



Oftalmolojide Büyük Dil Modelleri: Bibliyografik Bir Analiz

Large Language Models in Ophthalmology: A Bibliographic Analysis

© Neslihan Dilruba Köseoğlu¹, © T.Y. Alvin Liu²

¹Tufts Tıp Merkezi, Oftalmoloji Kliniği, Boston, MA, ABD

²Wilmer Göz Enstitüsü, Tıp Fakültesi, Johns Hopkins Üniversitesi, Baltimore, MD, ABD

Öz

Bu çalışma, PubMed'de Kasım 2024'e kadar yayımlanmış göz hastalıkları alanında büyük dil modellerinin (BDM) kullanımı ile ilgili araştırmaların bibliyografik analizini amaçlamaktadır. Çalışmalar, klinik karar desteği (yandallara göre ayrılmış), eğitim, hasta etkileşimleri ve diğer uygulamalar olmak üzere BDM yöntemlerinin dört ana uygulama alanına göre kategorize edilmiştir. Tanımlayıcı istatistikler kullanılarak çalışmaların, oftalmoloji alt uzmanlıkları, coğrafi bölge, dergi kalitesi ve yazar özellikleri (cinsiyet ve bilimsel etki, h-endeksi ve i10-endeksi dahil) açısından dağılımı analiz edilmiştir. Bulgularımız, klinik karar desteğinin en yaygın uygulama olduğunu (%43,7) ve bu alt grup içerisinde çalışmaların çoğunun retina üzerine odaklandığını (%39,5) ortaya koymuştur. Coğrafi olarak, araştırmaların çoğu Kuzey Amerika'dan (%48,3) yayımlanmakta olup sırasıyla Asya (%29,9) ve Avrupa (%20,7) Kuzey Amerika'yı izlemektedir. Çoğu çalışma, özellikle de klinik karar desteği alt grubunda, retina (%80,0), glokom (%100) ve birden fazla alt uzmanlık dalını kapsayan çalışmalar (%87,5), yüksek etki faktörüne sahip dergilerde yayımlanmıştır (Q1 dergileri: %74,7). Cinsiyet dağılımındaki eşitsizlik, tüm yazar rollerinde belirgin olarak gözlemlenmiştir. Çalışmalarda kadın yazarlar, ilk yazarların yalnızca %29,9'unu, son yazarların %25,3'ünü ve sorumlu yazarların %26,4'ünü oluşturmuştur. Sonuçlar, göz hastalıkları alanındaki BDM araştırmalarında cinsiyet ve coğrafi temsilde daha fazla çeşitliliğe ihtiyaç duyulduğunu ve bu çeşitliliğin alandaki kapsayıcı ilerlemeyi teşvik edeceğini göstermektedir.

Anahtar Kelimeler: Büyük dil modelleri, oftalmoloji, bibliyografik analiz

Abstract

This study evaluated the distribution of research on the use of large language models (LLMs) in ophthalmology through a bibliographic analysis of articles retrieved from PubMed through November 2024. Studies were categorized into four main areas of LLM application: clinical decision-making (further divided according to subspecialties), education, patient interactions, and miscellaneous applications. Descriptive statistics were used to analyze the distribution of studies by ophthalmic subspecialty, geographical region, journal quality, and author characteristics, including gender and scholarly impact (h-index and i10-index). The findings revealed that clinical decision-making was the most common application (43.7%), with the majority of studies in this subgroup focusing on the retina (39.5%). Geographically, most of the research originated from North America (48.3%), followed by Asia (29.9%) and Europe (20.7%). Most studies were published in high-impact journals (Q1 journals: 74.7%), particularly for those related to clinical decision-making in retina (80.0%), glaucoma (100%), and multiple subspecialties (87.5%). Gender disparities were evident across all author roles, with female authors accounting for only 29.9% of first authors, 25.3% of last authors, and 26.4% of corresponding authors. The results suggest a need for greater diversity in terms of gender and geographic representation in LLM research in ophthalmology to promote inclusive progress in the field.

Keywords: Large language models, ophthalmology, bibliographical analysis

Cite this article as: Köseoğlu ND, Liu TYA. Large Language Models in Ophthalmology: A Bibliographic Analysis. Turk J Ophthalmol. 2026;56:119-130

Yazışma Adresi/Address for Correspondence: Neslihan Dilruba Köseoğlu, Tufts Tıp Merkezi, Oftalmoloji Kliniği, Boston, MA, ABD

E-posta: dilruba33@yahoo.com

ORCID-ID: orcid.org/0000-0001-9268-1461

Geliş Tarihi/Received: 16.04.2025

Revizyon Talebi/Revision Requested: 06.06.2025

Son Revizyon Alınma/Last Revision Received: 01.12.2025

Kabul Tarihi/Accepted: 01.01.2026

Yayın Tarihi/Publication Date: 27.04.2026

DOI: 10.4274/tjo.galenos.2026.79484

Giriş

Son yıllarda, yapay zeka (YZ), özellikle büyük dil modelleri (BDM) şeklinde, bilimin birçok alanında köklü bir dönüşüme yol açmış ve araştırma süreçlerinin vazgeçilmez bir parçası haline gelmiştir.¹ ChatGPT, BERT ve LLaMA gibi modeller; kitaplar, makaleler ve diğer metin tabanlı kaynaklardan oluşan çeşitli veri kümelerinden beslenen milyonlarca parametre aracılığıyla insan benzeri metin tabanlı yanıtlar üretmektedir.² Bu BDM'lerin mimarileri arasında farklılıklar mevcuttur. ChatGPT, üretken dil görevlerine yönelik otoregresif dönüştürücüler kullanırken; BERT, bağlamsal kavrama için çift yönlü dönüştürücülerden



Telif Hakkı © 2026 Yazar(lar). Türk Oftalmoloji Derneği adına Galenos Yayınevi tarafından yayımlanmıştır.

Bu, Creative Commons Atıf-GayriTicari-TürevleriYaratılamaz 4.0 (CC BY-NC-ND) Uluslararası Lisansı kapsamında açık erişimli bir makaledir.

yararlanır. LLaMA ise eğitim kararlılığını en üst düzeyde tutmak amacıyla uzmanlar karışımı (“mixture of experts”) modellerinden kaçınarak, yalnızca kod çözücülerden (“decoder-only”) oluşan ve küçük uyarlamalar içeren standart bir dönüştürücü mimarisine sahiptir.^{3,4,5} Bu farklılıklara rağmen, bu modellerin tamamı doğal dil işleme görevlerinde yaygın biçimde kullanılmaktadır.

YZ tıpta onlarca yıldır kullanılmaktadır. BDM’ler günümüzde görüntü analizi yoluyla tanı süreçlerini dönüştürmekte; büyük hacimli verileri, klinik bilgileri ve hasta kayıtlarını hızlı ve etkin biçimde işleyerek klinik karar destek süreçlerini güçlendirmekte ve böylece kişiselleştirilmiş tıbbin önünü açmaktadır.⁶ BDM’ler son yıllarda radyoloji, dahiliye, pediatri, kardiyovasküler tıp ve diğer çeşitli tıbbi uzmanlık alanlarında giderek daha fazla benimsenmiştir.^{7,8,9} BDM’lere yönelik araştırmalar 2023 yılından itibaren belirgin biçimde çoğalmış olup bu durum, söz konusu modellerin yeteneklerine ve uygulama alanlarına duyulan ilginin giderek arttığını göstermektedir. Oftalmoloji, hasta bakımını geliştirmeye yönelik BDM kullanımına olan artan ilgiyle birlikte, YZ odaklı araştırmaların öncü alanları arasında yer almaktadır.¹⁰

Ancak, BDM ile ilgili araştırmalar da çeşitli kısıtlılıklar ve zorluklarla karşı karşıyadır. Öncelikle, BDM’ler büyük ölçüde İngilizce verilerle eğitilmektedir; bu durum, söz konusu modellerin İngilizce dışındaki dillerde kullanılması halinde ürettikleri yanıtların güvenilirliğini olumsuz yönde etkileyebilir. Bunun sonucu olarak, araştırmalar genellikle İngilizce’nin yaygın olarak konuşulduğu bölgelerde yoğunlaşmaktadır. Tüm YZ çalışmalarında bölgesel eşitsizlikler açıkça görülmekte olup yapılan yayınlar ve verilen araştırma desteklerinde ABD ve Çin’in üstünlüğü tartışılmazdır; 2014-2023 yılları arasında bu alanda yapılan tüm yayınların yaklaşık %50’sinin bu iki ülkeden geldiği gösterilmiştir.¹¹ Ayrıca, YZ Endeksi Raporu (2024) da 2023 yılında geliştirilen özgün YZ modellerinin çoğunluğunun başta ABD olmak üzere Çin ve Avrupa kaynaklı olduğunu ortaya koymuştur. Bu bulgu, YZ araştırmalarındaki coğrafi uçurumu bir kez daha gözler önüne sermektedir.¹² Araştırma çıktısındaki bu eşitsiz dağılıma, cinsiyet temsili gibi diğer eşitsizlikler de eşlik etmektedir. Dünya Ekonomik Forumu’nun yayımladığı en güncel Küresel Cinsiyet Farkı Raporu’na (2024) göre YZ alanındaki profesyonellerin yalnızca %22’si kadındır; bu cinsiyet eşitsizliği, YZ odaklı araştırma çıktılarına da yansımaktadır.¹³

YZ araştırmalarındaki bu bilinen eşitsizlikler ışığında, çalışmamızda oftalmoloji alanındaki BDM’ye ilişkin çalışmaların daha ayrıntılı biçimde incelenmesi amaçlanmıştır. Oftalmoloji alanındaki BDM araştırmalarının dağılımını cinsiyet, akademik etki ve çalışmaların coğrafi kökeni açısından değerlendirmek amacıyla bibliyografik bir analiz gerçekleştirmeyi amaçladık.

Yöntem

Kasım 2024 tarihine kadar yayımlanmış, oftalmoloji alanında BDM kullanımını konu alan çalışmaları belirlemek amacıyla PubMed veri tabanında kapsamlı bir literatür taraması yapıldı. Tarama yapılırken arama terimleri olarak *büyük dil modelleri* VE (*oftalmoloji* VEYA *retina* VEYA *glokom* VEYA *kornea* VEYA *uvea* VEYA *pediatrik oftalmoloji* VEYA *nöro-oftalmoloji*) kullanıldı (*large language models* AND [*ophthalmology* OR *retina* OR *glaucoma* OR *cornea* OR *uvea* OR *pediatric ophthalmology* OR *neuro-ophthalmology*]). Araştırma sonucunda toplam 194 çalışma elde edildi. Nihai analize yalnızca özgün araştırma makaleleri dahil edildi. BDM’ler veya oftalmoloji ile doğrudan ilişkili olmayan çalışmalar, derleme makaleleri, meta-analizler ve editöre mektup/yorum niteliğindeki yazılar nihai listeden çıkarıldı.

Dahil edilen çalışmalar, primer odak noktalarına ve uygulamalarına göre dört gruba ayrıldı:

- Klinik Karar Desteği Uygulamaları: BDM’lerin klinik karar destek süreçlerindeki kullanımını araştıran; göz hastalıklarının tanısı, tedavisi veya klinik desteklenmesinde BDM yanıtlarını değerlendiren çalışmalar. Bu kategori retina, glokom, kornea/ön segment, üveit, nöro-oftalmoloji ve pediatrik oftalmoloji alt uzmanlık dallarına göre ayrıca sınıflandırılmıştır. Birden fazla oftalmoloji alt dalını veya genel oftalmolojiyi kapsayan çalışmalar “çoklu alt uzmanlık” kategorisine alınmıştır.

- Eğitim Uygulamaları: BDM’lerin eğitim amaçlı kullanımını inceleyen çalışmalar; uzmanlık sınavı sorusu formatındaki soruların yanıtlanması ve hem hastalar hem de hekimler için eğitim materyali geliştirilmesi uygulamalarını kapsamaktadır.

- Hasta Etkileşimi Uygulamaları: BDM’lerin hasta iletişimini geliştirmek ve bir sohbet robotu formatında olduğu gibi göz sağlığı hakkında sıkça sorulan sorulara yanıt vermek için nasıl kullanıldığına odaklanan çalışmalar.

- Diğer Uygulamalar: Oftalmoloji alanında BDM’lerin diğer kategorilere tam olarak dahil edilemeyen, özellikle ileri düzey teknik nitelik taşıyan diğer uygulama alanlarını araştıran çalışmalar.

Çalışmaların ilk kategorizasyonu birinci yazar tarafından yapıldı ve kıdemli yazar tarafından gözden geçirildi. Görüş ayrılıkları, tartışma ve uzlaşa yoluyla giderildi.

Çalışmaların uygulama türü, oftalmoloji alt uzmanlık dalı, coğrafi bölge, dergi sıralaması ve yazar özellikleri açısından dağılımını incelemek için tanımlayıcı istatistiksel yöntemler kullanıldı. Coğrafi bölgeler, dünyayı altı kıta bölgesine bölen Birleşmiş Milletler Dünya Nüfusu Tahminleri’ne göre Afrika, Asya, Avrupa, Latin Amerika ve Karayipler, Kuzey Amerika ve Okyanusya olarak belirlendi.¹⁴

Coğrafi bölge, birinci yazarın bağlı olduğu kuruma göre atandı. Dergi sıralaması, SCImago veri tabanından elde edilen Q1 (birinci çeyrek), Q2, Q3 ve Q4 çeyreklik verilerine göre yapıldı. Yazar özellikleri kapsamında cinsiyet ve akademik etki incelendi. Cinsiyet bilgileri kurumsal web siteleri veya mesleki profillerden elde edildi. Akademik etki, her ikisi de Google Akademik'ten elde edilen h-endeksi ve i10-endeksi kullanılarak değerlendirildi. H-endeksi, bir yazarın yayınlarının bilimsel etkisini ölçen bir göstergedir. Yazarın en az n kez atıf almış n adet makalesi bulunduğu anlamına gelmektedir ve en yüksek n değeri yazarın h-endeksini oluşturur.¹⁵ i10-endeksi ise Google Akademik tarafından kullanılan ve en az 10 atıf almış yayınların toplam sayısını ölçen bir metriktir.¹⁶

Bulgular

Genel Bulgular

Toplam 87 özgün araştırma makalesi son değerlendirmeye dahil edildi. Uygulama türü açısından değerlendirildiğinde, en yaygın kategori klinik karar desteği (n=38, %43,7) olmuş; bunu sırasıyla eğitim uygulamaları (n=22, %25,3), hasta etkileşim uygulamaları (n=18, %20,7) ve diğer uygulamalar (n=9, %10,3) izlemiştir (Tablo 1). Klinik karar desteği kategorisinde retinaya ilişkin çalışmalar 15 çalışmayla (%39,5) en fazla paya sahip olurken, pediatrik oftalmoloji yalnızca 1 çalışmayla (%2,6) en az temsil edilen alt dal olmuştur. Dikkat çekici biçimde, birden fazla alt uzmanlık dalını birlikte ele alan çalışmalar bu grubun %21,1'ini (n=8) oluşturmuştur. Glukom, kornea/ön segment, üveit ve nöro-oftalmoloji alanları ise orta düzeyde katkı sağlamıştır. Her kategorideki çalışma sayıları Tablo 2'de sunulmuştur.

Birinci yazarın bağlı olduğu kuruma göre yapılan coğrafi dağılım analizi, araştırmaların büyük çoğunluğunun Kuzey Amerika'da (n=42, %48,3) yapıldığını ortaya koymuş; bunu Asya (n=26, %29,9) ve Avrupa (n=18, %20,7) izlemiştir. Okyanusya, Avustralya'dan yalnızca tek bir çalışmayla (%1,1) temsil edilirken; Latin Amerika ve Karayipler ile Afrika'dan oftalmoloji alanında BDM'lere ilişkin herhangi bir araştırma bildirilmemiştir. Çalışmaların bölgelere göre dağılımı Tablo 3 ve Şekil 1'de sunulmuştur.

Tablo 1. Çalışmaların uygulama türlerine göre dağılımı (n=87)

Kategori	Çalışma sayısı	Yüzde
Klinik karar desteği	38	%43,7
Eğitim uygulamaları	22	%25,3
Hasta etkileşimi uygulamaları	18	%20,7
Diğer uygulamalar	9	%10,3

Çalışmaların büyük çoğunluğu yüksek etkili dergilerde yayımlanmış olup 65 çalışma (%74,7) Q1, 10 çalışma (%11,5) Q2 ve 9 çalışma (%10,3) Q3 dergilerinde yer almıştır. Ayrıca, 4 çalışma (%4,6) için herhangi bir çeyreklik dilim bilgisi verilmemiş olup bunların 3'ü klinik karar verme alanındaki baskı öncesi yayınlardan (pre-print) oluşmaktadır. Çalışmaların büyük bölümü 2024 yılında yayımlanmış (n=71, %81,6) olup bunu 2023 yılında yayımlanan 15 çalışma (%17,2) ve 2022 yılında yayımlanan 1 çalışma (%1,1) izlemiştir.

Cinsiyet eşitsizliği tüm yazar rolleri için belirgin biçimde gözlemlenmiş olup kadınlar; birinci yazarların yalnızca

Tablo 2. Klinik karar desteği kategorisinde yer alan çalışmaların dağılımı (n=38)

Klinik karar desteği	Çalışma sayısı	Yüzde
Retina	15	%39,5
Glukom	6	%15,8
Kornea/ön segment	3	%7,9
Üveit	3	%7,9
Nöro-oftalmoloji	2	%5,3
Pediatri	1	%2,6
Çoklu alt uzmanlık	8	%21,1

Tablo 3. Çalışmaların ilk yazarın bağlı olduğu kuruma göre coğrafi dağılımı (n=87)

Bölge/kıta	Ülke	Çalışma sayısı	Yüzde
Kuzey Amerika	ABD	38	%43,7
	Kanada	4	%4,6
Avrupa	Birleşik Krallık	9	%10,3
	Almanya	3	%3,5
	İsviçre	2	%2,3
	İtalya	2	%2,3
	Romanya	1	%1,1
	Finlandiya	1	%1,1
Asya	Çin	12	%13,8
	Singapur	4	%4,6
	Türkiye	4	%4,6
	Hindistan	3	%3,5
	Japonya	1	%1,1
	İsrail	1	%1,1
	Ürdün	1	%1,1
Okyanusya	Avustralya	1	%1,1
Afrika		0	%0
Latin Amerika ve Karayipler		0	%0

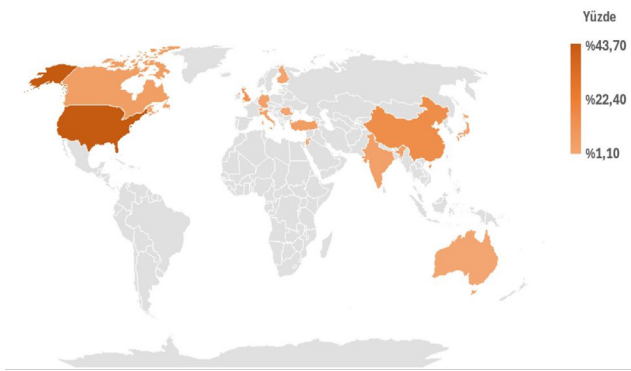
ABD: Amerika Birleşik Devletleri

%29,9'unu, son yazarların %25,3'ünü ve sorumlu yazarların %26,4'ünü oluşturmuştur (Şekil 2). Beklendiği üzere, son yazarlar ve sorumlu yazarlar genellikle birinci yazarlara kıyasla daha yüksek akademik etkiye sahipti. Son yazarlar ve sorumlu yazarlar için ortalama h-endeksi sırasıyla $39,7 \pm 37,7$ ve $28,7 \pm 33,7$ olarak bulunurken, birinci yazarlar için bu değer $8,9 \pm 6,2$ idi ($p < 0,001$). Benzer biçimde, son yazarlar ve sorumlu yazarlar için ortalama i10-endeksi sırasıyla $147,1 \pm 254,2$ ve $94,4 \pm 222,2$ iken, birinci yazarlar için $12,0 \pm 14,9$ olarak saptandı ($p = 0,005$) (Tablo 4).

Ayrıca yayınları her uygulama kategorisine göre değerlendirdik ve bu detaylı analizin sonucunda elde edilen bulgular şu şekildedir:

Klinik Karar Desteği

Bu kategorideki çalışmalar birçok kıtada yer alan çeşitli ülkelerden yayımlanmış olmakla birlikte, büyük çoğunluğu yine de ABD kaynaklıdır ($n = 17$, %44,7). Çalışmaların büyük çoğunluğu 2024 yılında yayımlanmıştır ($n = 32$, %84,2). Cinsiyet eşitsizliği tüm alt gruplarda belirgin biçimde gözlemlendi; bu durum özellikle kornea/ön segment alt dalında dikkat çekiciydi; bu alanda birinci yazar, son yazar veya sorumlu yazar olarak hiç kadın araştırmacı yoktu (Tablo 5).

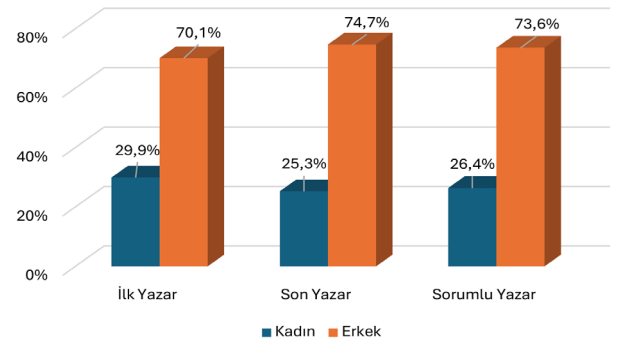


Şekil 1. Oftalmolojide BDM ile ilişkili çalışmaların küresel ısı haritası
BDM: Büyük dil modelleri

Çalışmaların büyük çoğunluğu yüksek etkili dergilerde yayımlanmıştır. Özellikle retina, glokom ve çoklu alt uzmanlık gruplarında Q1 dergi oranları sırasıyla %80,0, %100 ve %87,5 bulundu. Kornea/ön segment ve üveit alt grupları ise Q1, Q2 ve/veya Q3 dergilerde karma bir dağılım gösterdi. Öte yandan, Q4 dergilerinde yayımlanan herhangi bir çalışma yoktu. Bu kategorideki çalışmaların ayrıntılı bir listesi ek materyalde verilmiştir (Ek Tablo 1).¹⁷⁻⁵⁴

Eğitim Uygulamaları

Yirmi iki çalışmanın %45,5'i ile en yüksek paya sahip olan Kuzey Amerika'nın bu katkısı, ABD'den 8 ve Kanadadan 2 çalışmadan kaynaklanmıştır. Avrupadan yayımlanan çalışmalar toplamın %31,8'ini oluşturup bunların 6'sı Birleşik Krallık'tan, 1'i ise Almanya'dandır. Asya 5 çalışmayla (%22,7) katkı sağlamış olup bu çalışmaların 2'si Çin'den, 1'er tanesi ise Türkiye, İsrail ve Japonya'dandır. Bu çalışmaların büyük çoğunluğu Q1 dergilerinde (%72,7) yayımlanmıştır. Bunu Q3 (%18,2) ve Q2 (%9,1) dergileri izlemiştir. Çalışmaların büyük bölümü 2024 yılına aittir ($n = 18$, %81,8). Bu grupta da cinsiyet eşitsizliği belirgin biçimde gözlemlenmiştir. Tüm yazar rollerinde erkek yazarların ağırlıklı konumda olduğu saptanmıştır. Kadın araştırmacılar 4 birinci yazar (%18,2), 1 son yazar (%4,5) ve 3 sorumlu yazar (%13,6) ile temsil edilmiştir.



Şekil 2. Yazarlık rolleri arasında cinsiyet dağılımı (yüzde olarak sunulmuştur)

Tablo 4. Cinsiyet ve akademik etki göstergelerinin (h-endeksi ve i10-endeksi) birinci yazar, son yazar ve sorumlu yazar rollerine göre dağılımı

Yazar rolü	Kadın	Erkek	h-endeksi (Ortalama ± SD)	i10-endeksi (Ortalama ± SD)
İlk yazar	26 (%29,9)	61 (%70,1)	$8,9 \pm 6,2$	$12,0 \pm 14,9$
Son yazar	22 (%25,3)	65 (%74,7)	$39,7 \pm 37,7$	$147,1 \pm 254,2$
Sorumlu yazar	23 (%26,4)	64 (%73,6)	$28,7 \pm 33,7$	$94,4 \pm 222,2$

SD: Standart deviasyon

Tablo 5. Klinik karar desteği uygulamaları çalışmalarının birinci yazar ve son yazarın bağlı olduğu kuruma göre coğrafi bölge ve cinsiyet dağılımı						
	İlk yazar			Son yazar		
	Cinsiyet	Coğrafi bölge	En çok yayın yapan 3 ülke	Cinsiyet	Coğrafi bölge	En çok yayın yapan 3 ülke
Retina ^{17,31} (n=15)	Kadın: 4 (%26,7) Erkek: 11 (%73,3)	Asya: 6 (%40,0) Avrupa: 5 (%33,3) Kuzey Amerika: 4 (%26,7) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	Çin: 5 (%33,3) Birleşik Krallık: 3 (%20,0) ABD: 3 (%20,0)	Kadın: 3 (%20,0) Erkek: 12 (%80,0)	Asya: 6 (%40,0) Avrupa: 5 (%33,3) Kuzey Amerika: 4 (%26,7) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	Çin: 5 (%33,3) Birleşik Krallık: 4 (%26,7) ABD: 3 (%20,0)
Glökom ^{32,33,34,35,36,37} (n=6)	Kadın: 2 (%33,3) Erkek: 4 (%66,7)	Kuzey Amerika: 4 (%66,7) Asya: 1 (%16,7) Avrupa: 1 (%16,7) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 4 (%66,7) Çin: 1 (%16,7) İtalya: 1 (%16,7)	Kadın: 0 (%0) Erkek: 6 (%100)	Kuzey Amerika: 4 (%66,7) Asya: 1 (%16,7) Avrupa: 1 (%16,7) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 4 (%66,7) Çin: 1 (%16,7) İtalya: 1 (%16,7)
Kornea/ön segment ^{38,39,40} (n=3)	Kadın: 0 (%0) Erkek: 3 (%100)	Kuzey Amerika: 2 (%66,7) Avrupa: 1 (%33,3) Asya: 0 (%0) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 2 (%66,7) Almanya: 1 (%33,3)	Kadın: 0 (%0) Erkek: 3 (%100)	Kuzey Amerika: 2 (%66,7) Avrupa: 1 (%33,3) Asya: 0 (%0) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 2 (%66,7) Almanya: 1 (%33,3)
Üveit ^{41,42,43} (n=3)	Kadın: 2 (%66,7) Erkek: 1 (%33,3)	Asya: 1 (%33,3) Avrupa: 1 (%33,3) Kuzey Amerika: 1 (%33,3) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	Singapur: 1 (%33,3) İsviçre: 1 (%33,3) ABD: 1 (%33,3)	Kadın: 2 (%66,7) Erkek: 1 (%33,3)	Asya: 1 (%33,3) Avrupa: 1 (%33,3) Kuzey Amerika: 1 (%33,3) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	Singapur: 1 (%33,3) Birleşik Krallık: 1 (%33,3) ABD: 1 (%33,3)
Nöro-oftalmoloji ^{44,45} (n=2)	Kadın: 1 (%50) Erkek: 1 (%50)	Kuzey Amerika: 2 (%100) Asya: 0 (%0) Avrupa: 0 (%0) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 2 (%100)	Kadın: 0 (%0) Erkek: 2 (%100)	Kuzey Amerika: 2 (%100) Asya: 0 (%0) Avrupa: 0 (%0) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 2 (%100)

Tablo 5. Devamı	İlk yazar			Son yazar		
	Cinsiyet	Coğrafi bölge	En çok yayın yapan 3 ülke	Cinsiyet	Coğrafi bölge	En çok yayın yapan 3 ülke
Pediatriden ⁶ (n=1)	Kadın: 0 (%0) Erkek: 1 (%100)	Kuzey Amerika: 1 (%100) Asya: 0 (%0) Avrupa: 0 (%0) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 1 (%100)	Kadın: 0 (%0) Erkek: 1 (%100)	Kuzey Amerika: 1 (%100) Asya: 0 (%0) Avrupa: 0 (%0) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 1 (%100)
Çoklu alt uzmanlık ^{47,48,49,50,51,52,53,54} (n=8)	Kadın: 1 (%12,5) Erkek: 7 (%87,5)	Kuzey Amerika: 5 (%62,5) Asya: 3 (%37,5) Avrupa: 0 (%0) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 4 (%50) Çin: 3 (%37,5) Kanada: 1 (%12,5)	Kadın: 3 (%37,5) Erkek: 5 (%62,5)	Kuzey Amerika: 5 (%62,5) Asya: 3 (%37,5) Avrupa: 0 (%0) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 4 (%50) Çin: 3 (%37,5) Kanada: 1 (%12,5)

Hasta Etkileşimi Uygulamaları

Bu gruptaki çalışmalar ağırlıklı olarak Asya (%44,4) kaynaklı olup bunu sırasıyla Kuzey Amerika (%38,9), Avrupa (%11,1) ve Okyanusya (%5,6) izlemiştir. Çalışmaların büyük bir kısmı Q1 dergilerinde (%66,7), küçük bir kısmı ise Q2 ve Q3 dergilerinde yayımlanmıştır. Kadın yazarlar genel olarak bu kategoride daha fazla temsil edilmiş, ilk yazarların %33,3'ünü, son yazarların %5,6'sını ve sorumlu yazarların %44,4'ünü oluşturmuştur.

Diğer Uygulamalar

Dokuz çalışma arasında %66,7 ile en büyük katkısı Kuzey Amerika sağlamış olup, bu 6 çalışma ABD'de gerçekleştirilmiştir. Türkiye, Hindistan ve Finlandiya, birer çalışma ile katkıda bulunmuştur. Birinci yazar ve sorumlu yazar rolleri açısından cinsiyet dağılımı daha dengeli bir görünüm sergilemiştir. Birinci yazarların %44,4'ü erkek ve %55,6'sı kadın iken sorumlu yazarların %55,6'sı erkek ve %44,4'ü kadındır. Bununla birlikte, son yazarlar ağırlıklı olarak erkeklerdir (%66,7).

Birinci yazar, son yazar ve sorumlu yazarlara ait cinsiyet ile coğrafi bölge dağılımları; eğitim uygulamaları, hasta etkileşimi uygulamaları ve diğer uygulamalar kategorileri [Tablo 6](#)'da ayrıntılı biçimde sunulmuştur. Diğer kategorilerdeki çalışmaların kapsamlı listesi ve genel dergi metrikleri ek materyalde verilmiştir ([Ek Tablo 2](#) ve [3](#)).⁵⁵⁻¹⁰³

Tartışma

BDM'ler, insan dilini ve konuşmasını kavramak amacıyla doğal dil işleme yöntemlerini kullanan üretken YZ sistemleridir.^{104,105} ChatGPT, BERT ve LLaMA; kullanıcılarla anlamlı diyaloglar kurabilmek için derin öğrenme (DÖ) tekniklerinden yararlanan, yaygın biçimde kullanılan BDM'lerdir.¹⁰⁴ Bu uygulamalar büyük miktarda bilgiye erişim sağlamanın yanı sıra hastaların özel durumlarıyla ilgili tıbbi bilgilere erişmelerine, semptomlarının aciliyetini değerlendirmelerine veya uygun alt uzmanlığa yönlendirilmelerine yardımcı olabilir.¹⁰⁶

YZ uygulamalarının tıp alanındaki kapsamı genişledikçe, bu teknolojilerin oftalmolojide kullanımına yönelik ilgi de giderek artmaktadır. Geçtiğimiz yıl içinde oftalmoloji alanında klinik karar desteğinden hasta iletişiminin geliştirilmesine kadar uzanan geniş bir yelpazede BDM kullanımının belirgin biçimde arttığı gözlemlenmiştir. Çalışmamızın amacı, oftalmoloji alanındaki BDM araştırmalarının güncel durumunu bibliyografik bir analiz aracılığıyla ortaya koymaktır.

Oftalmoloji alanındaki BDM çalışmalarının büyük çoğunluğunun klinik karar desteğine, bunların da ağırlıklı olarak retinal uygulamalara odaklandığını gözlemledik.

Tablo 6. Hasta etkileşimi, eğitim ve diğer uygulamalarla ilgili çalışmaların birinci ve son yazarlarının bağlı olduğu kuruma göre coğrafi bölge ve cinsiyet dağılımı	İlk yazar			Son yazar		
	Cinsiyet	Coğrafi bölge	En çok yayın yapan 3 ülke	Cinsiyet	Coğrafi bölge	En çok yayın yapan 3 ülke
	Eğitim uygulamaları ⁵⁵⁻⁷⁶	Kadın: 4 (%18,2) Erkek: 18 (%81,8)	Kuzey Amerika: 10 (%45,5) Avrupa: 7 (%31,8) Asya: 5 (%22,7) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 8 (%36,4) Birleşik Krallık: 6 (%27,3) Kanada: 2 (%9,1) Çin: 2 (%9,1)	Kadın: 1 (%4,5) Erkek: 21 (%95,5)	Kuzey Amerika: 12 (%54,5) Avrupa: 5 (%22,7) Asya: 5 (%22,7) Afrika: 0 (%0) Latin Amerika/ Karayipler: 0 (%0) Okyanusya: 0 (%0)
Hasta etkileşimi uygulamaları ⁷⁷⁻⁹⁴	Kadın: 6 (%33,3) Erkek: 12 (%66,7)	Asya: 8 (%44,4) Kuzey Amerika: 7 (%38,9) Avrupa: 2 (%11,1) Okyanusya: 1 (%5,6) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0)	ABD: 7 (%38,9) Singapur: 3 (%16,7) Türkiye: 2 (%11,1)	Kadın: 8 (%55,6) Erkek: 10 (%44,4)	Asya: 8 (%44,4) Kuzey Amerika: 7 (%38,9) Avrupa: 2 (%11,1) Okyanusya: 1 (%5,6) Afrika: 0 (%0) Latin Amerika/ Karayipler: 0 (%0)	ABD: 7 (%38,9) Singapur: 3 (%16,7) Türkiye: 2 (%11,1)
Diğer uygulamalar ^{95,96,97,98,99,100,101,102,103}	Kadın: 5 (%55,6) Erkek: 4 (%44,4)	Kuzey Amerika: 6 (%66,7) Asya: 2 (%22,2) Avrupa: 1 (%11,1) Afrika: 0 (%0) Latin Amerika/Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 6 (%66,7) Hindistan: 1 (%11,1) Türkiye: 1 (%11,1)	Kadın: 3 (%33,3) Erkek: 6 (%66,7)	Kuzey Amerika: 6 (%66,7) Asya: 2 (%22,2) Avrupa: 1 (%11,1) Afrika: 0 (%0) Latin Amerika/ Karayipler: 0 (%0) Okyanusya: 0 (%0)	ABD: 6 (%66,7) Birleşik Krallık: 1 (%11,1) Türkiye: 1 (%11,1)

BDM'lerin bu karmaşık tanı ve tedavi kararlarına destek sağlama konusundaki umut verici potansiyeli göz önüne alındığında, klinik karar desteği uygulamalarına duyulan bu ilgi şaşırtıcı değildir. Bununla birlikte, BDM'lerin düzenleyici mevzuat çerçevesinde klinik hasta bakımına nasıl entegre edileceğine ilişkin önemli belirsizlikler devam etmektedir.¹⁰⁷ Retina ile ilgili BDM yayınlarının sayısının yüksek olmasının nedenleri net olmamakla birlikte, oftalmolojide DÖ uygulamalarının geçmişinden kaynaklanabileceğini düşünmekteyiz. Oftalmoloji alanında DÖ temelli görüntü analizlerinde retina alt uzmanlığı öncü rol üstlenmiştir. Dolayısıyla, retina görüntülerinin DÖ temelli analizine odaklanan araştırmacı grubunun BDM ile ilgili çalışmaları da ilk benimseyen kesim olması mümkündür.¹⁰⁸

Benzer şekilde, araştırma ekibimiz son zamanlarda BDM'lere odaklanmıştır. Örneğin; BDM'lerin uzmanlık sınavı formatındaki soruları yanıtlamaktaki doğruluğunu değerlendiren bir meta-analiz yayımladık.⁵⁵ Ayrıca, erişim destekli üretim (*retrieval-augmented generation*) yönteminin entegrasyonunun hem metin tabanlı hem de görüntü tabanlı uzmanlık sınavı sorularında BDM performansına katkısını da değerlendirdik.^{109,110} Bir diğer ilgi alanı, oftalmolojide sağlığın sosyal belirleyicileri ile ilgili soruları yanıtlamada farklı BDM'lerin performans farkı olmuştur.¹¹¹

BDM'lere ilişkin çalışmaların üst düzey dergilerde giderek daha fazla yer bulması, bu araştırmaların bilimsel niteliğini destekler niteliktedir ve çalışmamızın bulgularıyla da örtüşmektedir. Bu çalışmalar önemli bir akademik etkiye sahip olup kayda değer bir bölümü en üst düzey Q1 dergilerinde

yayımlanmıştır. Glukomla ilişkili uygulamaların klinik karar desteği alanında en yüksek akademik etkiye sahip olduğu saptanmıştır. Bu bulgu, söz konusu çalışmaların yazarlarının yüksek h-endeksi ve i10-endeksi değerlerine sahip olması ve çalışmaların Q1 dergilerde yayımlanması ile de desteklenmektedir. Bununla birlikte, genel değerlendirmede en yüksek akademik etkiyi sergileyen uygulamaların eğitim amaçlı kullanılanlar olduğu görülmüştür. Bu uygulamalar arasında BDM'lerin klinik bilgi sınavlarındaki başarısını uzmanlarla karşılaştıran ve eğitim materyali geliştirmeye yönelik çalışmalar öne çıkmaktadır. Bu çalışmaların ortalama h-endeksi ve i10-endeksi değerleri daha yüksek bulunmuş ve büyük çoğunluğu Q1 dergilerinde yayımlanmıştır. Eğitim uygulamalarının özellikle yüksek etki sergilemesi, bu uygulamaların oftalmoloji dışındaki tıp disiplinlerine de uyarlanabilir niteliği, eğitim çıktıları geliştirme potansiyeli ve tıp eğitimini iyileştirmedeki etkinliğiyle açıklanabilir.

BDM'lerin araştırma gündemindeki yeri giderek güçlenmesine karşın, kadın yazarların yeterli temsil edilmemesi önemli bir sorundur. Yakın tarihli bir çalışma, 2018'den bu yana araştırmalardaki kadın oranının birinci yazarlarda %33'ten %37'ye, son yazarlarda ise %27'den %30'a yükseldiğini ortaya koymuştur.^{112,113} Ancak, cinsiyet eşitsizliği, özellikle YZ alanında önemli bir sorun olmaya devam etmektedir. Literatürle benzer biçimde, çalışmamız tüm yazar rollerinde belirgin bir cinsiyet eşitsizliği olduğunu göstermiştir. Kadınlar birinci yazarların yalnızca %29,9'unu, son yazarların %25,3'ünü ve sorumlu yazarların %26,4'ünü oluşturmuştur. Çalışmamızdaki bu cinsiyet dağılımı, akademik yayıncılıkta kadın temsilini artırmaya yönelik son çabalara rağmen, kilit yazar rollerine ulaşmada kadınlar için halen önemli engeller olduğunu göstermektedir. Bu durum, toplumsal cinsiyet eşitliğini güçlendirmeye yönelik çabaların kesintisiz sürdürülmesi gerektiğini bir kez daha ortaya koymaktadır. Kadın yazarların dikkate değer derecede az temsil edilmesi, olasılıkla bilim, teknoloji, mühendislik ve matematik eğitimi ve liderlik rollerindeki cinsiyet eşitsizlikleri, YZ odaklı tıbbi araştırmalarda kadınlar için mentorluk fırsatlarının sınırlı olması ve olası kurumsal yanlılık gibi daha geniş sistemik sorunları yansıtmaktadır.¹¹⁴

Çalışmaların coğrafi dağılımında da benzer bir eşitsizlik görülmektedir. Araştırmalar, önde gelen tıp dergilerindeki yazarlığın başta ABD (%48,2) olmak üzere İngilizce konuşulan ülkeler tarafından domine edildiğini göstermektedir. Buna karşılık, son yıllarda coğrafi çeşitlilikte bir artış olmasına rağmen, gelişmekte olan ülkelere yazarlar yeterince temsil edilmemektedir.¹¹⁵ Yakın zamanda yapılan bir başka çalışmada, dergilere gönderilen araştırmaların reddedilmesinde benzer bir yanlılık olduğu bildirilmiştir. Batılı olmayan ülkelerdeki

kurumlardan gelen yazarlarla karşılaştırıldığında, Batılı ülkelere yazarların reddedilen makalelerinin daha sonra kabul edilme olasılığı %5,7 daha yüksektir. Bunun yanı sıra, Batılı ülkelere yazarlar makalelerini 23 gün daha hızlı yayımlamakta, %5,9 daha az revizyon yapmakta, ortak yazarlarını %12,0 daha az değiştirmekte ve sonuç itibarıyla etki faktörü %0,8 daha yüksek dergilerde yayın yapmaktadır.^{11,116}

Çalışmamızda da literatürle uyumlu olarak ilk yazara göre belirlenen araştırmaların coğrafi dağılımında benzer bir farklılık olduğu görülmüştür. Genel olarak bakıldığında, İngilizce konuşulan ülkeler araştırma çıktısında baskın konumda yer almış, çalışmaların %43,7'si ABD ve %10,3'ü Birleşik Krallık'ta yapılmıştır. Bununla birlikte, Çin %13,8'lik katkısıyla ABD'nin hemen ardından ikinci konumda yer almış ve alana sağladığı önemli katkıyı bir kez daha ortaya koymuştur. Bu bulgular, oftalmoloji alanındaki BDM araştırmalarında ABD ve Çin'in öncü konumunu sürdürdüğünü ancak özellikle farklı bölgelerin kültürel ve toplumsal gereksinimlerine özgü uygulamaların geliştirilmesi açısından diğer bölgelerden yapılacak katkılara halen ihtiyaç olduğunu göstermektedir. Bu dengesizlik, BDM'lerin eğitiminde kullanılan verilerin ağırlıklı olarak İngilizce olmasıyla kısmen açıklanabilir.¹¹⁷ Ayrıca, ABD ve Çin'de temel YZ altyapısına yapılan büyük ölçekli yatırımlar, araştırma finansmanına erişim ve ileri teknolojilerden yararlanabilme olanakları da BDM araştırmalarının küresel dağılımını belirleyen kritik etkenler arasında sayılabilir.¹¹⁸

Çalışmanın Kısıtlılıkları

Bu çalışmanın bazı kısıtlılıkları vardır. Bunların başında literatür taramasının yalnızca PubMed veri tabanı ile sınırlı tutulması gelmektedir. arXiv ve IEEE Xplore gibi diğer veri tabanları YZ alanında önemli sayıda çalışmaya ev sahipliği yapsa da hakemli içeriğe yönelik yüksek kalite standartları nedeniyle yalnızca PubMed'de endekslenmiş çalışmalar dahil edilmiştir. Bu tercih, analize alınan makalelerin bilimsel niteliğinden emin olmak amacıyla bilinçli olarak yapılmıştır. Öte yandan, YZ altyapısı veya finansman olanaklarına göre yapılacak bir sınıflandırmanın BDM araştırmalarının dağılımına farklı bir perspektif kazandırabileceğini ve bu yaklaşımın gelecekteki çalışmalarda ele alınması gerektiğini düşünüyoruz.

Sonuç

Oftalmoloji alanında BDM kullanımı hızla ilgi odağı haline gelmekte olup bu alandaki çalışmaların büyük çoğunluğu 2024 yılında ve üst düzey Q1 dergilerinde yayımlanmıştır. Yayın sayısı açısından Kuzey Amerika birinci sırada yer alırken bunu Asya ve Avrupa'nın artan katkıları takip etmektedir. Okyanusya, Latin Amerika ve Karayipler ile Afrika ise yeterince temsil edilemeyen bölgelerdir.

Yazarlık rollerindeki cinsiyet dağılımında da benzer bir dengesizlik dikkat çekmektedir. Kadın araştırmacılar tüm kilit yazar kategorilerinde belirgin biçimde yetersiz temsil edilmektedir. Bu coğrafi ve cinsiyet eşitsizlikleri, oftalmoloji alanındaki BDM araştırmalarında küresel ve demografik temsilde dengesizlik olduğunu göstermekte ve bu alanda ilerleme kaydedilmesinin gerekliliğini vurgulamaktadır.

Beyan

Yazarlık Katkıları

Konsept: N.D.K., T.Y.A.L., Dizayn: N.D.K., T.Y.A.L., Veri Toplama veya İşleme: N.D.K., T.Y.A.L., Analiz veya Yorumlama: N.D.K., T.Y.A.L., Literatür Arama: N.D.K., T.Y.A.L., Yazan: N.D.K., T.Y.A.L.

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

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

Kaynaklar

- Amisha, Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. *J Family Med Prim Care*. 2019;8:2328-2331.
- Ray P. ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*. 2023;3:121-154.
- Large Language Model Introducing LLaMA 3.1: Our most capable models to date, 2024. <https://ai.meta.com/blog/meta-llama-3-1/>
- Kocoń J, Cichecki I, Kaszyca O, Kochanek M, Szydło D, Baran J, Bielaniec J, Gruza M, Janz A, Kanclerz K, Kocoń A, Koptyra B, Mieleśczenko-Kowszewicz W, Miłkowski P, Oleksy M, Piasecki M, Radliński Ł, Wojtasik K, Woźniak S, Kazienko P. ChatGPT: Jack of all trades, master of none. *Information Fusion*. 2023;99:101861.
- Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. 2018.
- Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, Aldairem A, Alrashed M, Bin Saleh K, Badreldin HA, Al Yami MS, Al Harbi S, Albekairy AM. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med Educ*. 2023;23:689.
- Nakaura T, Ito R, Ueda D, Nozaki T, Fushimi Y, Matsui Y, Yanagawa M, Yamada A, Tsuboyama T, Fujima N, Tatsugami F, Hirata K, Fujita S, Kamagata K, Fujioka T, Kawamura M, Naganawa S. The impact of large language models on radiology: a guide for radiologists on the latest innovations in AI. *Jpn J Radiol*. 2024;42:685-696.
- Wyatt KD, Alexander N, Hills GD, Liang WH, Kadauke S, Volchenboum SL, Mian A, Phillips CA. Making sense of artificial intelligence and large language models-including ChatGPT-in pediatric hematology/oncology. *Pediatr Blood Cancer*. 2024;71:e31143.
- Quer G, Topol EJ. The potential for large language models to transform cardiovascular medicine. *Lancet Digit Health*. 2024;6:e767-e771.
- Tan TF, Thirunavukarasu AJ, Campbell JP, Keane PA, Pasquale LR, Abramoff MD, Kalpathy-Cramer J, Lum F, Kim JE, Baxter SL, Ting DSW. Generative artificial intelligence through ChatGPT and other large language models in ophthalmology: clinical applications and challenges. *Ophthalmol Sci*. 2023;3:100394.
- Draux, H. Research on artificial intelligence – the global divides. January 4, 2024. <https://www.digital-science.com/blog/research-on-artificial-intelligence-the-global-divides/>
- Stanford Institute for Human-Centered AI (HAI). Artificial Intelligence Index Report 2024. <https://aiindex.stanford.edu/report/>
- Pal S, Lazzaroni RM, Mendoza P. AI's missing link: the gender gap in the talent pool. October 10, 2024. <https://www.stiftung-nv.de/publications/ai-gender-gap/>
- United Nations, D. o. E. a. S. A., Population Division World Population Prospects 2024. <https://population.un.org/wpp/definition-of-regions>
- Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A*. 2005;102:16569-16572.
- Cornell University Library. Measuring your research impact: i10-Index. <https://guides.library.cornell.edu/c.php?g=32272&p=203393>
- Ferro Desideri L, Roth J, Zinkernagel M, Anguita R. Application and accuracy of artificial intelligence-derived large language models in patients with age related macular degeneration. *Int J Retina Vitreous*. 2023;9:71.
- Li J, Guan Z, Wang J, Cheung CY, Zheng Y, Lim LL, Lim CC, Ruamviboonsuk P, Raman R, Corsino L, Echouffo-Tcheugui JB, Luk AOY, Chen LJ, Sun X, Hamzah H, Wu Q, Wang X, Liu R, Wang YX, Chen T, Zhang X, Yang X, Yin J, Wan J, Du W, Quek TC, Goh JHL, Yang D, Hu X, Nguyen TX, Szeto SKH, Chotcomwongse P, Malek R, Normatova N, Ibragimova N, Srinivasan R, Zhong P, Huang W, Deng C, Ruan L, Zhang C, Zhang C, Zhou Y, Wu C, Dai R, Koh SWC, Abdullah A, Hee NKY, Tan HC, Liew ZH, Tien CS, Kao SL, Lim AYL, Mok SF, Sun L, Gu J, Wu L, Li T, Cheng D, Wang Z, Qin Y, Dai L, Meng Z, Shu J, Lu Y, Jiang N, Hu T, Huang S, Huang G, Yu S, Liu D, Ma W, Guo M, Guan X, Yang X, Bascaran C, Cleland CR, Bao Y, Ekinci EI, Jenkins A, Chan JCN, Bee YM, Sivaprasad S, Shaw JE, Simó R, Keane PA, Cheng CY, Tan GSW, Jia W, Tham YC, Li H, Sheng B, Wong TY. Integrated image-based deep learning and language models for primary diabetes care. *Nat Med*. 2024;30:2886-2896.
- Anguita R, Makuloluwa A, Hind J, Wickham L. Large language models in vitreoretinal surgery. *Eye (Lond)*. 2024;38:809-810.
- Anguita R, Downie C, Ferro Desideri L, Sagoo MS. Assessing large language models' accuracy in providing patient support for choroidal melanoma. *Eye (Lond)*. 2024;38:3113-3117.
- Antaki F, Chopra R, Keane PA. Vision-language models for feature detection of macular diseases on optical coherence tomography. *JAMA Ophthalmol*. 2024;142:573-576.
- Chen X, Zhang W, Xu P, Zhao Z, Zheng Y, Shi D, He M. FFA-GPT: an automated pipeline for fundus fluorescein angiography interpretation and question-answer. *NPJ Digit Med*. 2024;7:111.
- Carlà MM, Gambini G, Baldascino A, Giannuzzi F, Boselli F, Crincoli E, D'Onofrio NC, Rizzo S. Exploring AI-chatbots' capability to suggest surgical planning in ophthalmology: ChatGPT versus Google Gemini analysis of retinal detachment cases. *Br J Ophthalmol*. 2024;108:1457-1469.
- Mohammadi SS, Nguyen QD. A User-friendly approach for the diagnosis of diabetic retinopathy using ChatGPT and automated machine learning. *Ophthalmol Sci*. 2024;4:100495.
- Ghalibafan S, Taylor Gonzalez DJ, Cai LZ, Graham Chou B, Panneerselvam S, Conrad Barrett S, Djulbegovic MB, Yannuzzi NA. Applications of multimodal generative artificial intelligence in a real-world retina clinic setting. *Retina*. 2024;44:1732-1740.
- Chen X, Xu P, Li Y, Zhang W, Song F, He M, Shi D. ChatFFA: an ophthalmic chat system for unified vision-language understanding and question answering for fundus fluorescein angiography. *iScience*. 2024;27:110021.

27. Chen X, Zhang W, Zhao Z, Xu P, Zheng Y, Shi D, He M. ICGA-GPT: report generation and question answering for indocyanine green angiography images. *Br J Ophthalmol*. 2024;108:1450-1456.
28. Gopalakrishnan N, Joshi A, Chhablani J, Yadav NK, Reddy NG, Rani PK, Pulipaka RS, Shetty R, Sinha S, Prabhu V, Venkatesh R. Recommendations for initial diabetic retinopathy screening of diabetic patients using large language model-based artificial intelligence in real-life case scenarios. *Int J Retina Vitreous*. 2024;10:11.
29. Balas M, Mandelcorn ED, Yan P, Ing EB, Crawford SA, Arjmand P. ChatGPT and retinal disease: a cross-sectional study on AI comprehension of clinical guidelines. *Can J Ophthalmol*. 2024;60:e117-e123.
30. Liu X, Wu J, Shao A, Shen W, Ye P, Wang Y, Ye J, Jin K, Yang J. Uncovering language disparity of ChatGPT on retinal vascular disease classification: cross-sectional study. *J Med Internet Res*. 2024;26:e51926.
31. Tailor PD, Dalvin LA, Chen JJ, Iezzi R, Olsen TW, Scruggs BA, Barkmeier AJ, Bakri SJ, Ryan EH, Tang PH, Parke DW 3rd, Belin PJ, Sridhar J, Xu D, Kuriyan AE, Yonekawa Y, Starr MR. A comparative study of responses to retina questions from either experts, expert-edited large language models, or expert-edited large language models alone. *Ophthalmol Sci*. 2024;4:100485.
32. Carlà MM, Gambini G, Baldascino A, Boselli F, Giannuzzi F, Margollicci F, Rizzo S. Large language models as assistance for glaucoma surgical cases: a ChatGPT vs. Google Gemini comparison. *Graefes Arch Clin Exp Ophthalmol*. 2024;262:2945-2959.
33. Delsoz M, Raja H, Madadi Y, Tang AA, Wirostko BM, Kahook MY, Yousefi S. The use of ChatGPT to assist in diagnosing glaucoma based on clinical case reports. *Ophthalmol Ther*. 2023;12:3121-3132.
34. Huang X, Raja H, Madadi Y, Delsoz M, Poursoroush A, Kahook MY, Yousefi S. Predicting glaucoma before onset using a large language model chatbot. *Am J Ophthalmol*. 2024;266:289-299.
35. Xue X, Zhang D, Sun C, Shi Y, Wang R, Tan T, Gao P, Fan S, Zhai G, Hu M, Wu Y. Xiaoqing: a Q&A model for glaucoma based on LLMs. *Comput Biol Med*. 2024;174:108399.
36. Raja H, Huang X, Delsoz M, Madadi Y, Poursoroush A, Munawar A, Kahook MY, Yousefi S. Diagnosing glaucoma based on the ocular hypertension treatment study dataset using chat generative pre-trained transformer as a large language model. *Ophthalmol Sci*. 2025;5:100599.
37. AlRiyal SA, Musleh AM, Kahook MY. Evaluating the strengths and limitations of multimodal ChatGPT-4 in detecting glaucoma using fundus images. *Front Ophthalmol (Lausanne)*. 2024;4:1387190.
38. Ćirković A, Katz T. Exploring the Potential of ChatGPT-4 in predicting refractive surgery categorizations: comparative study. *JMIR Form Res*. 2023;7:e51798.
39. Delsoz M, Madadi Y, Raja H, Munir WM, Tamm B, Mehravaran S, Soleimani M, Djalilian A, Yousefi S. Performance of ChatGPT in diagnosis of corneal eye diseases. *Cornea*. 2024;43:664-670.
40. Tuttle JJ, Moshirfar M, Garcia J, Altaf AW, Omidvarnia S, Hoopes PC. Learning the Randleman criteria in refractive surgery: utilizing ChatGPT-3.5 versus internet search engine. *Cureus*. 2024;16:e64768.
41. Rojas-Carabali W, Sen A, Agarwal A, Tan G, Cheung CY, Rousselot A, Agrawal R, Liu R, Cifuentes-González C, Elze T, Kempen JH, Sobrin L, Nguyen QD, de-la-Torre A, Lee B, Gupta V, Agrawal R. Chatbots Vs. human experts: evaluating diagnostic performance of chatbots in uveitis and the perspectives on AI adoption in ophthalmology. *Ocul Immunol Inflamm*. 2024;32:1591-1598.
42. Schumacher I, Bühler VMM, Jaggi D, Roth J. Artificial intelligence derived large language model in decision-making process in uveitis. *Int J Retina Vitreous*. 2024;10:63.
43. Marshall RF, Mallem K, Xu H, Thorne J, Burkholder B, Chaon B, Liberman P, Berkenstock M. Investigating the accuracy and completeness of an artificial intelligence large language model about uveitis: an evaluation of ChatGPT. *Ocul Immunol Inflamm*. 2024;32:2052-2055.
44. Madadi Y, Delsoz M, Lao PA, Fong JW, Hollingsworth TJ, Kahook MY, Yousefi S. ChatGPT assisting diagnosis of neuro-ophthalmology diseases based on case reports. *medRxiv [Preprint]*. 2023:2023.
45. Tailor PD, Dalvin LA, Starr MR, Tajfirouz DA, Chodnicki KD, Brodsky MC, Mansukhani SA, Moss HE, Lai KE, Ko MW, Mackay DD, Di Nome MA, Dumitrascu OM, Pless ML, Eggenberger ER, Chen JJ. A comparative study of large language models, human experts, and expert-edited large language models to neuro-ophthalmology questions. *J Neuroophthalmol*. 2025;45:71-77.
46. Upadhyaya DP, Shaikh AG, Cakir GB, Prantzos K, Golnari P, Ghasia FF, Sahoo SS. A 360° view for large language models: early detection of amblyopia in children using multi-view eye movement recordings. *medRxiv [Preprint]*. 2024:2024.
47. Luo MJ, Pang J, Bi S, Lai Y, Zhao J, Shang Y, Cui T, Yang Y, Lin Z, Zhao L, Wu X, Lin D, Chen J, Lin H. Development and evaluation of a retrieval-augmented large language model framework for ophthalmology. *JAMA Ophthalmol*. 2024;142:798-805.
48. Zheng C, Ye H, Guo J, Yang J, Fei P, Yuan Y, Huang D, Huang Y, Peng J, Xie X, Xie M, Zhao P, Chen L, Zhang M. Development and evaluation of a large language model of ophthalmology in Chinese. *Br J Ophthalmol*. 2024;108:1390-1397.
49. Huang AS, Hirabayashi K, Barna L, Parikh D, Pasquale LR. Assessment of a large language model's responses to questions and cases about glaucoma and retina management. *JAMA Ophthalmol*. 2024;142:371-375.
50. Deng Z, Gao W, Chen C, Niu Z, Gong Z, Zhang R, Cao Z, Li F, Ma Z, Wei W, Ma L. OphGLM: an ophthalmology large language-and-vision assistant. *Artif Intell Med*. 2024;157:103001.
51. Haghghi T, Gholami S, Sokol JT, Kishnani E, Ahsaniyan A, Rahmanian H, Hedayati F, Leng T, Alam MN. EYE-LLaMA, an in-domain large language model for ophthalmology. *bioRxiv [Preprint]*. 2025.
52. Chen JS, Reddy AJ, Al-Sharif E, Shoji MK, Kalaw FGP, Eslani M, Lang PZ, Arya M, Koretz ZA, Bolo KA, Arnett JJ, Roginiel AC, Do JL, Robbins SL, Camp AS, Scott NL, Rudell JC, Weinreb RN, Baxter SL, Granet DB. Analysis of ChatGPT responses to ophthalmic cases: can ChatGPT think like an ophthalmologist? *Ophthalmol Sci*. 2024;5:100600.
53. Milad D, Antaki F, Milad J, Farah A, Khairy T, Mikhail D, Giguère CÉ, Touma S, Bernstein A, Szigiato AA, Nayman T, Mullie GA, Duval R. Assessing the medical reasoning skills of GPT-4 in complex ophthalmology cases. *Br J Ophthalmol*. 2024;108:1398-1405.
54. Hu X, Ran AR, Nguyen TX, Szeto S, Yam JC, Chan CKM, Cheung CY. What can GPT-4 do for diagnosing rare eye diseases? A pilot study. *Ophthalmol Ther*. 2023;12:3395-3402.
55. Wu JH, Nishida T, Liu TYA. Accuracy of large language models in answering ophthalmology board-style questions: a meta-analysis. *Asia Pac J Ophthalmol (Phila)*. 2024;13:100106.
56. Raimondi R, Tzoumas N, Salisbury T, Di Simplicio S, Romano MR; North East Trainee Research in Ophthalmology Network (NETRION). Comparative analysis of large language models in the Royal College of Ophthalmologists fellowship exams. *Eye (Lond)*. 2023;37:3530-3533.

57. Sakai D, Maeda T, Ozaki A, Kanda GN, Kurimoto Y, Takahashi M. Performance of ChatGPT in board examinations for specialists in the Japanese Ophthalmology Society. *Cureus*. 2023;15:e49903.
58. Fowler T, Pullen S, Birkett L. Performance of ChatGPT and Bard on the official part 1 FRCOphth practice questions. *Br J Ophthalmol*. 2024;108:1379-1383.
59. Yaïci R, Cieplucha M, Bock R, Moayed F, Bechrakis NE, Berens P, Feltgen N, Friedburg D, Gräf M, Guthoff R, Hoffmann EM, Hoerauf H, Hintschich C, Kohnen T, Messmer EM, Nentwich MM, Pleyer U, Schaudig U, Seitz B, Geerling G, Roth M. ChatGPT und die deutsche Facharztprüfung für Augenheilkunde: eine Evaluierung [ChatGPT and the German board examination for ophthalmology: an evaluation]. *Ophthalmologie*. 2024;121:554-564.
60. Ming S, Guo Q, Cheng W, Lei B. Influence of model evolution and system roles on ChatGPT's performance in Chinese medical licensing exams: comparative study. *JMIR Med Educ*. 2024;10:e52784.
61. Bahir D, Zur O, Attal L, Nujeidat Z, Knaanie A, Pikkal J, Mimouni M, Plopsky G. Gemini AI vs. ChatGPT: a comprehensive examination alongside ophthalmology residents in medical knowledge. *Graefes Arch Clin Exp Ophthalmol*. 2025;263:527-536.
62. Gill GS, Tsai J, Moxam J, Sanghvi HA, Gupta S. Comparison of Gemini advanced and ChatGPT 4.0's performances on the ophthalmology resident ophthalmic knowledge assessment program (OKAP) examination review question banks. *Cureus*. 2024;16:e69612.
63. Antaki F, Milad D, Chia MA, Giguère CÉ, Touma S, El-Khoury J, Keane PA, Duval R. Capabilities of GPT-4 in ophthalmology: an analysis of model entropy and progress towards human-level medical question answering. *Br J Ophthalmol*. 2024;108:1371-1378.
64. Antaki F, Touma S, Milad D, El-Khoury J, Duval R. Evaluating the performance of ChatGPT in ophthalmology: an analysis of its successes and shortcomings. *Ophthalmol Sci*. 2023;3:100324.
65. Kianian R, Sun D, Crowell EL, Tsui E. The use of large language models to generate education materials about uveitis. *Ophthalmol Retina*. 2024;8:195-201.
66. Dihan Q, Chauhan MZ, Eleiwa TK, Brown AD, Hassan AK, Khodeiry MM, Elsheikh RH, Oke I, Nihalani BR, VanderVeen DK, Sallam AB, Elhusseiny AM. Large language models: a new frontier in paediatric cataract patient education. *Br J Ophthalmol*. 2024;108:1470-1476.
67. Dihan Q, Chauhan MZ, Eleiwa TK, Hassan AK, Sallam AB, Khouri AS, Chang TC, Elhusseiny AM. Using large language models to generate educational materials on childhood glaucoma. *Am J Ophthalmol*. 2024;265:28-38.
68. Dihan QA, Brown AD, Zaldivar AT, Chauhan MZ, Eleiwa TK, Hassan AK, Solyman O, Gise R, Phillips PH, Sallam AB, Elhusseiny AM. Advancing patient education in idiopathic intracranial hypertension: the promise of large language models. *Neurol Clin Pract*. 2025;15:e200366.
69. Singer MB, Fu JJ, Chow J, Teng CC. Development and evaluation of aeyeconsult: a novel ophthalmology chatbot leveraging verified textbook knowledge and GPT-4. *J Surg Educ*. 2024;81:438-443.
70. Waisberg E, Ong J, Masalkhi M, Lee AG. Large language model (LLM)-driven chatbots for neuro-ophthalmic medical education. *Eye (Lond)*. 2024;38:639-641.
71. Durmaz Engin C, Karatas E, Ozturk T. Exploring the Role of ChatGPT-4, BingAI, and Gemini as virtual consultants to educate families about retinopathy of prematurity. *Children (Basel)*. 2024;11:750.
72. Jung H, Oh J, Stephenson KAJ, Joe AW, Mammo ZN. Prompt engineering with ChatGPT3.5 and GPT4 to improve patient education on retinal diseases. *Can J Ophthalmol*. 2025;60:e375-e381.
73. Cai LZ, Shaheen A, Jin A, Fukui R, Yi JS, Yannuzzi N, Alabiad C. Performance of generative large language models on ophthalmology board-style questions. *Am J Ophthalmol*. 2023;254:141-149.
74. Thirunavukarasu AJ, Mahmood S, Malem A, Foster WP, Sanghera R, Hassan R, Zhou S, Wong SW, Wong YL, Chong YJ, Shakeel A, Chang YH, Tan BKJ, Jain N, Tan TF, Rauz S, Ting DSW, Ting DSJ. Large language models approach expert-level clinical knowledge and reasoning in ophthalmology: a head-to-head cross-sectional study. *PLOS Digit Health*. 2024;3:e0000341.
75. Sevgi M, Antaki F, Keane PA. Medical education with large language models in ophthalmology: custom instructions and enhanced retrieval capabilities. *Br J Ophthalmol*. 2024;108:1354-1361.
76. Chen X, Zhao Z, Zhang W, Xu P, Wu Y, Xu M, Gao L, Li Y, Shang X, Shi D, He M. EyeGPT for patient inquiries and medical education: development and validation of an ophthalmology large language model. *J Med Internet Res*. 2024;26:e60063.
77. Baxter SL, Longhurst CA, Millen M, Sitapati AM, Tai-Seale M. Generative artificial intelligence responses to patient messages in the electronic health record: early lessons learned. *JAMIA Open*. 2024;7:ooae028.
78. Cohen SA, Brant A, Fisher AC, Pershing S, Do D, Pan C. Dr. Google vs. Dr. ChatGPT: exploring the use of artificial intelligence in ophthalmology by comparing the accuracy, safety, and readability of responses to frequently asked patient questions regarding cataracts and cataract surgery. *Semin Ophthalmol*. 2024;39:472-479.
79. Muntean GA, Marginean A, Groza A, Damian I, Roman SA, Hapca MC, Sere AM, Mănoiu RM, Muntean MV, Nicoară SD. A qualitative evaluation of ChatGPT4 and PaLM2's response to patient's questions regarding age-related macular degeneration. *Diagnostics (Basel)*. 2024;14:1468.
80. Strzalkowski P, Strzalkowska A, Chhablani J, Pfau K, Errera MH, Roth M, Schaub F, Bechrakis NE, Hoerauf H, Reiter C, Schuster AK, Geerling G, Guthoff R. Evaluation of the accuracy and readability of ChatGPT-4 and Google Gemini in providing information on retinal detachment: a multicenter expert comparative study. *Int J Retina Vitreous*. 2024;10:61.
81. Kayabaşı M, Köksaldı S, Durmaz Engin C. Evaluating the reliability of the responses of large language models to keratoconus-related questions. *Clin Exp Optom*. 2025;108:784-791.
82. Shi R, Liu S, Xu X, Ye Z, Yang J, Le Q, Qiu J, Tian L, Wei A, Shan K, Zhao C, Sun X, Zhou X, Hong J. Benchmarking four large language models' performance of addressing Chinese patients' inquiries about dry eye disease: A two-phase study. *Heliyon*. 2024;10:e34391.
83. Tan DNH, Tham YC, Koh V, Loon SC, Aquino MC, Lun K, Cheng CY, Ngiam KY, Tan M. Evaluating Chatbot responses to patient questions in the field of glaucoma. *Front Med (Lausanne)*. 2024;11:1359073.
84. Ichhpujani P, Parmar UPS, Kumar S. Appropriateness and readability of Google Bard and ChatGPT-3.5 generated responses for surgical treatment of glaucoma. *Rom J Ophthalmol*. 2024;68:243-248.
85. Pushpanathan K, Lim ZW, Er Yew SM, Chen DZ, Hui'En Lin HA, Lin Goh JH, Wong WM, Wang X, Jin Tan MC, Chang Koh VT, Tham YC. Popular large language model chatbots' accuracy, comprehensiveness, and self-awareness in answering ocular symptom queries. *iScience*. 2023;26:108163.
86. Tailor PD, Xu TT, Fortes BH, Iezzi R, Olsen TW, Starr MR, Bakri SJ, Scruggs BA, Barkmeier AJ, Patel SV, Baratz KH, Bernhisel AA, Wagner LH, Tooley AA, Roddy GW, Sit AJ, Wu KY, Bothun ED, Mansukhani SA, Mohney BG, Chen JJ, Brodsky MC, Tajfirouze DA, Chodnicki KD, Smith WM, Dalvin LA. Appropriateness of ophthalmology recommendations from an online chat-based artificial intelligence model. *Mayo Clin Proc Digit Health*. 2024;2:119-128.

87. Alqudah AA, Aleshawi AJ, Baker M, Alnajjar Z, Ayasrah I, Ta'ani Y, Al Salkhadi M, Aljawarneh S. Evaluating accuracy and reproducibility of ChatGPT responses to patient-based questions in Ophthalmology: an observational study. *Medicine (Baltimore)*. 2024;103:e39120.
88. Zandi R, Fahey JD, Drakopoulos M, Bryan JM, Dong S, Bryar PJ, Bidwell AE, Bowen RC, Lavine JA, Mirza RG. Exploring diagnostic precision and triage proficiency: a comparative study of GPT-4 and bard in addressing common ophthalmic complaints. *Bioengineering (Basel)*. 2024;11:120.
89. Bernstein IA, Zhang YV, Govil D, Majid I, Chang RT, Sun Y, Shue A, Chou JC, Schehlein E, Christopher KL, Groth SL, Ludwig C, Wang SY. Comparison of ophthalmologist and large language model chatbot responses to online patient eye care questions. *JAMA Netw Open*. 2023;6:e2330320.
90. Wang H, Masselos K, Tong J, Connor HRM, Scully J, Zhang S, Rafla D, Posarelli M, Tan JCK, Agar A, Kalloniatis M, Phu J. ChatGPT for addressing patient-centered frequently asked questions in glaucoma clinical practice. *Ophthalmol Glaucoma*. 2024;8:157.
91. Wu G, Zhao W, Wong A, Lee DA. Patients with floaters: answers from virtual assistants and large language models. *Digit Health*. 2024;10:20552076241229933.
92. Cappellani F, Card KR, Shields CL, Pulido JS, Haller JA. Reliability and accuracy of artificial intelligence ChatGPT in providing information on ophthalmic diseases and management to patients. *Eye (Lond)*. 2024;38:1368-1373.
93. Lim ZW, Pushpanathan K, Yew SME, Lai Y, Sun CH, Lam JSH, Chen DZ, Goh JHL, Tan MCJ, Sheng B, Cheng CY, Koh VTC, Tham YC. Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard. *EBioMedicine*. 2023;95:104770.
94. Reyhan AH, Mutaf Ç, Uzun İ, Yükseyayla F. A performance evaluation of large language models in keratoconus: a comparative study of ChatGPT-3.5, ChatGPT-4.0, Gemini, Copilot, Chatsonic, and Perplexity. *J Clin Med*. 2024;13:6512.
95. Raja H, Munawar A, Mylonas N, Delsoz M, Madadi Y, Elahi M, Hassan A, Abu Serhan H, Inam O, Hernandez L, Chen H, Tran S, Munir W, Abd-Alrazaq A, Yousefi S. Automated category and trend analysis of scientific articles on ophthalmology using large language models: development and usability study. *JMIR Form Res*. 2024;8:e52462.
96. Deiner MS, Deiner NA, Hristidis V, McLeod SD, Doan T, Lietman TM, Porco TC. Use of large language models to assess the likelihood of epidemics from the content of tweets: infodemiology study. *J Med Internet Res*. 2024;26:e49139.
97. Wang SY, Huang J, Hwang H, Hu W, Tao S, Hernandez-Boussard T. Leveraging weak supervision to perform named entity recognition in electronic health records progress notes to identify the ophthalmology exam. *Int J Med Inform*. 2022;167:104864.
98. Jaskari J, Sahlsten J, Summanen P, Moilanen J, Lehtola E, Aho M, Säpyskä E, Hietala K, Kaski K. DR-GPT: a large language model for medical report analysis of diabetic retinopathy patients. *PLoS One*. 2024;19:e0297706.
99. Shaheen A, Afflitto GG, Swaminathan SS. ChatGPT-assisted classification of postoperative bleeding following microinvasive glaucoma surgery using electronic health record data. *Ophthalmol Sci*. 2025;5:100602.
100. Aykut A, Sezenoz AS. Exploring the potential of code-free custom GPTs in ophthalmology: an early analysis of GPT store and user-creator guidance. *Ophthalmol Ther*. 2024;13:2697-2713.
101. Wu JH, Nishida T, Moghimi S, Weinreb RN. Effects of prompt engineering on large language model performance in response to questions on common ophthalmic conditions. *Taiwan J Ophthalmol*. 2024;14:454-457.
102. Singh S, Djalilian A, Ali MJ. ChatGPT and ophthalmology: exploring its potential with discharge summaries and operative notes. *Semin Ophthalmol*. 2023;38:503-507.
103. Marshall R, Xu H, Dalvin LA, Mishra K, Edalat C, Kirupaharan N, Francis JH, Berkenstock M. Accuracy and completeness of large language models about antibody-drug conjugates and associated ocular adverse effects. *Cornea*. 2024;44:851-855.
104. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med*. 2023;29:1930-1940.
105. Kedia N, Sanjeev S, Ong J, Chhablani J. ChatGPT and Beyond: an overview of the growing field of large language models and their use in ophthalmology. *Eye (Lond)*. 2024;38:1252-1261.
106. Clusmann J, Kolbinger FR, Muti HS, Carrero ZI, Eckardt JN, Laleh NG, Löffler CML, Schwarzkopf SC, Unger M, Veldhuizen GP, Wagner SJ, Kather JN. The future landscape of large language models in medicine. *Commun Med (Lond)*. 2023;3:141.
107. U.S. Food and Drug Administration. 24 Hour Summary of the Digital Health Advisory Committee November 20-21, 2024, Available from: <https://www.fda.gov/media/184078/download>
108. Goutam B, Hashmi MF, Geem ZW, Bokde ND. A comprehensive review of deep learning strategies in retinal disease diagnosis using fundus images. *IEEE Access*. 2022;10:57796-57823.
109. Song S, Peng K, Wang E, Liu TYA. Enhancing large language model performance on ophthalmology board-style questions with retrieval-augmented generation. *Invest Ophthalmol Vis Sci*. 2025;66:3930.
110. Peng K, Wang E, Song S, Liu TYA. Leveraging retrieval-augmented generation with large language models in answering image-based board-style ophthalmology questions. *Invest Ophthalmol Vis Sci*. 2025;66:3931.
111. Wang E, Song S, Peng K, Liu TYA. Performance of large language models in answering questions regarding social determinants of health in ophthalmology. *Invest Ophthalmol Vis Sci*. 2025;66:3936.
112. Meyer A, Streichert T. Twenty-five years of progress-lessons learned from JMIR publications to address gender parity in digital health authorships: bibliometric analysis. *J Med Internet Res*. 2024;26:e58950.
113. Holman L, Stuart-Fox D, Hauser CE. The gender gap in science: How long until women are equally represented? *PLoS Biol*. 2018;16:e2004956.
114. Shah SS. Gender bias in artificial intelligence: empowering women through digital literacy. *Premier Journal of Artificial Intelligence*. 2024;1:1000088.
115. Brück O. A bibliometric analysis of geographic disparities in the authorship of leading medical journals. *Commun Med (Lond)*. 2023;3:178.
116. Chen H, Rider CI, Jurgens D, Teplitskiy M. Geographical disparities in navigating rejection in science drive disparities in its file drawer. *Journal of Criminal Justice Education*. 2024:45.
117. Navigli R, Conia S, Ross B. Biases in large language models: origins, inventory, and discussion. *J Data Inf Qual*. 2023;15:1-21.
118. Rahkovsky I, Toney A, Boyack KW, Klavans R, Murdick DA. AI research funding portfolios and extreme growth. *Front Res Metr Anal*. 2021;6:630124.