the FKRE averaged 56.6 ± 2.1, indicating moderately difficult but generally comprehensible language for the average educated patient.
A detailed distribution of expert ratings and readability indices is presented in Table 1.

Procedure-specific analysis

The highest total quality scores were observed for arthroplasty-related summaries: THA (14.4) and TKA (14.0). These notes were consistently rated as the most clear, complete, and clinically safe. Summaries for spine (lumbar disc surgery, 13.8) and fracture fixation (13.2) procedures also performed well. The lowest overall score was recorded for rotator cuff repair (12.7), mainly due to minor omissions in postoperative rehabilitation details.

Minor discrepancies between raters were identified but did not alter the interpretation of results. No summary contained unsafe or misleading recommendations. In two summaries (ACL reconstruction and proximal femoral nailing), experts noted that the duration of thromboprophylaxis was not explicitly mentioned, representing a minor but clinically relevant omission. No factual inaccuracies or potentially harmful content were detected in any document.

Statistical analysis

A Kruskal–Wallis test comparing total scores among the ten procedures revealed a statistically significant difference ($p = 0.03$). Pairwise post-hoc analysis showed that THA and TKA summaries scored significantly higher than those for rotator cuff repair and carpal tunnel release (adjusted $p<0.05$).
The correlation between readability and clarity was moderately negative (Spearman $\rho = -0.54$, $p = 0.04$), indicating that summaries written in simpler language tended to receive higher clarity ratings from experts. No significant correlation was found between readability and clinical safety ($\rho = -0.22$, $p = 0.47$).

The inter-rater reliability for the total quality score was excellent, with an intraclass correlation coefficient (ICC [3,1]) of 0.92 (95% CI 0.85–0.97). Domain-specific ICCs were also high: 0.88 for clarity, 0.91 for completeness, and 0.93 for safety.

Readability outcomes

Readability scores demonstrated minor variability across procedure types. The most accessible summaries were those for carpal tunnel release (FKGL 8.5; FKRE 59.3) and meniscus repair (FKGL 8.9; FKRE 58.1), while the most linguistically complex were for THA (FKGL 9.8) and lumbar disc surgery (FKGL 9.7). Despite slight variation, all summaries remained within the range considered understandable for general health information (FKRE 50–60).

Table 1. Quality and readability scores of ChatGPT-generated discharge summaries for ten orthopaedic procedures

Procedure	Clarity (Mean ± SD)	Completeness (Mean ± SD)	Clinical Safety (Mean ± SD)	Total Score (0–15)	FKGL	FKRE
Total Hip Arthroplasty (THA)	4.8 ± 0.2	4.7 ± 0.3	4.9 ± 0.1	14.4	9.8	54.0
Total Knee Arthroplasty (TKA)	4.6 ± 0.3	4.6 ± 0.4	4.8 ± 0.2	14.0	9.5	55.2
Meniscus Repair	4.4 ± 0.4	4.3 ± 0.5	4.5 ± 0.3	13.2	8.9	58.1
Distal Radius Fixation	4.3 ± 0.5	4.1 ± 0.6	4.4 ± 0.4	12.8	8.6	59.0
Proximal Femoral Nailing	4.4 ± 0.4	4.2 ± 0.5	4.6 ± 0.3	13.2	9.0	56.4
Achilles Tendon Repair	4.5 ± 0.3	4.3 ± 0.4	4.7 ± 0.3	13.5	8.8	57.6
Rotator Cuff Repair	4.3 ± 0.4	4.0 ± 0.6	4.4 ± 0.4	12.7	9.1	56.9
Lumbar Discectomy	4.6 ± 0.3	4.5 ± 0.4	4.7 ± 0.2	13.8	9.7	54.5
Carpal Tunnel Release	4.4 ± 0.3	4.1 ± 0.5	4.6 ± 0.3	13.1	8.5	59.3
ACL Reconstruction	4.5 ± 0.3	4.2 ± 0.4	4.6 ± 0.3	13.3	9.3	55.8
Median (IQR)	**4.5 (4.3–4.6)**	**4.3 (4.1–4.6)**	**4.6 (4.4–4.8)**	**13.4 (12.9–14.0)**	**9.1 (8.7–9.6)**	**56.6 (55–58)**

FKGL, Flesch–Kincaid Grade Level; FKRE, Flesch Reading Ease.

Table 2. Comparative quality scores by procedure category

Category	Included Procedures	n	Median Total Score (IQR)	Median Clarity	Median Completeness	Median Safety	p value (Kruskal–Wallis)
Arthroplasty	THA, TKA	2	14.2 (13.8–14.4)	4.7	4.6	4.8	—
Trauma	Radius Fx, Femoral Nail	2	13.0 (12.8–13.2)	4.3	4.1	4.5	—
Soft-tissue / Sports	Meniscus, Achilles, Rotator Cuff, ACL	4	13.1 (12.7–13.3)	4.4	4.2	4.6	—
Spine / Neuro	Lumbar Discectomy	1	13.8	4.6	4.5	4.7	—
Hand / Nerve	Carpal Tunnel	1	13.1	4.4	4.1	4.6	—
Overall	—	10	13.4 (12.9–14.0)	4.5	4.3	4.6	0.03*

Kruskal–Wallis test among procedures, p = 0.03

Table 3. Correlation and reliability analysis

Parameter Pair	Correlation Coefficient (ρ)	p Value	Interpretation
FKGL ↔ Clarity	−0.54	0.04*	Simpler language correlated with higher clarity
FKGL ↔ Completeness	−0.31	0.20	Not significant
FKGL ↔ Safety	−0.22	0.47	Not significant
FKRE ↔ Clarity	+0.49	0.05	Marginally significant
FKRE ↔ Total Score	+0.41	0.09	Weak positive correlation
Clarity ↔ Safety	+0.72	0.01*	Strong positive relationship
Completeness ↔ Safety	+0.69	0.01*	Good internal consistency
Inter-rater ICC (total)	**0.92 (95% CI 0.85–0.97)**	—	Excellent agreement

Abbreviations: *FKGL = Flesch–Kincaid Grade Level, FKRE = Flesch Reading Ease.*
Spearman correlation; ICC(3,1); significance p < 0.05.

Discussion

This study provides one of the first structured evaluations of ChatGPT-generated discharge summaries within the field of orthopaedic surgery. Our findings demonstrate that ChatGPT can produce clear, clinically safe, and largely complete discharge documentation across a wide range of surgical procedures, with particularly high performance in arthroplasty-related cases. The overall quality scores approached expert-level benchmarks, and inter-rater reliability was excellent (ICC = 0.92), underscoring the model's consistency and linguistic coherence. Importantly, no unsafe or misleading instructions were identified, and readability metrics (mean FKGL = 9.1; FKRE = 56.6) indicated that the language was accessible to most educated patients. By systematically quantifying both linguistic and clinical aspects of AI-generated discharge notes, this study fills a critical evidence gap in the emerging interface between LLMs and postoperative communication.

The median total quality score across all summaries was 13.4/15 (IQR 12.9–14.0), reflecting overall high performance in our study. Among subdomains, clinical safety achieved the highest median value of 4.6 (IQR 4.4–4.8), followed by clarity at 4.5 (IQR 4.3–4.6) and completeness at 4.3 (IQR 4.1–4.6). The average FKGL was 9.1 ± 0.4, corresponding approximately to a high-school reading level, while the FKRE averaged 56.6 ± 2.1, indicating moderately difficult but generally comprehensible language for the average educated patient. These results are consistent with previous studies demonstrating that LLMs and AI-assisted documentation tools can improve readability and documentation quality [5] and the transformation of discharge summaries into more patient-friendly formats [9].

However, our completeness score is somewhat higher than in several reports where omissions or safety concerns were more pronounced—for example, the study of discharge summary transformation found safety concerns in 18% of reviews and only 56% of summaries rated "entirely complete" [9]. Moreover, open-source LLM-based generation of discharge summaries reported lower alignment metrics and higher error rates compared to human-written notes, which suggests that context (specialty, prompt design) matters substantially [10]. Similarly, a systematic review of AI in clinical documentation highlighted that while AI tools show promise, issues of accuracy, completeness, and documentation burden remain [3].

Additionally, in a comparison of junior residents vs LLMs in medical record documentation, Huang et al. [11] found trade-offs between readability, completeness, and accuracy. In contrast, a prompt-engineering study in cardiology discharge summaries showed that LLM-generated summaries reached significant readability improvements and high expert ratings for correctness and completeness when carefully engineered [12]. Taken together, our high clarity and safety scores suggest that—with careful prompt engineering and postoperative-care context—the use of ChatGPT-generated summaries in orthopaedic surgery can match or exceed prior benchmarks. Yet the persistence of completeness and omission risks reinforces the need for clinician review before deployment.

The higher performance observed in arthroplasty-related summaries (THA 14.4; TKA 14.0) likely reflects the advantage of standardized postoperative pathways and well-defined care protocols in facilitating consistent AI output [13].

This mirrors findings by LLM Assistant for Emergency Department Discharge Documentation, where an LLM assisted ED discharge notes and achieved higher completeness and correctness scores in a fast-paced environment [14]. Conversely, the lower score for rotator cuff repair (12.7), due to omissions in rehabilitation guidance, aligns with the results of a pilot feasibility study comparing LLMs in extracting key information from ICU patient text records, where LLMs achieved only ~41% accuracy in capturing key clinical events in ICU summarisation tasks [15]. The absence of unsafe or misleading recommendations in our generated summaries is echoed by the cross-sectional study, which found LLM- and physician-generated narratives comparable in potential for harm despite slightly higher error frequency in the LLM group [16]. On the flip side, our finding of near-expert clarity and safety contrasts somewhat with the German study by Automated generation of discharge summaries: leveraging LLMs with clinical data, where only 60% of the summaries were rated "good" for comprehensiveness and numerous omissions were documented (ROUGE-1 only 0.25) when open-source LLMs were used [10]. The divergence may be attributable to differences in input data (structured vs narrative), language context, and prompt engineering. Further, the improvement in readability via LLMs seen in evaluation of a large language model to simplify discharge summaries and provide cardiological lifestyle recommendations (demonstrating significant readability gains in cardiology discharge summaries) supports our FKGL ≈ 9 and FKRE ≈ 56 findings [12]. Overall, these comparisons suggest that AI-generated summaries perform best in surgical domains with clear post-operative regimens, but procedural nuances (such as thromboprophylaxis duration and rehab instructions) remain vulnerable to omission. This reinforces the need for clinician review and specialty-specific prompt tuning before broad deployment.

The finding that discharge summaries for THA and TKA achieved significantly higher scores than those for rotator cuff repair and carpal tunnel release highlights the impact of procedural standardization on the quality of AI-generated documentation. In major joint replacement surgeries such as THA and TKA, perioperative protocols—including pain management, thromboprophylaxis, and structured rehabilitation—are well established and widely integrated into clinical practice and literature.

For instance, recent research has demonstrated that implementing comprehensive clinical pathways encompassing preoperative preparation, intraoperative care, and postoperative rehabilitation significantly improves documentation quality in arthroplasty cases [17].

Such structured care models provide an advantageous framework for large LLMs to generate accurate, contextually appropriate discharge summaries. Indeed, studies have shown that properly fine-tuned or domain-adapted LLMs can perform as well as, or even outperform, human experts in clinical summarization tasks [18]. In contrast, upper-extremity soft-tissue surgeries such as rotator cuff repair and carpal tunnel release often involve greater variability in rehabilitation duration, immobilization strategies, and patient-specific factors. This heterogeneity can compromise the model's ability to produce complete and coherent outputs [19,20]. Consistent with this, literature in clinical natural-language generation has emphasized that unstructured or variable input data negatively affect model performance [21,22]. Therefore, the superior quality observed in THA and TKA summaries likely stems from the protocol-driven, clearly defined nature of arthroplasty care, which facilitates safer and more comprehensive AI-generated outputs [23,24]. Our findings suggest that LLM-based summarization still faces challenges in procedures with heterogeneous care pathways and that its applicability may vary depending on the degree of clinical standardization inherent to each surgical domain.

Clinical implications

LLMs such as ChatGPT can assist in generating accurate and coherent discharge summaries, particularly for standardized procedures like THA and TKA. Structured clinical pathways and well-defined postoperative protocols enhance the reliability of AI-generated medical documentation. Human supervision remains essential for variable or individualized procedures (e.g., soft-tissue or upper-extremity surgeries) to ensure clinical safety and contextual accuracy.

Limitations and strengths

This study has several limitations that warrant consideration. First, the discharge summaries were generated using a single AI model (ChatGPT-4), and the findings may not generalize to other LLMs or future model versions with different training datasets or prompt structures. Second, the evaluation process, although based on validated scoring instruments (DISCERN questionnaire, Journal of the American Medical Association, the Global Quality Scale, FKGL, FKRE), relied on expert assessment and may still contain subjective elements inherent to qualitative rating. Third, the study covered a limited set of orthopaedic procedures and did not include emergency or pediatric conditions, which could influence the generalizability of the results across broader surgical contexts. Additionally, AI-generated content was not tested in real-time clinical environments, where integration with electronic health records and user feedback might further impact output quality.

Despite these limitations, the study has notable strengths. It is among the first to systematically evaluate the documentation quality of ChatGPT across multiple orthopaedic procedures using standardized assessment metrics. The inclusion of both arthroplasty and upper-extremity soft-tissue procedures provided insight into how procedural standardization affects AI performance. Moreover, the multidisciplinary evaluation design—combining readability, reliability, and medical accuracy metrics—offers a comprehensive framework for assessing AI-generated clinical communication. Collectively, these strengths enhance the validity and practical relevance of our findings for future applications of AI in orthopaedic documentation.

Conclusions

ChatGPT demonstrated a strong capacity to generate coherent and reliable discharge summaries across a range of orthopaedic procedures. The model performed best in arthroplasty-related summaries, likely reflecting the benefits of standardized postoperative care pathways. While these findings support the potential use of LLMs in clinical documentation, human oversight remains essential, particularly for procedures with variable rehabilitation protocols. Integrating AI-assisted documentation into orthopaedic workflows may enhance efficiency and consistency, provided its outputs are validated within real clinical settings. AI-generated discharge templates could reduce clinician workload while maintaining patient safety and communication quality.

Acknowledgment and Funding Statement: No funding was received for conducting this study

Conflict of Interest: We declare that all authors are free of financial or any other conflicts of interest in this article. As this study evaluated publicly available AI-generated text and did not involve human subjects or patient data, ethical approval was not applicable.

Author contributions: Conceptualization: [MY, GK, FS, MB, AT]; Methodology: [MY, GK, FS, MB, AT]; Formal analysis and investigation: [MY, GK, FS, MB, AT]; Writing- original draft preparation: [MY, GK, FS, MB, AT]; Writing- review and editing: [MY, GK, FS, MB, AT]; Resources: [MY, GK, FS, MB, AT]; Supervision: [MY].

References

1. Gonçalves-Bradley DC, Lannin NA, Clemson L, Cameron ID, Shepperd S. Discharge planning from hospital. Cochrane Database Syst Rev. 2022;2(2):CD000313. doi:10.1002/14651858.CD000313.pub6

2. Cam H, Wennlöf B, Gillespie U, et al. The complexities of communication at hospital discharge of older patients: a qualitative study of healthcare professionals' views. BMC Health Serv Res. 2023;23(1):1211. doi:10.1186/s12913-023-10192-5

3. Perkins SW, Muste JC, Alam T, Singh RP. Improving Clinical Documentation with Artificial Intelligence: A Systematic Review. Perspect Health Inf Manag. 2024;21(2):1d.

4. Meng X, Yan X, Zhang K, et al. The application of large language models in medicine: A scoping review. iScience. 2024;27(5):109713. doi: 10.1016/j.isci.2024.109713.

5. Baker HP, Dwyer E, Kalidoss S, Hynes K, Wolf J, Strelzow JA. ChatGPT's Ability to Assist with Clinical Documentation: A Randomized Controlled Trial. J Am Acad Orthop Surg. 2024;32(3):123-9. doi: 10.5435/JAAOS-D-23-00474.

6. Rodler S, Cei F, Ganjavi C, et al. GPT-4 generates accurate and readable patient education materials aligned with current oncological guidelines: A randomized assessment. PLoS One. 2025;20(6):e0324175. doi: 10.1371/journal.pone.0324175.

7. Sterling NW, Brann F, Frisch SO, Schrager JD. Patient-Readable Radiology Report Summaries Generated via Large Language Model: Safety and Quality. J of Patient Exp. 2024;11. doi: 10.1177/23743735241259477.

8. Wei Q, Yao Z, Cui Y, Wei B, Jin Z, Xu X. Evaluation of ChatGPT-generated medical responses: A systematic review and meta-analysis. J Biomed Inform. 2024;151:104620. doi: 10.1016/j.jbi.2024.104620.

9. Zaretsky J, Kim JM, Baskharoun S, et al. Generative Artificial Intelligence to Transform Inpatient Discharge Summaries to Patient-Friendly Language and Format. JAMA Netw Open. 2024;7(3):e240357. doi: 10.1001/jamanetworkopen.2024.0357.

10. Ganzinger M, Kunz N, Fuchs P, et al. Automated generation of discharge summaries: leveraging large language models with clinical data. Sci Rep. 2025;15(1):16466. doi: 10.1038/s41598-025-01618-7.

11. Huang TY, Hsieh PH, Chang YC. Performance Comparison of Junior Residents and ChatGPT in the Objective Structured Clinical Examination (OSCE) for Medical History Taking and Documentation of Medical Records: Development and Usability Study. JMIR Med Educ. 2024;10:e59902. doi: 10.2196/59902.

12. Rust P, Frings J, Meister S, Fehring L. Evaluation of a large language model to simplify discharge summaries and provide cardiological lifestyle recommendations. Commun Med (Lond). 2025;5(1):208. doi: 10.1038/s43856-025-00927-2.

13. Featherall J, Brigati DP, Faour M, Messner W, Higuera CA. Implementation of a Total Hip Arthroplasty Care Pathway at a High-Volume Health System: Effect on Length of Stay, Discharge Disposition, and 90-Day Complications. J Arthroplasty. 2018;33(6):1675-80. doi: 10.1016/j.arth.2018.01.038.

14. Song JW, Park J, Kim JH, You SC. Large Language Model Assistant for Emergency Department Discharge Documentation. JAMA Netw Open. 2025;8(10):e2538427. doi: 10.1001/jamanetworkopen.2025.38427.

15. Urquhart E, Ryan J, Hartigan S, et al. A pilot feasibility study comparing large language models in extracting key information from ICU patient text records from an Irish population. Intensive Care Med Exp. 2024;12(1):71. doi: 10.1186/s40635-024-00656-1.

16. Williams CYK, Subramanian CR, Ali SS, et al. Physician- and Large Language Model-Generated Hospital Discharge Summaries. JAMA Intern Med. 2025;185(7):818-25. doi: 10.1001/jamainternmed.2025.0821.

17. Godinho M, Maçães F, Gonçalves H, Silva F. Clinical Pathways in Knee and Hip Arthroplasty: Narrative Review on Sustainability, Quality, and Resource Management. JMIR Perioper Med. 2025;8:e78174. doi: 10.2196/78174.

18. Van Veen D, Van Uden C, Blankemeier L, et al. Adapted large language models can outperform medical experts in clinical text summarization. Nat Med. 2024;30(4):1134-42. doi: 10.1038/s41591-024-02855-5.

19. Mazuquin B, Moffatt M, Gill P, et al. Effectiveness of early versus delayed rehabilitation following rotator cuff repair: Systematic review and meta-analyses. PLoS One. 2021;16(5):e0252137. doi: 10.1371/journal.pone.0252137.

20. Galetta MD, Keller RE, Sabbag OD, et al. Rehabilitation variability after rotator cuff repair. J Shoulder Elbow Surg. 2021;30(6):322-33. doi: 10.1016/j.jse.2020.11.016.

21. Wu H, Wang M, Wu J, et al. A survey on clinical natural language processing in the United Kingdom from 2007 to 2022. NPJ Digit Med. 2022;5(1):186. doi: 10.1038/s41746-022-00730-6.

22. Zhang X, Wang Y, Jiang Y, Pacella CB, Zhang W. Integrating structured and unstructured data for predicting emergency severity: an association and predictive study using transformer-based natural language processing models. BMC Med Inform Decis Mak. 2024;24(1):372. doi: 10.1186/s12911-024-02793-9.

23. Bonner HV, Jones JR, Arguello AM, et al. Quality Improvement Tools in Total Joint Arthroplasty: A Systematic Review. J Surg Orthop Adv. 2021;30(3):125-30.

24. Minick KI, Hunter SJ, Capin JJ, et al. Improved Outcomes Following a Care Guideline Implementation: Part 1 of an Analysis of 12 355 Patients After Total Knee Arthroplasty. J Orthop Sports Phys Ther. 2023;53(3):143-50. doi: 10.2519/jospt.2022.11369.

VAKA TAKDİMİ | CASE REPORT

Dravet Syndrome with Neurological Progression and Systemic Findings: A Case Report

Ozge COSKUN SAGLAM [1,2] , Sevim YENER [3] , Nurullah YUCEL [4] , Mete BUYUKERTAN [4]

[1] Istanbul Bilgi University, Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, Istanbul, Turkey.
[2] University of Health Sciences, Institute of Health Sciences, Department of Anatomy, Istanbul, Turkey.
[3] University of Health Sciences, Umraniye Training and Research Hospital, Pediatric Urology Clinic, Istanbul, Turkey.
[4] University of Health Sciences, Hamidiye International School of Medicine, Department of Anatomy, Istanbul, Turkey.

Abstract

Objective: This case report aims to evaluate the clinical course, EEG, imaging findings, developmental status, and treatment of a patient with Dravet Syndrome associated with an SCN1A mutation from a multidisciplinary perspective.

Case Report: A 36-week premature female infant experienced a febrile hemiclonic seizure, followed by refractory epileptic attacks including status epilepticus, tonic-clonic, absence, and myoclonic seizures. Initially normal or mild, the EEG findings worsened over time, evolving from Lennox-Gastaut-like patterns to modified hypsarrhythmias reflecting epileptic encephalopathy. Genetic analysis confirmed an asymptomatic (heterozygous) SCN1A mutation in the father. Developmental assessments revealed a mild initial delay but later progressed to a marked, widespread impairment in all domains. Despite multiple combinations of antiepileptic drugs and a ketogenic diet, complete seizure control was not achieved.

Conclusion: This case highlights the multifaceted clinical deterioration of Dravet syndrome, characterized by treatment-resistant seizures, marked neurophysiological progression, and systemic organ involvement. The findings highlight the critical importance of early genetic diagnosis and a comprehensive, multidisciplinary approach in the management of this complex and progressive syndrome to optimize long-term outcomes.

Keywords: Case Report, Dravet Syndrome, Mutation, Pediatric Epileps, SCN1A

Introduction

Dravet Syndrome was first described in 1978 by Charlotte Dravet as "severe myoclonic epilepsy of infancy (SMEI)" [1]. However, over time, it was realized that this disease includes different seizure types such as febrile seizures, focal seizures, tonic-clonic seizures, and absence seizures, and also exhibits movement disorders and developmental and cognitive delays, and its name was changed to "Dravet Syndrome" in 1989. In 2001, the syndrome was shown to be associated with the SCN1A gene mutation, and the International League Against Epilepsy classified it among early-onset, treatment-resistant epilepsy syndromes [1-3].

Brief seizures, particularly in infancy, prolong over time, and status epilepticus can occur in up to 75% of cases throughout life (4). Developmental slowdown or regression becomes evident after the age of two; walking and balance problems, loss of motor skills, speech and developmental delays, behavioral problems, and sleep disturbances are common [5-7]. Patients with Dravet syndrome are among the patients with the highest risk of sudden unexpected death (SUDEP), with a reported rate as high as 15-20%. [8].

The variability between genetic mutation and clinical presentation, systemic comorbidities, and the frequent presence of severe seizures despite normal imaging findings make diagnosis and management of the disease challenging. This case report aims to provide a detailed evaluation of the neurological, developmental, and systemic characteristics of a patient diagnosed with Dravet Syndrome.

Corresponding Author: Ozge COSKUN SAGLAM, PhD

Address: *Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Istanbul Bilgi University, Istanbul, Türkiye, Department of Anatomy, Institute of Health Sciences, University of Health Sciences, Ph.D. Program, Istanbul, Türkiye*
E-mail: *ozge.0493@gmail.com*
ORCID ID: *0000-0002-6721-9029*

Case Presentation

The patient was a 36-week preterm female baby born via cesarean section weighing 2700 grams and did not require neonatal intensive care. The family history includes consanguinity between the parents (cousins). The grandfather had been diagnosed with schizophrenia and the grandmother with bipolar disorder. Both of the mother's brothers had a history of epilepsy with febrile seizures. The first seizure occurred at 7 months of age, a hypotonic seizure triggered by fever for 10 seconds, characterized by gazing up, staring, turning blue, and throwing the head back. The family reported that subsequent seizures were triggered by fever, fear, and excitement, and could even occur during defecation. The family did not receive vaccinations due to increased seizure frequency following vaccinations and infections. Seizures were sometimes reported to occur up to three times a day, and intensive care unit admissions due to status epilepticus were observed. Absence, tonic-clonic, and myoclonic seizure types were observed.

Genetic testing revealed a mutation in the SCN1A gene. It has been reported that the same mutation was found heterozygous in the father.

Electroencephalography (EEG) Results

Following the patient's first seizure at 7 months of age, the first EEG following status epilepticus was taken while awake. A sleep EEG was taken at 11 months. Following recurrent febrile seizures at 13 and 16 months of age, a spontaneous sleep EEG was taken. Based on these initial EEG results, although frequency slowing and disorganization were observed in the initial EEG, no epileptiform abnormalities were observed. The results were consistent with the EEG findings seen in the early stages of Dravet Syndrome. EEG findings were extracted retrospectively from the original routine clinical EEG reports generated during standard care at the treating institution. Archived EEG tracings were not reinterpreted for the purposes of this manuscript.
The EEG results taken at 36 months of age showed marked epileptiform activity. The EEG results taken at 37 months of age showed ictal activity originating on the right side and generalizing, along with fixed gaze, nystagmoid movements, and tonic contractions, consistent with status epilepticus. This was one of the most clinically severe seizures, and the EEG findings were found to confirm severe epileptic activity. EEG recordings taken at 38 months revealed widespread disorganization and slow wave paroxysms in the frontal hemispheres. It was noted that rhythmic activity in the brain was disrupted and functional organization was affected.

The EEGs taken at 41 months revealed a marked transition to an encephalopathic course. EEGs at this age reported generalized epileptiform paroxysms, left temporal sharp slow wave activity, and a suggested correlation with Lennox-Gastaut syndrome.

The EEG taken at age 5 revealed frequently repetitive generalized discharges and myoclonic seizure activity in both hemispheres, and a correlation of the myoclonic seizures with the EEG was detected. The EEG at age 6 revealed evidence of modified hypsarrhythmia and epileptic encephalopathy. The last available clinical and EEG follow-up in our records was at 6 years of age. No subsequent EEG recordings or structured developmental assessments were available.

Imaging Results

The patient's cranial CT (computed tomography) results taken at 7 and 12 months were reported as normal, with no evidence of ventricular enlargement, structural abnormalities in the brain, or pathological density. A diffusion-weighted magnetic resonance imaging (MRI) taken at age 3.5 revealed no evidence of restricted diffusion in the cerebral and cerebellar regions (Figure 1). The patient was diagnosed with arrhythmias, particularly those caused by ventricular extrasystoles (VES) and premature atrial contractions (AES). Therefore, the patient underwent cardiac monitoring with an electrocardiogram (ECG) and rhythm Holter monitoring. The corrected QT interval (QTc) was 400 ms and was within normal limits. Transthoracic echocardiography performed at ages 2, 4, and 4.5 years revealed normal cardiac structure and rhythm.

Urinary system ultrasound (US) was performed due to the patient's recurring complaints of crying, burning sensation, and frequent urination. Kidney dimensions and parenchymal thickness were generally normal. However, over time, parenchymal echogenicity increased from grade 1 to grade 2, and crystalloid foci were observed in the right kidney. Grade 1 nephropathy was considered. A hand-wrist radiograph taken at age 3.5 revealed a bone age of 2 years, indicating that the patient was lagging behind her chronological age. The patient's existing developmental delay was also supported radiologically (Figure 2).

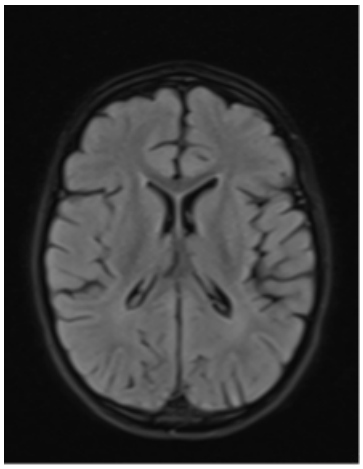


Figure 1. *Diffusion-weighted MRI showing no restricted diffusion in the cerebral and cerebellar regions.*

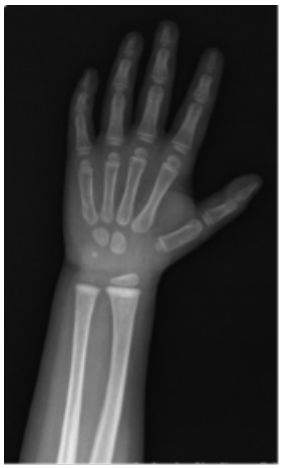


Figure 2. *Radiograph demonstrates delayed bone age of approximately 2 years compared to the chronological age of 3.5 years, supporting the clinical developmental delay.*

Developmental Results

The developmental status of the patient was evaluated according to the results of the Denver II Developmental Screening Test performed at regular intervals since birth and clinical observations. The results of the evaluation are given in Table 1, along with the details.

Drugs Used and Responses to Treatment

The patient has been treated with various antiepileptic drugs and treatment combinations at different times since diagnosis. The primary medications used were valproic acid, levetiracetam, topiramate, phenobarbital, clobazam, stiripentol, and zonisamide. A combination of valproic acid, levetiracetam, topiramate, and phenobarbital was initiated in the first year, but the regimen was repeatedly adjusted due to persistent seizures and status epilepticus. Although partial seizure control was achieved with valproic acid and phenobarbital in the second year, generalized ictal activity and status attacks persisted even after levetiracetam and clobazam were added around the third year of age.

At the age of four, Lennox-Gastaut-like patterns were detected on the EEG, and despite the addition of stiripentol, the seizures remained uncontrolled. Between the ages of five and six, multiple drug combinations, including zonisamide, were tried. The ketogenic diet, which was used as an alternative treatment for approximately two years, was discontinued because it did not provide significant clinical benefit.

Systemic Findings and Comorbidities

The patient's recurring complaints included constipation and anal fissures. She had a history of parasitic infestations (pinworms), and antiparasitic treatment was administered. Constipation was also a contributing factor to the patient's stress. Because she had stress-related seizures, constipation affected her quality of life.

The patient underwent an allergy evaluation due to eczema, blotchy pigmentation in the left inguinal region, and recurring rashes on the face and legs. No specific allergen was identified, but symptoms were monitored. Hearing test and eye examination results were normal. Motor symptoms such as impaired balance and intention tremor were found and these symptoms still persist.

Discussion

SCN1A mutations have been reported in approximately 80% of cases, and a small proportion of these variants are inherited from a parent (9). In contrast, de novo SCN1A mutations have been reported to arise more frequently on the paternal chromosome (10). As emphasized by Česká et al., the clinical spectrum in SCN1A carriers can range from simple febrile seizures to typical Dravet syndrome (11). In our case, an SCN1A mutation was also detected, and the fact that the same variant was heterozygous in the father, who was apparently unaffected, suggests variable penetrance and expressivity.

Dravet syndrome usually begins with prolonged febrile/afebrile focal clonic or generalized clonic seizures; seizures often appear between 1 and 4 years of age, can be refractory to treatment, and can progress to status epilepticus [12, 13]. In our case, the first finding was a febrile seizure that began at 7 months of age; tonic, myoclonic, and apneic seizures were subsequently added to the picture. Regarding the EEG, it has been reported that the EEG is frequently normal in the early stages of Dravet syndrome, and epileptiform abnormalities become more pronounced with age [14]. Kapoor et al. have shown that children with initially normal EEG and brain MRI develop significant neurological deterioration over time [15]. In our case, while the early EEG recordings were mostly normal or mildly abnormal, follow-up revealed generalized epileptiform discharges, patterns suggestive of Lennox–Gastaut syndrome, and ultimately modified hypsarrhythmia; this course reflects the progressive nature of the disease, which progresses to epileptic encephalopathy.

Developmentally, our Denver II findings revealed a developmental delay that partially parallels the epileptic activity but spreads to all areas over time. It has been reported that language and communication delays in children with Dravet syndrome can begin early, independent of seizures, and that development slows significantly after the age of two [16]. A mild delay in early language development was observed in this case, and the lack of new age-appropriate acquisitions after the age of 2–3 was consistent with progressive language and developmental delays described in the literature.

Despite the use of multiple antiseizure medication combinations and a ketogenic diet during the treatment period, complete seizure control was never achieved. Dravet syndrome management in contemporary practice is largely guided by international expert consensus, which emphasizes early recognition, individualized polytherapy, and avoidance of sodium channel blocking antiseizure medications in SCN1A related disease (17). A Norwegian series reported that the vast majority of patients with Dravet syndrome were taking at least two different antiseizure medications, and a substantial proportion were taking four or five different antiseizure medications [18]. Randomized controlled trials also support syndrome specific therapies. Cannabidiol has been shown to significantly reduce convulsive seizure frequency in patients with Dravet syndrome, and open label extension data suggest that clinically meaningful seizure reduction can be sustained over longer follow up [19,20].

Stiripentol has shown efficacy in a syndrome focused placebo controlled trial when added to clobazam and valproate in severe myoclonic epilepsy in infancy or Dravet syndrome [21]. Ketogenic dietary therapies remain an important non pharmacological option for refractory seizures in Dravet syndrome, and pooled evidence from a recent systematic review and meta analysis suggests clinically meaningful seizure reduction, although adherence may vary over time [22, 23]. Similarly, in our case, a wide range of medications were tried, but refractory seizures and epileptic encephalopathy persisted.

A survey conducted by the Dravet Syndrome Foundation reported that constipation, allergic diseases, and respiratory tract infections, along with sleep disorders, photosensitivity, gait disturbances, and cardiac problems, were common comorbidities [24]. In our case, comorbidities such as constipation, allergic skin findings, balance disorders, tremors, and cardiac arrhythmias were observed. The complex nature of Dravet syndrome, which affects multiple organ systems, further emphasizes the importance of multidisciplinary patient follow-up. Dravet syndrome is increasingly recognized as a disorder with multisystem comorbidities. Gastrointestinal problems, including constipation, are commonly reported and may adversely affect quality of life and stress levels, which can in turn influence seizure susceptibility [25]. Dermatologic symptoms were monitored and may represent common pediatric conditions or treatment-related effects rather than a direct manifestation of SCN1A-related disease. Renal ultrasonography changes (progressive increase in parenchymal echogenicity and crystalloid foci, with grade 1 nephropathy considered) were non-specific and were followed clinically; these findings may be incidental or related to intercurrent urinary symptoms, metabolic factors rather than a direct manifestation of SCN1A-related disease. Given the reported autonomic involvement and cardiorespiratory vulnerability in Dravet syndrome, the observed arrhythmias were managed as clinically relevant comorbidities requiring surveillance, without implying a syndrome-specific causal link.

Conclusion

This case report of a patient with Dravet Syndrome associated with an SCN1A mutation meticulously documents the progressive nature of the disease across multiple dimensions. Given that delays in diagnosis and treatment are associated with poorer long-term cognitive and behavioral outcomes, this report advocates for the rapid implementation of early genetic testing and diagnostic protocols. Early diagnosis and an interdisciplinary approach are crucial to improving the long-term prognosis of patients struggling with this complex and devastating syndrome.

Limitations

Incomplete documentation of some antiseizure medications during the treatment course, which is frequently encountered in routine clinical practice, represents a key limitation of retrospective case reports. A further limitation is the lack of clinical and developmental follow-up data beyond 6 years of age due to incomplete medical records.

Acknowledgments and Funding Statement

None to declare.

7. Conflicts of Interest

The authors declare that there are no conflicts of interest regarding this case report.

8. Ethics and consent to participate

Ethical approval was obtained from the Istanbul Bilgi University Ethics Committee (Decision date: 22 December 2025; Project/Protocol No: 2025-40162-267; Document No: 01.01.2026-56135). This case report was prepared retrospectively using de-identified clinical data obtained during routine care. All identifiable information has been removed, and the report does not contain any images that could identify the patient.

References

1. Dravet C. Dravet syndrome history. Dev Med Child Neurol. 2011;53(Suppl 2):S1-S6.

2. He Z, Li Y, Zhao X, Li B. Dravet syndrome: advances in etiology, clinical presentation, and treatment. Epilepsy Res. 2022;188:107041. doi:10.1016/j.eplepsyres.2022.107041.

3. Li W, Schneider AL, Scheffer IE. Defining Dravet syndrome: an essential pre-requisite for precision medicine trials. Epilepsia. 2021;62(9):2205-2217. doi:10.1111/epi.17015.

4. Strzelczyk A, Kalski M, Bast T, Wiemer-Kruel A, Bettendorf U, Kay L, et al. Burden-of-illness and cost-driving factors in Dravet syndrome patients and carers: a prospective, multicenter study from Germany. Eur J Paediatr Neurol. 2019;23(3):392-403. doi:10.1016/j.ejpn.2019.02.014.

5. Han Z, Chen C, Christiansen A, Ji S, Lin Q, Anumonwo C, et al. Antisense oligonucleotides increase Scn1a expression and reduce seizures and SUDEP incidence in a mouse model of Dravet syndrome. Sci Transl Med. 2020;12(558):eaaz6100. doi:10.1126/scitranslmed.aaz6100.

6. Cooper MS, McIntosh A, Crompton DE, McMahon JM, Schneider A, Farrell K, et al. Mortality in Dravet syndrome. Epilepsy Res. 2016;128:43-47. doi:10.1016/j.eplepsyres.2016.10.006.

7. Shao W, Liu L, Gu J, Yang Y, Wu Y, Zhang Z, et al. Spotlight on mechanism of sudden unexpected death in epilepsy in Dravet syndrome. Transl Psychiatry. 2025;15(1):84. doi:10.1038/s41398-025-03304-8.

8. Kearney J. Sudden unexpected death in Dravet syndrome: SUDEP in Dravet mouse model. Epilepsy Curr. 2013;13(6):264-265.

9. Cetica V, Chiari S, Mei D, Parrini E, Grisotto L, Marini C, et al. Clinical and genetic factors predicting Dravet syndrome in infants with SCN1A mutations. Neurology. 2017;88(11):1037-1044. doi:10.1212/WNL.0000000000003716.

10. Heron SE, Scheffer IE, Iona X, Zuberi SM, Birch R, McMahon JM, et al. De novo SCN1A mutations in Dravet syndrome and related epileptic encephalopathies are largely of paternal origin. J Med Genet. 2010;47(2):137-141.

11. Ceska K, Danhofer P, Horak O, Spanelova K, Kolar S, Oslejskova H, et al. Phenotypic spectrum of the SCN1A mutation (from febrile seizures to Dravet syndrome). Bratisl Lek Listy. 2022;123(7):483-486.

12. Fan HC, Yang MT, Lin LC, Chiang KL, Chen CM. Clinical and genetic features of Dravet syndrome: a prime example of the role of precision medicine in genetic epilepsy. Int J Mol Sci. 2023;25(1):31. doi:10.3390/ijms25010031.

13. Zuberi SM, Wirrell E, Yozawitz E, Wilmshurst JM, Specchio N, Riney K, et al. ILAE classification and definition of epilepsy syndromes with onset in neonates and infants: Position statement by the ILAE Task Force on Nosology and Definitions. Epilepsia. 2022;63(6):1349-1397. doi:10.1111/epi.17239.

14. Minato E, Myers KA. Age-related evolution of EEG in Dravet syndrome: Meta-analysis of 155 patients. Seizure. 2021;91:108-111. doi:10.1016/j.seizure.2021.06.007.
15. Kapoor D, Anand A, Sharma S, Mukherjee SB, Marini C, Mei D, et al. Dravet syndrome: a case series. Indian J Pediatr. 2021;88:82. doi:10.1007/s12098-020-03383-z.

16. Fine AL. ENVISIONing a critical period to preserve development: communication delays in SCN1A+ Dravet syndrome. Epilepsy Curr. 2024;24(5):342-344. doi:10.1177/15357597241280687.

17. Wirrell EC, Hood V, Knupp KG, Meskis MA, Nabbout R, Scheffer IE, et al. International consensus on diagnosis and management of Dravet syndrome. Epilepsia. 2022;63(7):1761-1777. doi:10.1111/epi.17274.

18. Heger K, Lund C, Larsen Burns M, Bjørnvold M, Sætre E, Johannessen SI, et al. A retrospective review of changes and challenges in the use of antiseizure medicines in Dravet syndrome in Norway. Epilepsia Open. 2020;5(3):432-441. doi:10.1002/epi4.12413.

19. Scheffer IE, Halford JJ, Miller I, Nabbout R, Sanchez-Carpintero R, Shiloh-Malawsky Y, et al. Add-on cannabidiol in patients with Dravet syndrome: results of a long-term open-label extension trial. Epilepsia. 2021;62(10):2505-2517. doi:10.1111/epi.17036.

20. Sullivan J, Lagae L, Cross JH, Devinsky O, Guerrini R, Knupp KG, et al. Fenfluramine in the treatment of Dravet syndrome: results of a third randomized, placebo-controlled clinical trial. Epilepsia.
2023;64(10):2653-2666. doi:10.1111/epi.17737.

21. Chiron C, Marchand MC, Tran A, Rey E, d'Athis P, Vincent J, et al. Stiripentol in severe myoclonic epilepsy in infancy: a randomised placebo-controlled syndrome-dedicated trial. Lancet. 2000;356(9242):1638-1642.

22. Li Y, Dai S, Deng Q, Liu Z. The impact and adherence of ketogenic dietary therapies for Dravet syndrome: a systematic review and meta-analysis. Seizure. 2025;130:16-24. doi:10.1016/j.seizure.2025.05.001.

23. Yu M, Li H, Sun D, Li D, Zhong J, Gu Q, et al. The ketogenic diet for Dravet syndrome: a multicenter retrospective study. Nutrition. 2023;110:111976.

24. Villas N, Meskis MA, Goodliffe S. Dravet syndrome: characteristics, comorbidities, and caregiver concerns. Epilepsy Behav. 2017;74:81-86.

25. Minderhoud CA, Postma A, Jansen FE, Verhoeven JS, Schrijver JJ, Goudswaard J, et al. Gastrointestinal and eating problems in SCN1A-related seizure disorders. Epilepsy Behav. 2023;146:109361. doi:10.1016/j.yebeh.2023.109361.

YALOVA ÜNİVERSİTESİ
TIP FAKÜLTESİ