

Occupational Health and Safety in Urological Surgeries: A Comparison of the Performance of Artificial Intelligence-Based Large Language Models (ChatGPT, Gemini, Perplexity, DeepSeek)

Ürolojik Ameliyatlarda İş Sağlığı ve Güvenliği: Yapay Zeka Tabanlı Büyük Dil Modellerinin (ChatGPT, Gemini, Perplexity, DeepSeek) Performanslarının Karşılaştırılması

Şeyma Demir¹ , Murat Demir² 

¹ Department of Occupational Health and Safety, Van Yüzüncü Yıl University, Van, Türkiye

² Department of Urology, Van Yüzüncü Yıl University, Van, Türkiye

ABSTRACT

Objective: This study aims to comparatively evaluate the knowledge level, guideline compliance, scope, and consistency of responses provided by artificial intelligence-based large language models (ChatGPT, Gemini, Perplexity, DeepSeek) to questions related to occupational health and safety (OHS) in urological surgeries.

Materials and Methods: Based on the literature and international guidelines (EAU, WHO, OSHA, etc.), 10 standardized questions covering biological, chemical, physical, ergonomic, environmental, and psychosocial risks were developed. These questions were presented to four different models on the same day and in the same format. The responses were blindly evaluated by three independent experts using DISCERN and Global Quality Score (GQS) criteria. Statistical analyses were performed using ANOVA/Kruskal-Wallis tests and Fleiss' kappa.

Results: ChatGPT and Gemini achieved the highest performance in both DISCERN and GQS scores. Perplexity demonstrated moderate performance, while DeepSeek showed the lowest quality and consistency. ChatGPT scored significantly higher than Perplexity and DeepSeek ($p < 0.01$), and Gemini was also significantly superior to Perplexity and DeepSeek ($p < 0.05$). No significant difference was found between ChatGPT and Gemini.

Conclusion: AI-based large language models can provide supportive information on OHS in urological surgeries; however, there are notable quality differences among models. While ChatGPT and Gemini deliver more reliable and comprehensive content, the limitations of Perplexity and especially DeepSeek are evident. Therefore, AI tools should not be used as standalone decision-makers in clinical practice but should be supported by expert oversight and guideline-based validation.

Keywords: artificial intelligence, ChatGPT, DeepSeek, Gemini, large language model, occupational health and safety, Perplexity, urological surgery

Cite As: Demir S, Demir M. Occupational Health and Safety in Urological Surgeries: A Comparison of the Performance of Artificial Intelligence-Based Large Language Models (ChatGPT, Gemini, Perplexity, DeepSeek). Endourol Bull. 2026;18(2):73-79. <https://doi.org/10.54233/endourolbull-1792399>

Corresponding Author: Şeyma Demir, Van Yüzüncü Yıl University, Van, Türkiye

e-mail: seymacankurtarandemir@gmail.com

Received : September 27, 2025

Accepted : April 16, 2026



ÖZET

Amaç: Bu çalışma, yapay zeka (YZ) tabanlı büyük dil modellerinin (ChatGPT, Gemini, Perplexity, DeepSeek) ürolojik ameliyatlarda iş sağlığı ve güvenliği (İSG) ile ilgili sorulara verdikleri yanıtların bilgi düzeyi, rehberlerle uyumu, kapsamı ve tutarlılığını karşılaştırmalı olarak değerlendirmeyi amaçlamaktadır.

Gereç ve Yöntemler: Literatür ve uluslararası kılavuzlar (EAU, WHO, OSHA vb.) temel alınarak biyolojik, kimyasal, fiziksel, ergonomik, çevresel ve psikososyal riskleri kapsayan 10 standart soru hazırlanmıştır. Bu sorular dört farklı modele aynı gün ve aynı formatta yöneltilmiş, yanıtlar üç bağımsız uzman tarafından DISCERN ve GQS kriterleriyle kör olarak değerlendirilmiştir.

Bulgular: ChatGPT ve Gemini, hem DISCERN hem de GQS skorlarında en yüksek performansı sergilemiştir. Perplexity orta düzeyde, DeepSeek ise en düşük kalite ve tutarlılığı göstermiştir. ChatGPT, Perplexity ve DeepSeek'ten anlamlı derecede daha yüksek skor alırken ($p<0,01$), Gemini de Perplexity ve DeepSeek'ten anlamlı şekilde üstündür ($p<0,05$). ChatGPT ile Gemini arasında anlamlı fark bulunmamıştır.

Sonuç: YZ tabanlı büyük dil modelleri, ürolojik ameliyatlarda İSG konusunda bilgi desteği sağlayabilmektedir; ancak modeller arasında önemli kalite farklılıkları mevcuttur. ChatGPT ve Gemini daha güvenilir ve kapsamlı içerikler sunarken, Perplexity ve özellikle DeepSeek'in sınırlılıkları belirgindir. Bu nedenle YZ araçları, klinik karar desteğinde tek başına kullanılmamalı, uzman denetimi ve kılavuz temelli doğrulamalarla desteklenmelidir.

Anahtar Kelimeler: büyük dil modeli, ChatGPT, DeepSeek, Gemini, iş sağlığı ve güvenliği, Perplexity, ürolojik cerrahi, yapay zeka

GİRİŞ

Ameliyathane ortamı, sağlık çalışanlarının biyolojik, kimyasal, fiziksel, ergonomik ve psikososyal pek çok risk faktörüyle karşılaştığı yüksek riskli bir çalışma alanıdır. Ürolojik ameliyatlarda ise lazer, radyasyon ve elektrocerrahi cihazlarının yoğun kullanımı nedeniyle iş sağlığı ve güvenliği (İSG) açısından özel bir önem taşımaktadır (1). Bu operasyonlarda cerrah ve ameliyathane ekibi; kan yoluyla bulaşan enfeksiyonlar, kesici-delici yaralanmalar, cerrahi duman, anestezi gazları, radyasyon ve lazer ışınları gibi tehlikelere maruz kalabilmektedir. Ayrıca, uzun süren operasyonlarda ergonomik sorunlar, iş stresi ve atık yönetimi de göz ardı edilmemesi gereken risk alanlarıdır (2,3).

Son yıllarda yapay zeka (YZ) tabanlı büyük dil modellerinin (Large Language Models, LLM) sağlık alanında kullanımı hızla artmaktadır. Bu modeller, kullanıcılarla doğal dilde etkileşim kurarak bilgi sağlayabilmekte, literatür özetleme, klinik eğitim, karar destek ve hasta bilgilendirme gibi çeşitli alanlarda potansiyel faydalar sunmaktadır (4). Ancak, bu sistemlerin güvenilirliği ve rehberlerle uyumluluğu özellikle hasta güvenliği ve çalışan sağlığı açısından kritik bir konudur (5).

Mevcut dil modelleri arasında ChatGPT (OpenAI), Gemini (Google DeepMind), DeepSeek ve Perplexity AI öne çıkan sistemlerdir. Her biri farklı veri setleri, eğitim teknikleri ve kullanıcı arayüzleri ile geliştirilmiştir. ChatGPT doğal dilde yüksek akıcılığı ile bilinirken, Gemini çok modlu yetenekleriyle dikkat çekmektedir. DeepSeek araştırma tabanlı yaklaşımıyla öne çıkarken, Perplexity ise literatür tarama ve kaynak göstermeye yönelik özellikleriyle farklılaşmaktadır (6). Bu modellerin ürolojik cerrahilerde İSG konularında verdikleri yanıtların doğruluk, kapsam ve tutarlılığının karşılaştırılması, hem tıp eğitimi hem de klinik karar destek sistemleri açısından önemli bir boşluğu dolduracaktır.

Bu çalışmanın amacı, ChatGPT, DeepSeek, Gemini ve Perplexity YZ modellerinin ücretsiz versiyonlarının ürolojik ameliyatlarda İSG ile ilgili sorulara verdikleri yanıtların bilgi düzeyi, rehberlerle uyumu, kapsamı ve tutarlılığını karşılaştırmalı olarak değerlendirmektir. Böylelikle, farklı YZ platformlarının güçlü ve zayıf yönleri ortaya konarak sağlık alanında güvenli ve etkin kullanım potansiyelleri hakkında bilimsel katkı sağlanması amaçlanmaktadır.

GEREÇ VE YÖNTEMLER

Çalışma Tasarımı

Bu çalışma, büyük dil modellerinin (ChatGPT, DeepSeek, Gemini, Perplexity) ürolojik ameliyatlarda İSG ile ilgili sorulara verdikleri yanıtların karşılaştırmalı olarak değerlendirilmesi amacıyla yapılmış kesitsel ve metodolojik bir içerik analizi olarak tasarlanmıştır.

Soru Seti

Literatür taraması (EAU kılavuzları, WHO/CDC, OSHA/NIOSH, Türkiye İSG mevzuatı) ve uzman görüşleri temel alınarak altı ana kategoride sorular oluşturulmuştur:

1. Biyolojik riskler
2. Fiziksel riskler
3. Kimyasal riskler
4. Ergonomik riskler
5. Çevresel riskler
6. Psikososyal riskler

Her kategoriden en az bir soru olacak şekilde toplam 10 standart soru hazırlanmıştır (Tablo 1). Bu sorular açık uçlu ve senaryo temelli olacak şekilde düzenlenmiştir.

Tablo 1. Yapay Zeka modellerine sorulan sorular

Kategori	Soru
Biyolojik Riskler	Ürolojik ameliyatlarda cerrah ve hemşireler için en önemli biyolojik iş sağlığı riskleri nelerdir?
Biyolojik Riskler	TURP sırasında kan yoluyla bulaşabilecek enfeksiyonlardan korunmak için hangi önlemler alınmalıdır?
Fiziksel Riskler	Perkütan nefrolitotomi (PNL) sırasında yoğun floroskopi kullanımı nedeniyle çalışanların radyasyondan korunma yöntemleri nelerdir?
Fiziksel Riskler	Holmium:YAG lazer ile taş kırma sırasında cerrahi ekip hangi güvenlik önlemlerini almalıdır?
Kimyasal Riskler	Ameliyathanede kullanılan anestezi gazlarının sağlık çalışanları üzerindeki etkilerini azaltmak için hangi yöntemler uygulanmalıdır?
Ergonomi	Laparoskopik radikal prostatektomi sırasında cerrahın uzun süreli ayakta çalışmasına bağlı ergonomik riskler ve çözüm önerileri nelerdir?
Çevresel Riskler	Ürolojik ameliyatlarda ortaya çıkan kesici-delici atıkların güvenli şekilde yönetilmesi için hangi prosedürler uygulanmalıdır?
Psikososyal Riskler	Ameliyathanede zaman baskısı altında çalışan cerrahi ekibin iş güvenliği açısından yaşadığı psikososyal riskler nelerdir?
Senaryo Temelli	Bir hemşire, TURP sırasında eldivenine iğne batması durumunda hangi adımları izlemelidir?
Senaryo Temelli	Robotik radikal prostatektomi sırasında ekipman arızası yaşanırsa İSG açısından öncelikli yapılması gerekenler nelerdir?

Veri Toplama

- Sorular, her bir dil modeline (ChatGPT, DeepSeek, Gemini, Perplexity) aynı gün ve aynı standart istem formatı kullanılarak yöneltilmiştir.
- Her modelden alınan yanıtlar zaman damgası ile birlikte kaydedildi.
- Tutarlılığı değerlendirmek için aynı soru seti, her model için 3 farklı oturumda tekrarlanmıştır.
- Tüm çıktılar ham veri olarak arşivlendi.

Değerlendirme Süreci

Yanıtlar iki bağımsız uzman (1 ürolog, 1 İSG doktora öğrencisi) tarafından körleme yöntemiyle değerlendirilmiştir. Sonrasında skorlar üzerine ortak puan verilmeye kadar tartışılmıştır. Değerlendirme kriterleri:

Brief DISCERN (Kısa Versiyon) (7) (Tablo 2)

- 6 temel sorudan oluşur.
- Amaç, sağlık bilgisinin kalite ve güvenilirliğini hızlı bir şekilde değerlendirmektir.
- Sorular; bilginin güvenilirliği, kaynak gösterme, güncellik, tarafsızlık, tedavi seçeneklerinin açıklığı ve genel kaliteyi kapsar.
- Her soru 1–5 arasında puanlanır.

Tablo 2. Discern skor maddeleri

Madde	Soru	Puanlama (1–5)
1	Kaynağın amaçları açıkça belirtilmiş mi?	1–5
2	Verilen bilgilerin kaynakları güvenilir ve açıkça belirtilmiş mi?	1–5
3	Bilgiler güncel mi?	1–5
4	Bilgi tarafsız mı, yoksa ticari/sübjektif yönlendirme içeriyor mu?	1–5
5	Tedavi/önlem seçenekleri, faydaları ve riskleriyle birlikte dengeli şekilde sunulmuş mu?	1–5
6	Genel olarak bu kaynağın sunduğu bilgi kalitesini nasıl değerlendirirsiniz?	1–5

GQS (Global Quality Score) (8) (Tablo 3)

GQS, sağlık bilgisi sağlayan dijital içeriklerin genel kalite, kullanılabilirlik ve bilgilendiricilik düzeyini ölçmek için kullanılan basit ve yaygın bir ölçektir.

- 1–5 arasında tek maddeden oluşan bir değerlendirme ölçeğidir.
- Kullanıcıya verilen bilginin anlaşılabilirliği, kapsamı, klinik açıdan faydası ve güvenilirliği göz önünde bulundurulur.
- Daha çok video, web sayfası veya YZ yanıtlarının kullanıcıya pratik fayda düzeyini değerlendirmede kullanılır.

Tablo 3. GQS Tablosu

Skor	Açıklama
1	Çok kötü kalite – Hiç faydalı değil, güvenilmez, yanlış/eksik bilgiler
2	Zayıf kalite – Kısmen faydalı ama ciddi eksiklikler içeriyor
3	Orta kalite – Bir miktar faydalı, ama önemli eksiklikler var
4	İyi kalite – Genellikle faydalı, yeterli kapsam ve güvenilirlik mevcut
5	Mükemmel kalite – Çok faydalı, kapsamlı, güvenilir, klinik açıdan güçlü

İstatistiksel Analiz

Skor değerleri için tanımlayıcı istatistikler; Ortalama, Standart Sapma, Medyan, Q1, Q3, Minimum ve Maksimum değerler olarak ifade edilmiştir. Bu değerler bakımından gruplar arasında fark olup olmadığını belirlemek üzere Kruskal-Wallis testi yapılmıştır. Sonrasında farklı grupları belirlemek üzere, ikili karşılaştırmalar için Dunn-Benferroni testi kullanılmıştır. Hesaplamalarda istatistik anlamlılık düzeyi %5 olarak alınmış ve hesaplamalar için SPSS (versiyon 21) istatistik paket programı kullanılmıştır

BULGULAR

Çalışmamızda ChatGPT için DISCERN skoru $4,8 \pm 0,422$ (medyan [Q1-Q3]: 5,0 [4,75 – 5,0]) olarak tespit edilirken, Gemini için bu değer $4,6 \pm 0,516$ (medyan [Q1-Q3]: 5,0 [4,00–5,0]) olarak saptandı. Perplexity için DISCERN skoru $3,7 \pm 0,483$ (medyan [Q1-Q3]: 4,0 [3,00–4,0]) iken, DeepSeek için $4,1 \pm 0,738$ (medyan [Q1-Q3]: 4,0 [3,75–5,0]) olarak tespit edildi (Tablo 3). GQS skorları incelendiğinde ChatGPT $4,7 \pm 0,483$ (medyan [Q1-Q3]: 5,0 [4,00–5,0]) bulunurken, Gemini için $4,5 \pm 0,707$ (medyan [Q1-Q3]: 5,0 [4,00–5,0]), Perplexity için $3,6 \pm 0,516$ (medyan [Q1-Q3]: 4,0 [3,00–5,0]), DeepSeek için ise $4,0 \pm 0,816$ (medyan [Q1-Q3]: 4,0 [3,00–4,0]) olarak tespit edildi (Tablo 4).

Tablo 4. Yapay zeka modellerinin DİSCERN ve GQS karşılaştırılması

	Gruplar	Mean ± SD	Median [Q1 – Q3]	Min.	Max.	p
DİSCERN	CHATGPT	4,8 ± 0,422	a 5,0 [4,75 – 5,0]	4,0	5,0	0,001
	GEMİNİ	4,6 ± 0,516	a 5,0 [4,00 – 5,0]	4,0	5,0	
	DEEPSEEK	4,1 ± 0,738	b 4,0 [3,75 – 5,0]	3,0	5,0	
	PERPLEXİTY	3,7 ± 0,483	b 4,0 [3,00 – 4,0]	3,0	4,0	
GQS	CHATGPT	4,7 ± 0,483	a 5,0 [4,00 – 5,0]	4,0	5,0	0,002
	GEMİNİ	4,5 ± 0,707	a 5,0 [4,00 – 5,0]	3,0	5,0	
	DEEPSEEK	4,0 ± 0,816	b 4,0 [3,00 – 5,0]	3,0	5,0	
	PERPLEXİTY	3,6 ± 0,516	b 4,0 [3,00 – 4,0]	3,0	4,0	

a, b: Farklı harfi alan gruplar arası fark anlamlıdır

GQS: Global Quality Score

Genel olarak ChatGPT ve Gemini en yüksek performansı gösterirken, DeepSeek en düşük, Perplexity ise orta düzey performans sergiledi.

TARTIŞMA

Ameliyathane, sağlık çalışanlarının çok çeşitli biyolojik, kimyasal, fiziksel, ergonomik ve psikososyal risklerle karşılaştığı yüksek riskli bir çalışma ortamıdır. Ürolojik ameliyatlarda bu riskler lazer, radyasyon, cerrahi duman ve kesici-delici yaralanmalar gibi özel durumlarla daha da artmaktadır (9). Bu nedenle çalışanların İSG konusunda bilgi düzeylerinin yüksek olması kritik öneme sahiptir. Ancak literatürde, sağlık çalışanlarının İSG konusunda farkındalıklarının ve uygulamalarının çoğu zaman yetersiz kaldığı gösterilmiştir. Enfeksiyon kontrol önlemlerine uyumda eksiklikler, kişisel koruyucu ekipman kullanımında ihlaller ve ergonomik risklerin göz ardı edilmesi sık rapor edilmektedir (10).

Son yıllarda YZ tabanlı sistemlerin sağlık alanında yaygınlaşmasıyla birlikte, çalışanlar bilgi eksikliği yaşadıklarında veya emin olmadıklarında ChatGPT, Gemini, Perplexity gibi YZ uygulamalarına yönelmektedir (11). Özellikle hızlı yanıt alabilme, kolay erişilebilirlik ve pratik bilgi ihtiyacı bu eğilimi güçlendirmektedir. Ancak bu noktada kritik soru şudur: YZ modellerinin verdiği yanıtların düzeyi, çalışanların güvenli ve doğru uygulamalar yapması için yeterli midir? (11).

Çalışmamızın sonuçları, dört farklı YZ modelinin ürolojik cerrahilerde İSG konularında verdikleri yanıtların kalitesinin değişken olduğunu ortaya koymuştur. ChatGPT ve Gemini, literatür ve rehberlere uyumlu, kapsamlı ve pratik olarak uygulanabilir yanıtlar üretmiş; buna karşın Perplexity dalgali kalite göstermiş, DeepSeek ise yüzeysel ve yetersiz yanıtlar üretmiş ve en düşük performansı sergilemiştir. Bu bulgu, çalışanların YZ'ye yöneldiklerinde her zaman güvenilir ve yeterli bilgiye ulaşamayabileceklerini göstermektedir.

Buradan çıkarılacak en önemli sonuç, YZ'nin çalışanların bilgi eksikliklerini gidermede yardımcı olabileceği, ancak tek başına yeterli olmadığıdır. Özellikle düşük performans gösteren modellerin verdiği yüzeysel ya da eksik yanıtlar, çalışanların hatalı uygulamalara yönelmesine ve dolayısıyla hem kendi sağlıklarının hem de hasta güvenliğinin riske girmesine yol açabilir. Bu nedenle YZ araçlarının, çalışanların İSG eğitimlerini tamamlayıcı nitelikte, uzman denetimiyle ve kılavuz temelli doğrulamalarla desteklenerek kullanılması gereklidir.

SONUÇLAR

Sonuç olarak, çalışanların İSG konusundaki bilgi eksiklikleri göz önüne alındığında, YZ tabanlı sistemlerin bu boşluğu doldurmak için cazip bir seçenek olduğu açıktır. Ancak modeller arasında ciddi performans farklılıkları bulunduğundan, ChatGPT ve Gemini gibi daha yüksek doğruluk gösteren sistemler tercih edilmeli, fakat hiçbir zaman

tek başına karar verici veya rehber yerine kullanılmamalıdır. Gelecekte yapılacak çalışmalar, çalışanların İSG eğitimine YZ entegrasyonunun hangi yöntemlerle daha güvenli ve etkili olabileceğini araştırmalıdır.

Çalışmanın Güçlü Yönleri ve Sınırlılıkları

Bu çalışmanın en önemli güçlü yönü, ürolojik ameliyatlarda İSG gibi spesifik ve kritik bir alanda dört farklı YZ modelinin (ChatGPT, Gemini, Perplexity, DeepSeek) sistematik biçimde karşılaştırılmasıdır. Hem DISCERN hem de GQS skorlamalarının birlikte kullanılması, yanıtların yalnızca içerik kalitesini değil, aynı zamanda genel güvenilirlik ve pratik uygulanabilirliğini de ölçmeye imkân tanımıştır. Ayrıca, soruların literatür ve uluslararası kılavuzlar temel alınarak hazırlanması ve uzman görüşleriyle desteklenmesi, çalışmanın metodolojik sağlamlığını artırmıştır.

Bununla birlikte, çalışmanın bazı sınırlılıkları da bulunmaktadır. Öncelikle çalışma yalnızca ürolojik cerrahi ve ameliyathane İSG konularıyla sınırlıdır; bu nedenle elde edilen sonuçların diğer cerrahi branşlara veya farklı klinik alanlara genellenebilirliği kısıtlıdır. İkinci olarak, YZ modelleri sürekli güncellenmekte ve geliştirilmektedir. Bu nedenle çalışmanın bulguları, belirli bir zaman dilimindeki performansı yansıtmaktadır ve ilerleyen dönemlerde farklı sonuçlar elde edilebilir. Üçüncü olarak, değerlendirmeler uzman görüşlerine dayalı olup, her ne kadar körleme uygulanmış olsa da öznel unsurlar içerebilir. Son olarak, çalışanların gerçek klinik pratikte YZ yanıtlarını nasıl kullandıkları ve bu yanıtların karar verme süreçlerine etkisi bu çalışmada ölçülmemiştir.

Çıkar Çatışması: Yazarlar, bu çalışma ile ilgili olarak herhangi bir çıkar çatışması bulunmadığını beyan ederler.

Finansal Destek: Bu çalışma için herhangi bir finansal destek/fon alınmamıştır.

Etik Kurul: Bu çalışmada herhangi bir hasta verisi kullanılmadığı için etik kurul onayı gerekmemiştir.

Bilgilendirilmiş Onam: Hasta verilerine yer verilmediğinden, bilgilendirilmiş onam formu/izni gerekmemektedir.

Yazar Katkıları: Kavramsallaştırma, çalışma tasarımı ve taslak metnin yazımı: Ş.D. ve M.D. Her iki yazar da makalenin son halini incelemiş ve onaylamıştır.

KAYNAKLAR

1. Landford WN, Ngaage LM, Lee E, et al. Occupational exposures in the operating room: Are surgeons well-equipped? PLoS One. 2021;16(7):e0253785. <https://doi.org/10.1371/journal.pone.0253785>
2. Dixon K, Dasgupta P, Vasdev N. A systematic review of the harmful effects of surgical smoke inhalation on operating room personnel. Health Sci Rev. 2023;6:100077. <https://doi.org/10.1016/j.hsr.2022.100077>
3. Memon AG, Naeem Z, Zaman A, Zahid F. Occupational health related concerns among surgeons. Int J Health Sci (Qassim). 2016;10(2):279-91.
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. <https://doi.org/10.1016/j.isci.2024.109713>
5. Hager P, Jungmann F, Holland R, et al. Evaluation and mitigation of the limitations of large language models in clinical decision-making. Nat Med. 2024;30(9):2613-22. <https://doi.org/10.1038/s41591-024-03215-7>
6. Rossettini G, Barger S, Cook C, et al. Accuracy of ChatGPT-3.5, ChatGPT-4o, Copilot, Gemini, Claude, and Perplexity in advising on lumbosacral radicular pain against clinical practice guidelines: Cross-sectional study. Front Digit Health. 2025;7:1574287. <https://doi.org/10.3389/fgth.2025.1574287>

7. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health*. 1999;53(2):105-11. <https://doi.org/10.1136/jech.53.2.105>
8. Bromage S, Batis C, Bhupathiraju SN, et al. Development and validation of a novel food-based global diet quality score (GDQS). *J Nutr*. 2021;151(Suppl 1):75S-92S. <https://doi.org/10.1093/jn/nxab244>
9. van den Berg-Dijkmeijer ML, Frings-Dresen MH, Sluiter JK. Risks and health effects in operating room personnel. *Work*. 2011;39(3):331-44. <https://doi.org/10.3233/WOR-2011-1181>
10. Johnson D, Goodman R, Patrinely J, et al. Assessing the accuracy and reliability of AI-generated medical responses: An evaluation of the ChatGPT model. *Res Sq*. 2023;rs.3.rs-2566942. <https://doi.org/10.21203/rs.3.rs-2566942/v1>
11. Wang C, Liu S, Yang H, et al. Ethical considerations of using ChatGPT in health care. *J Med Internet Res*. 2023;25:e48009. <https://doi.org/10.2196/48009>