

KANSER İLİŞKİLİ LENFÖDEMDE CHATGPT’NİN FİZYOTERAPİ, EGZERSİZ VE YAŞAM STİLİ ÖNERİLERİ İLE İLGİLİ REHBERLİĞİNİN DOĞRULUK, KALİTE VE OKUNABİLİRLİK AÇISINDAN DEĞERLENDİRİLMESİ

Miray HASPOLAT¹, Canan ERGİN²

¹ Çukurova Üniversitesi Sağlık Bilimleri Fakültesi, Fizyoterapi ve Rehabilitasyon Bölümü, Adana, Türkiye

 0000-0002-9002-0253

² Gazi Üniversitesi Sağlık Bilimleri Fakültesi, Fizyoterapi ve Rehabilitasyon Bölümü, Ankara, Türkiye

 0009-0003-9133-7951

ÖZ

Giriş: ChatGPT, sağlıkla ilişkili bilgi arayışında giderek daha yaygın kullanılan büyük dil modellerinden biridir. Bununla birlikte, yapay zekâ sistemlerinin sunduğu bilgilerin doğruluğu ve kalitesi özellikle klinik durumlar açısından önemli bir tartışma konusudur.

Amaç: Bu çalışma ChatGPT’nin kanser ilişkili lenfödemde fizyoterapi, egzersiz ve yaşam stili önerileri ile ilgili rehberliğinin doğruluk, kalite ve okunabilirlik açısından değerlendirilmesi amaçladı.

Yöntem: Kanser ilişkili lenfödemde fizyoterapi, egzersiz ve yaşam stili önerileri ile ilgili olarak bu alanda deneyimli fizyoterapistlere hastalarının sıklıkla sorduğu 10 soru belirlendi. Her soru 2026 yılı Nisan ayında ChatGPT-5.3’e Türkçe ve bağımsız oturumlar kullanılarak yöneltildi. Yanıtlar lenfödem terapisi sertifikalı 2 bağımsız fizyoterapist tarafından değerlendirildi. Doğruluk, kalite ve okunabilirliğin değerlendirmesinde sırasıyla 5 puanlık Likert ölçeği, Global Kalite Ölçeği ve Flesch Okuma Kolaylığı Formülü kullanıldı. Değerlendiriciler arası uyum, yüzde uyum ve Cohen’s kappa katsayısı ile değerlendirildi.

Bulgular: Tüm yanıtlar için doğruluk ve kalite skorları 4 ve üzerindeydi. Doğruluk, kalite ve okunabilirlik ortanca (çeyrekler arası aralık) değerleri sırasıyla 4,5 (4-5), 4 (4-5), 56.2 (54.5-67.1) idi. Doğruluk ve kalite için değerlendiriciler arası uyum sırasıyla %70 ve %80 olarak bulundu. Cohen’s kappa katsayıları ise sırasıyla 0.40 ve 0.60 olup, zayıf-orta ile orta arasında uyuma işaret etmektedir.

Sonuç: Bulgular, ChatGPT-5.3’ün yüksek doğruluk ve kaliteye sahip yanıtlar ürettiğini düşündürmektedir ancak soru ve değerlendirici sayısının az olması genellenebilirliği sınırlayabilir. Ayrıca yanıtların okunabilirliği en az lise düzeyine uygun olup, bu durum erişilebilirliği sınırlayabilir. Bu nedenle ChatGPT-5.3, genel bilgi edinme amacıyla yararlı bir araç olmakla birlikte, bireyselleştirilmiş tedavi yaklaşımları için sağlık profesyoneline danışılması önerilir.

Anahtar Kelimeler: Yapay zeka, lenfödem, egzersiz, rehabilitasyon

* Bu çalışma, 16-17 Nisan 2026’da III. Ulusal Onkolojik Rehabilitasyon Kongresi’nde sözel bildiri olarak sunulmuştur.

EVALUATION OF THE ACCURACY, QUALITY, AND READABILITY OF CHATGPT'S GUIDANCE ON PHYSIOTHERAPY, EXERCISE, AND LIFESTYLE RECOMMENDATIONS IN CANCER-RELATED LYMPHEDEMA

Miray HASPOLAT¹, Canan ERGİN²

¹ Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Cukurova University, Adana, Turkey

 0000-0002-9002-0253

² Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Gazi University, Ankara, Turkey

 0009-0003-9133-7951

ABSTRACT

Background: ChatGPT is an increasingly used large language model for obtaining health-related information. However, the accuracy and quality of information provided by artificial intelligence systems remain a concern, particularly in clinical settings.

Aim: This study aimed to evaluate the accuracy, quality, and readability of ChatGPT's guidance on physiotherapy, exercise, and lifestyle recommendations in cancer-related lymphedema.

Method: In cancer related lymphedema, 10 frequently asked questions by patients to physiotherapists experienced in this field regarding physiotherapy, exercise, and lifestyle recommendations were identified. Each question was presented once to ChatGPT-5.3 in Turkish using independent sessions in April 2026. Responses were evaluated by two independent physiotherapists certified in lymphedema therapy. Accuracy, quality, and readability were assessed using a 5-point Likert scale, the Global Quality Scale, and the Flesch Reading Ease Formula, respectively. Inter-rater agreement was analyzed using percentage agreement and Cohen's kappa.

Results: Accuracy and quality scores were 4 and above for all responses. Median (interquartile range) values for accuracy, quality, and readability were 4.5 (4–5), 4 (4–5), and 56.2 (54.5–67.1), respectively. Inter-rater agreement was 70% for accuracy and 80% for quality. Cohen's kappa values were 0.40 and 0.60, indicating fair to moderate agreement.

Conclusion: The findings suggest that ChatGPT-5.3 produces responses with high accuracy and quality; however, the small number of questions and evaluators may limit generalizability. Furthermore, the readability of the responses is at least suitable for high school level, which may limit its accessibility. Therefore, ChatGPT-5.3 may be useful for general information, but professional consultation remains essential for individualized treatment.

Keywords: Artificial intelligence, lymphedema, exercise, rehabilitation

* This study was presented as an oral presentation at the 3rd National Oncological Rehabilitation Congress on 16-17 April 2026.