Impact of Artificial Intelligence on academic competence in medical research: A Scoping Review

Impact de l'Intelligence Artificielle sur la compétence académique en recherche médicale: Une revue exploratoire

تأثير الذّكاء الاصطناعي على الكفاءة الأكاديميّة في البحث الطبّي: مراجعة استكشافية

Mouhamed Baheddine Soula[1], Hamza Gazzeh[1], Salima Bradai[1,2], Youssef Zanina[1,2], Mohamed Khelil[1,2], Ahmed Ben Abdelaziz[1,2,3]

1. Research Laboratory "Measurement and Support for Hospital Performance" (LR19SP01)
2. Medical Information Systems Department, Sahloul Hospital, Sousse (Tunisia)
3. Faculty of Medicine, Sousse, University of Sousse (Tunisia)

Abstract

Introduction: The effect of Artificial Intelligence on scientific research is currently controversial (over-use, refusal). The objective of this scoping review was to summarize the impact of Generative Pre-Trained Bots (*ChatGPT*) on medical research, for optimal use.

Methods: In accordance with the PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) guidelines, three electronic databases (*MEDLINE, Web of Science, and Scopus*) were searched for publications on the use of ChatGPT in medical research up to September 30, 2024. The inclusion criteria encompassed systematic reviews, meta-analyses, and reviews published in English or French. A synthesized and consensual perspective was then derived using the SWOT (*Strengths, Weaknesses, Opportunities, and Threats*) approach.

Results: Among the 120 articles identified during the study period, 33 publications were reviewed to describe the impact of *ChatGPT* on medical research skills. It emerged that *ChatGPT* was considered a tool for generating innovative ideas and analyzing "*Big Data*," helping researchers execute manually time-consuming processes. However, its responses may be accompanied by "hallucinations," posing a threat to the scientific integrity of medical research.

Conclusion: *ChatGPT* serves as a valuable aid in medical research, particularly in its conceptualization and writing phases. However, its potential drawbacks, such as "hallucinations," highlight the need to strengthen young researchers' skills in the proper use of Artificial Intelligence.

Keywords: Medical writing - Research - Artificial Intelligence- Publication - Scoping Review.

Résumé

Introduction: L'effet de l'Intelligence Artificielle sur la recherche scientifique est actuellement controversé (sur-usage, refus). L'objectif de cette revue exploratoire a été de synthétiser l'impact de *Generative Pre-Trained Bots (ChatGPT)* sur la recherche médicale, pour une utilisation optimale.

Méthodes: Conformément aux directives PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), trois bases de données électroniques (*MEDLINE, Web of Science et Scopus*) ont été explorées pour l'inclusion des publications spécifiques a l'usage de *ChatGPT* en recherche médicale jusqu'au 30 septembre 2024. Suite à l'inclusion des publications de types revue systématique, méta-analyse et essai contrôlé randomisé, publiées en anglais ou en français, une vision synthétique et consensuelle a été extraite selon l'approche SWOT (*Strength, Weakness, Opportunity, Threat*).

Résultats: Parmi 120 articles identifiés au cours de la période d'étude, 33 publications ont été examinées pour décrire l'impact de ChatGPT sur les aptitudes en recherche médicale. Il ressort qu'il a été considéré comme un outil générateur d'idées innovantes et d'analyses de «Big-Data», facilitant aux chercheurs les procédures manuellement chronophages. Cependant, ses réponses seraient accompagnées par des «hallucinations», menaçant l'intégrité scientifique de la recherche médicale.

Conclusion: *ChatGPT* constitue une aide précieuse en recherche médicale dans ses phases conceptuelle et rédactionnelle. Ses effets pervers («hallucinations») imposent aujourd'hui, le renforcement des capacités des jeunes chercheurs aux bonnes pratiques d'usage de l'Intelligence Artificielle.

Mots-clés: Rédaction médicale – Recherche - Intelligence artificielle - Publication - Revue exploratoire.

Correspondance

Salima Bradai

Research Laboratory "Measurement and Support for Hospital Performance". (LR19SP01) Medical Information Systems Department, Sahloul Hospital, Sousse (Tunisia).
Email: salima_bradai@outlook.com

LA TUNISIE MEDICALE-2025; Vol 103 (10): 1362-1374 DOI: 10.62438/tunismed.v103i10.6077

الملخص

المقدّمة: تُعدّ انعكاسات الذّكاء الاصطناعي على البحث العلمي قضيّة مثيرة للجدل في الوقت الحالي (الإفراط في الاستخدام، الرّفض). هدف هذه المراجعة الاستكشافية هو تلخيص تأثير الروبوتات التوليدية المدرّبة مسبقًا (*ChatGPT*) على البحث الطبّي، من أجل تحقيق استخدام أمثل.

الطرق: وفقًا لإرشادات*PRISMA* (*عناصر* الإبلاغ التفضيلية للمراجعات المنهجيّة والتحليلات التلوية)، تمّ البحث في ثلاث قواعد بيانات إلكترونية (*MEDLINE*, *Web of Science* و*Scopus*) عن مقالات تتعلّق باستخدام *ChatGPT* في البحث الطبي إلى حدّ 30 سبتمبر 2024. شملت معايير الإدراج المراجعات المنهجيّة، والتحليلات التلوية، والمراجعات المنشورة باللغة الإنجليزية أو الفرنسية. ثمّ استخلص منظور مُلخّص وتوافقي باستخدام نهج (*SWOT* نقاط القوّة، نقاط الضّعف، الفرص، و الإكراهات).

النتائج: من بين 120 مقالًا حدّدت خلال فترة الدّراسة، تمّ مراجعة 33 مقالا لوصف تأثير ChatGPT على مهارات البحث الطبّي. تبيّن أنه يُعتبر أداة لتوليد الأفكار المبتكرة وتحليل "البيانات الكبرى"، ممّا يساعد الباحثين في تنفيذ العمليات التّي تستغرق وقتًا طويلاً عند أدائها يدويًا. ومع ذلك، قد تكون استجاباته مصحوبة بـ "هلوسات"، ممّا يشكّل تهديدًا للنزاهة العلميّة في البحث الطبّي.

الخلاصة: يعدّ ChatGPT أداة ممتازة عند استخدامه بشكل صحيح. لكن مع ذلك، يجب أن تظلّ الأوساط العلميّة يقظة تجاه المخاطر التّي يطرحها الذّكاء الاصطناعي.

الكلمات المفتاحية: الكتابة الطبية – البحث – الذّكاء الاصطناعي– النشر –مراجعة استكشافية.

Introduction

Medical research is worldwide considered a systematic and rational approach to discover new scientific facts and understand the environment [1]. Usually, research follows a methodical and rigorous approach to collect and analyze data, especially in the medical field often employing experimental and observational methods [2]. Academic research has typically relied on manual methods to arrange Big-Data [3]. In the age of Artificial Intelligence (AI) various fields have been reformed, including the field of medical research and medical writing [4]. More recently, Large Language Models (LLM) have demonstrated better performance on exhaustive tasks [5]. An LLM is a machine-learning model that can imitate human intelligence by analyzing vast amounts of data and predicting responses based on the text it was trained on. Chatbots or the Generative Pre-Trained transformers (GPT) are a well-known LLM that was developed with various performances and different technologies [1]. This Tool has achieved extraordinary growth, exceeding 100 million users by January and more than one billion visits by February 2024 [6]. The first article published in MEDLINE on ChatGPT was in the 9th of December 2022. Over 1 600 article on the chatbot were suitable to be published in MEDLINE during only the year 2023. Its use by researchers for generating content in academic publishing raised a debate between experts of the field [7]. ChatGPT has already been used in different tasks to execute a medical study (identifying literature gaps, summarizing literature, drafting and editing manuscripts and even performing statistical analysis etc…) [7]. Some of the most important scientific institutions such as the International Committee of Medical Journal Editors (ICMJE) recommend its use in research [8], while others explicitly discourage attributing authorship to it [5]. On one hand, ChatGPT launched by OpenAI is likely chosen by users due to its accessibility and its ability to produce correct answers in different Specialties [5], Deepseek, developed by a former student, gives the opportunity to users to view and modify the source code to improve or customize it [9]. Even students are aware of these advantages and the risks offered by AI as it was shown in a survey conducted in 2025 by 118 qualified researchers and joined their professors in the debate [10]. This controversy arises from concerns regarding the potential for ChatGPT to produce artificial erroneous responses that may appear plausible yet lack fidelity to actual data [7]. Therefore it appears essential to define shared rules, standards, and methodologies to guide the conduct of studies and mitigate the risk of misconduct or errors in the era of AI tools and particularly ChatGPT [1]. The purpose of this Scoping review is to synthesize the medical literature on the strengths and opportunities of employing ChatGPT in medical research, its weaknesses and threats involved and ultimately to refine strategies for its optimal utilization.

Methods

This scoping review was reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Reviews (PRISMA-ScR) guidelines [11]. Institutional Review Board approval was not required due to the use of publicly available data. The research question is as follows: What are the intrinsic and extrinsic benefits and risks of using ChatGPT in the biomedical research process? Three electronic databases: MEDLINE, Scopus, and Web of Science Core Collection, were searched to extract the published articles from inception to September 30, 2024. The search query was developed by consensus among the authors to identify articles containing "ChatGPT" in the title, along with other terms related to biomedical research. To maximize sensitivity, various synonyms and truncations of keywords were employed (**Box 1**). Only synthetic studies: meta-analyses, systematic reviews, and reviews were included. Exclusion criteria are: (1) non-English or French records, (2) papers not related to the medical field, (3) opinion or commentary articles, and (4) off-topic papers not

meeting the eligibility criteria. The identified studies were uploaded to Zotero to eliminate duplicates and facilitate screening. Two independent reviewers initially screened the titles and abstracts of each retrieved study, followed by a full-text screening to confirm eligibility for inclusion. Disagreements were resolved through discussion and consensus. If unresolved, a third reviewer a Full Professor of Preventive Medicine with expertise in Health Sciences Research and Publishing provided the final decision.

The characteristics of each record were systematically extracted, including the article's specialty, type, country of affiliation, first author and their h-index identified from the Scimago Journal Rank (SJR) database, and the journal's Impact Factor retrieved through Google Scholar. The specialty of the article was assigned based on the specialty of the first author. For each article, the strengths and weaknesses both intrinsic and extrinsic of using ChatGPT at each stage of the research process were analyzed. To structure the data, the research process was divided into three main stages: 1. Conceptualize; 2. Conduct; 3. Communicate. Each of these stages was further subdivided into three sub-stages, resulting in a total of nine stages in the final research process as shown in **Box 2**. This framework, inspired by a methodological guide published by experts in biomedical scientific writing [12], is summarized in **Box 3**. The information was analyzed to make a list of evidence from the articles selected showing the advantages and Shortcomings of using ChatGPT in every step previously explained. The impact of ChatGPT on the field was collected according to the SWOT (Strengths, Weaknesses, Opportunities, and Threats) approach (**Box 4**). SWOT analysis examines the internal and external factors that affect a tool's performance. Strengths and weaknesses evaluate what a tool does well or ineffectively internally, while opportunities and threats explore positive and negative external factors that might impact the tool's effectiveness. This method helps identify the tool's competitive advantages and vulnerabilities, as well as the external factors that can either enhance or undermine its performance.

Box 1. Search queries for identifying records related to "*ChatGPT*" AND "Biomedical Research"

a. Search Query for MEDLINE

("chatgpt"[Title] AND ("writ*"[Title] OR "research*"[Title] OR "publis*"[Title] OR "publication*"[Title] OR "article*"[Title] OR "study"[Title] OR "studies"[Title])) AND (meta-analysis[Filter] OR review[Filter] OR systematic review[Filter])

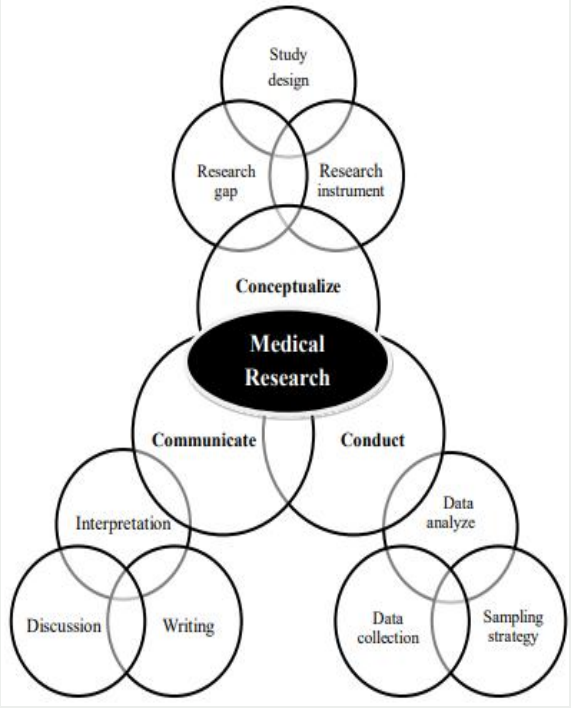
b. Search Query for Scopus

TITLE(chatgpt AND(writ* OR research* OR publis* OR publication* OR article* OR study OR studies)) AND (LIMIT-TO (DOCTYPE , "re"))

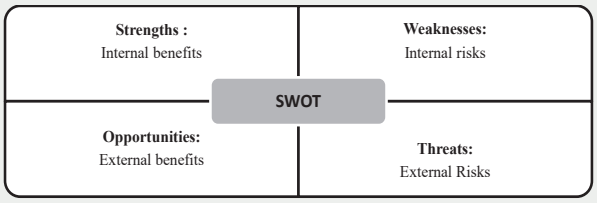
c. Search Query for Web of Science Core Collection

(TI=(chatgpt)) AND (TI=(writ*) OR TI=(research*) OR TI=(publis*) OR TI=(publication*) OR TI=(writ*) OR TI=(article*) AND TI=(study) OR TI=(studies))

We applied the following filter: Document Types: Review Article



Box 2. Core competencies in medical research: three primary aptitudes and their essential skills



Box 4. The four pillars of the SWOT Framework: Strengths, Opportunities, Weaknesses, and Threats.

Box 3. The 3C Framework of the Research Process: Conception, Conducting, and Communication with Their Sub-Stages

1. Conceptualize

1.1 Identifying the Research Gap

This first step involves identifying an under-explored or unanswered area in the existing literature through a comprehensive review, highlighting gaps, contradictions, or overlooked aspects, and thereby providing an opportunity for novel contributions to the field.

1.2 Choosing a Research Instrument

This step involves selecting a data collection tool that aligns with the research objectives, ensuring reliability, validity, and ethical suitability. The choice should also consider ease of use, participant consent, and minimizing biases or errors.

1.3 Selecting the Study Design

The final sub-stage involves evaluating and selecting the most appropriate study design. The researcher must critically assess different study types, considering their respective strengths and limitations concerning the study's objectives to ensure weather the chosen design effectively addresses the research question.

2. Conducting

2.1 Sampling

This sub-stage involves defining and selecting a representative sample from the target population. The sampling method must align with the research objectives, ensuring that the chosen sample is appropriate for addressing the research question while minimizing selection bias.

2.2 Data Collection

This step focuses on gathering information using the selected measurement instrument. Data collection can involve multiple sources, such as surveys, interviews, observational data, electronic health records…etc.

2.3 Data Analysis

This final sub-stage involves processing and analyzing the collected data using appropriate statistical methods. It includes the tabular and graphical presentation of results.

3. Communication

3.1 Interpretation

This sub-stage involves interpreting the study's results in relation to the research objectives, assessing their significance, and evaluating their implications within the context of the study's hypothesis.

3.2 Discussion

The discussion compares the study's results with similar research, highlighting consistencies, explaining discrepancies, and situating findings within the existing body of knowledge. It should also address study limitations and suggest directions for future research.

3.3 Writing

Writing entails drafting the manuscript following a predefined structure, typically the Introduction, Methods, Results, and Discussion (IMRAD) format. This step requires clear and concise presentation of findings while adhering to journal guidelines and academic standards.

Results

A total of 120 published medical records were extracted from the following Data bases: MEDLINE, Scopus, and Web of Science. The review was conducted on 33 publications. The flowchart for our systematic review is shown in **figure 1**. The characteristics of the included records are summarized in **table 1**. **Tables 2, 3** and **4** present the SWOT analysis applied to the findings obtained from the included articles. The following section highlights the key results, and more details are in the tables already mentioned.

Conceptualize: Not only ChatGPT is useful for quickly understanding and explaining scientific information for their research questions but also assists medical and health sciences researchers, especially novices, in selecting good research ideas relatively accurate and appropriate. ChatGPT also can assist in creating questionnaires to gather data from participants and can recommend study designs, but human knowledge and expertise should always guide the process. However, the most reported risk by researchers, from different fields, is "hallucinations" or "stochastic parroting" for the most part with specialized topics, where inaccurate texts might be confidently generated. Also, ChatGPT is basically a mirror to the data it was trained on. If the data contains bias, it will be spread by the chatbot relentlessly with the highly convincing language in which it can generate text.

Conduct: ChatGPT has the ability to collect vast amounts of data and extract key information from it making various research processes easier and less challenging. The capacity of analyzing a huge amount of Data by ChatGPT, can expedite result generation, allowing researchers to allocate more time to more important aspects of the research process. On the scale of universities and medical publisher, it is a huge opportunity to improve the rankings quickly. Using ChatGPT and AI tools boosts fluidity in terms of frequency of publishing. This benefit is

particularly pronounced for non-native English-speaking researchers. It is the tool that will "neutralize" the language barrier.

Communicate: ChatGPT has excellent performance in terms of article writing and editing. Articles generated by the AI model obtained higher scores than those written by humans due to the highly standardized and structured and also logical writing. Furthermore, it is undeniable that LLM can help suggesting ideas for discussion and can provide a relatively acceptable interpretation especially when given the right prompts. Nevertheless, extensive reliance on ChatGPT, a tool limited in its capacity for moral and value-based judgments, reduces trustworthiness and harms scientific integrity. If a scientist relies on ChatGPT for thinking, he or she may lose the ability to innovate. This tendency could inhibit the development of critical thinking, problem- solving, communication, and interpersonal skills among researchers ultimately weakening their motivation to engage in original thought.

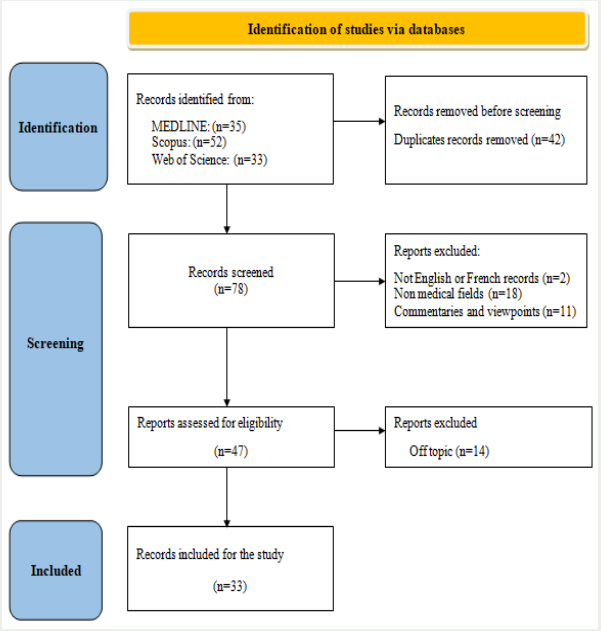


Figure 1. Flow Chart of the study selection process included in the scoping review on the Impact of Artificial Intelligence (AI) on medical academic competence in medical research conducted until the 30th of September 2024.

Table 1. Bibliometric analysis of articles included in a scoping review on the impact of Artificial Intelligence on academic competence in medical research, based on studies conducted up to September 30, 2024.

Article	Specialty/Field	First author	*h-index**	Article type	Country	Journal	IF**
[13]	Nursing Studies	Shefaly Shorey	54	Review article	Singapore	Nurse Education Today	8.6
[14]	Periodontology	Hiroj Bagde	10	Meta-analysis	India	Heliyon	3.4
[15]	Ophthalmology	Ding-Qiao Wang	18	Review article	China	MedComm – Future Medicine	8.6
[16]	Gynecology	Neli Semrl	NA	Journal article	Austria	Human Reproduction	8.6
[17]	Radiology	Li Mofan	NA	Review article	China	Journal advanced ultrasound	5.1
[18]	Orthopedics	AhmedA.Khalifa	12	Review article	Egypt	Arab Gulf Journal of Scientific Research	0.1
[19]	Family Medicine	Hunny Sharma	13	Review article	India	Perspectives in Clinical Research	2.7
[20]	Cardiology	Hassaan B Arshad	6	Review article	USA	Methodist Debakey cardiovascular journal	2.4
[21]	Orthopedics	Srijan Chatterjee	9	Review article	Korea	Journal of Experimental Orthopedics	1.8
[22]	Administration	Ahmed Keykha	5	Systematic review	Iran	Journal of Advances in Medical Education	11.0
[23]	Radiology	Sandeep Singh Awal	5	Review article	India	Journal of Public Health	0.9
[24]	Dentistry	Bader Fatani	7	Review article	KSA	Cureus	1.2
[25]	Physiology	Mondal, Himel	21	Review article	India	Indian journal of ophthalmology	2.1
[26]	Orthopedics	Madhan Jeyaraman	24	Review article	India	The World Journal of Methodology	NA
[27]	Orthopedics	Madhan Jeyaraman	24	Review article	USA	Cureus	1.2
[28]	Toxicology	Bhargava Daideepya Chandra	5	Review article	India	Medico-Legal Journal	1.5
[29]	Medicine	Afia Fatima	NA	Review article	Pakistan	Medicine	1.5
[30]	Medical Sciences	Songtao Tan	NA	Review article	China	International Journal of Surgery	12.0
[31]	Medical Sciences	Pradeep Sahu	14	Review article	Trinidad	Postgraduate Medical Journal	3.6
[32]	Information Systems	Maria Ijaz Baig	NA	Systematic review	Malaysia	International Journal of Educational Research	8.6
[33]	Microbiology	Malik Sallam	34	Review article	Jordan	Healthcare	1.9
[34]	Aesthetic Surgery	Rohun Gupta	10	Review article	USA	Aesthetic Surgery Journal	3.1
[35]	Neurology	Ravindra Garg	57	Review article	India	Health Promotion Perspectives	2.4
[36]	Civil Engineering	Ahmad Alshami	5	Review article	USA	Systems	2.5
[37]	Radiology	Sung Il Hwang	25	Review	Korea	Korean Journal of Radiology	7.1
[38]	Neurosurgery	Pietro Zangrossi	NA	Review article	Italy	Journal of Neurosurgical Sciences	2.3
[39]	Pharmacology	Chetna R. Patel	NA	Review article	India	Journal of Pharmacology	0.4
[40]	Biotechnology	Sang-Jun Kim	NA	Review article	Korea	Science Editing	1.3
[41]	Surgery	Zubair Muhammad Mojadeddi	NA	Review article	Denmark	New Zealand Medical Journal	0.7
[42]	Computing/ Medicine	Jingshan Huang	17	Review article	USA	The American Journal of Cancer Research	6.6
[43]	Orthopedic Surgery	Tyler Margetts	NA	Review article	USA	Current Osteoporosis Reports	4.2
[44]	Medical Nursing	Suebsarn Ruksakulpiwat	9	Review article	Thailand	Journal of Multidisciplinary Healthcare	4.6
[45]	Anesthesiology	Sang-Wook Lee	NA	Review article	Korea	Anesthesia and Pain Medicine	3.7

*Collected from Google Scholar **Impact Factor Collected from SCImago Journal Rank (SJR) **NA**: Not Available **USA**: United States of America **KSA**: Kingdom of Saudi Arabia

Table 2. SWOT analysis of ChatGPT's impact on the conceptualization of medical research, based on studies published up to September 30, 2024.

	Research gap	Research instrument	Study design
Strength			
	ChatGPT can help scientists to select a suitable topic for their literature review by generating relevant keywords and suggesting related and meaningful research areas. *ChatGPT* can generate a list of subtopics related to the main topic [17,22,24,30,32,40,42,45]. *ChatGPT*'s accuracy in answering queries about cancer reported an accuracy of 96.9% for the obtained responses and 100% for completeness, 39.4% accuracy for perfect answers and 18.3% accuracy for near-perfect answers [16].	If the attributes are provided, they can draft a questionnaire for a survey. In addition, they can translate the questionnaire or consent form into various languages [25].	Based on the research objectives, *ChatGPT* can suggest the most suitable study design. *ChatGPT* can assist in designing clinical trials by identifying potential inclusion and exclusion criteria, selecting appropriate endpoints, and analyzing the potential risks and benefits of the study [25,27]. *ChatGPT* was much more accurate resulting in a 35% accuracy rate for general topics and 75% for specific topics [38].
Weaknesses			
	There is a risk for 'hallucinations' or 'stochastic parroting' for specialized topics, where nonsensical or unfaithful texts are confidently generated LLM such as GPT are prone to hallucinating due to three reasons: 1) limitations in the training data, 2) transformer architecture, and 3) reinforced learning from human feedback (RLHF) [21–23,31,36,39,40,44,45,47]. If AI tools are only familiar with information about superficial spreading melanoma and not with other types of melanomas, the content they produce may be irrelevant because superficial spreading melanoma only accounts for about 70% of all melanomas [23].	-	-
Opportunities			
	Furthermore, it can help scholars save time and energy by generating ideas [25]. We found that *ChatGPT* was able to create a total of 105 novel ideas, which corresponded to a 49.3% accuracy rate [34].	-	-
Threats			
	ChatGPT has authoritative writing style and tendency to avoid admitting when it lacks an answer or makes errors, posing a challenge for healthcare students who lack specialized knowledge to discern between credible knowledge and unverified content [13].	-	-

Table 3. SWOT analysis of ChatGPT's impact on conducting medical research, based on studies published up to September 30, 2024.

	Sampling strategy	Data Collection	Data analyze
Strength			
	If the relevant information is provided, it can calculate the sample [25].	Due to time-consuming and complex processes, *ChatGPT* is becoming an appealing tool where the researcher instructs these models to conduct a literature search based on inputs reflecting the research question [19,25,30,33,35,42,43] For healthcare researchers, *ChatGPT*'s ability to quickly, efficiently, and accurately extract and analyze vast amounts of data enables various research processes to be streamlined [13,25,35,38,45].	*ChatGPT* can run tests such as t-tests, and Pearson correlation. Hence, ChatGPT can help in data analysis in Excel. Those who are using SPSS or other statistical software can ask ChatGPT [25,35]. *ChatGPT* could suggest suitable statistical tests such as Chi-square tests, t-tests for comparing means, or regression [13,17,19,22,27,30,33,34,39,44,46]. *ChatGPT* demonstrates excellent performance in writing R or Python code. Error messages are occasionally encountered when running R or Python code. In such cases, *ChatGPT* can be a valuable resource for correcting and improving the code [45].

Table 3. SWOT analysis of ChatGPT's impact on conducting medical research, based on studies published up to September 30, 2024. (continue)

	Sampling strategy	Data Collection	Data analyze
Weaknesses			
	However, we observed that although the formulas were correct, some calculations were wrong [13].	As *ChatGPT*'s knowledge base remains static with (cutoff point in September 2021) and is not linked to the internet, it is unable to acquire and integrate constantly evolving research findings, guidelines, and treatment protocols in real-time, unlike conventional search engines [13]. Only 8% of the references were retrievable via Google Scholar and Mendeley. In other studies of article writing, the accuracy of references was low, with only 7% being correct and 70% of the cited references being inaccurate [40]. Out of the 120 references provided by *ChatGPT* 3.5, only 16.66% of the references were completely correct, and most of them were nonexistent or inaccurate, with various errors [17,30,39,43,46]. Some neuroscientists pointed out that when *ChatGPT* was asked to help with literature retrieval, it could not judge which documents were more important [32].	-
Opportunities			
		It has the capability to simplify complex datasets into readable summaries, thereby saving researchers considerable time and effort. The efficiency advantage of *ChatGPT* can help scientists save time and focus on more creative efforts [13,17,20,46]. *ChatGPT* has the powerful ability to collect and summarize information, which can help researchers keep abreast of the latest progression in related fields [30].	The generation of novel protein structures and sequencing.20 Drawing from data, *ChatGPT* can predict the safety and efficacy of potential drugs and contributes to the evolution of new protein designs through its capacity to analyze data quickly [39].
Threats			
		Moreover, healthcare students and educators face potential threats to the confidentiality and security of their personal information, as well as that of their patients, when inadvertently disclosing sensitive data through user prompts [13,21,27,35,38,46]. In addition, paying to use *ChatGPT* may also affect the fairness of scientific research. The leaders of OpenAI have confirmed that their products will be available for a fee in the future. This will place a financial burden on institutions in low-income countries and regions, widening the gap in knowledge dissemination and scientific publishing [13,17]. Relying solely on AI-generated content without human verification can lead to potential inaccuracies, misinformation, or flawed conclusions in research [19,37]. Notably, many journals that were removed from the Web of Science list in 2023 belonged to Hindawi, a publisher acquired by Wiley. Along with removing these journals, the number of articles produced by "paper mills" was estimated to be in the hundreds of thousands. In 2023, Wiley-Hindawi retracted over 8,000 articles from a total of approximately 10,000 due to issues such as incoherent text and irrelevant references [40].	Limitations persist concerning accuracy, transparency, predictive outcomes, and the requirement for reliable data within clinical trials. The chief executive officer (CEO) of OpenAI has noted that *ChatGPT* is not infallible and may not consistently deliver accurate information [39].

Table 4. SWOT analysis of ChatGPT's impact on the communication of medical research, based on studies published up to September 30, 2024.

	Data interpretation	Discussion	Report writing
Strength			
	ChatGPT's main advantage is its ability to understand information quickly and connect evidence to reach conclusions faster than humans who have limitations in reading a wide range of literature comprehensively and connecting seemingly disparate pieces of information [19,33,44]. It might identify patterns such as a higher prevalence of heart disease among males over age 60 or correlations between obesity and heart disease incidence [20].	The Discussion section is primarily for explaining the main observations and interpreting the study in comparison with prior research. For these purposes, LLM can be of great help when given appropriate prompts [13,20,37,42].	*ChatGPT* capacity to curate appropriate language, and terminology also expedites the publication process and makes research findings more accurate, readable, and promptly available (which is especially beneficial during urgent public health crises e.g., corona virus pandemic) [13,17,19–22,24,35,38,42,45]. It could help with different manuscript preparation steps, including formatting, linguistic correction and translation for non-English-speaking authors [17]. Overall, it helps improve the writing quality by detecting areas of inconsistencies and potential errors [17,32]. In many situations, authors need to paraphrase the sentences of a part of the text to present the message in a more coherent manner to avoid textual plagiarism [25]. Articles generated by *ChatGPT* have been found to obtain higher scores than those written by humans. Its writing is highly standardized and structured and is also logical [30,33]. LLM can also help overcome the "blank page syndrome" that authors occasionally experience, particularly in the introduction section [37].
Weaknesses			
	Other *ChatGPT* drawbacks include the problems with interpretability, repeatability, and the treatment of uncertainty, all of which can have negative effects in healthcare settings and health care research[14]. The model cannot access contextual knowledge, which could complicate presenting persuasive research results [22]. ChatGPT is a machine and cannot think critically like a human, meaning it cannot interpret or analyze medical information beyond what is programmed into its system [27].	Biomedical research often requires critical analysis, interpretation of complex data, and making informed judgments. Currently, all the AI models are in the infancy stage and lack the ability to reason, evaluate evidence, or apply expert review in the same way human researchers can. This can generate responses that may not align with the field's best practices or accepted standards [17,19,23,27,47].	Moreover, there are ongoing debates concerning authorship rights for *ChatGPT*, with most prominent publishers and editors (Nature, Science, and the International Committee of Medical Journal Editors), advocating against its inclusion as a coauthor [13]. Since *ChatGPT* is unable to adhere to the World Association of Medical Editors (WAMEs) Recommendations on Chatbots and Generative AI or to the fourth ICMJE guideline that ad dresses accountability, it should be cited as a source of information rather than a co-author on scientific papers [13]. *ChatGPT* is likely to provide wrong information to readers. In addition, the training of *ChatGPT* will end at a certain point in time so the knowledge that it has will be limited[17]. The text generated by *ChatGPT* may contain articles published by other researchers. Using such text is equivalent to inadvertently plagiarism [13,17,24,37–39,42,43]. In addition, the use of AI products such as *ChatGPT* can undermine the fairness of scientific research. The convenience of *ChatGPT* may vary for researchers who use different languages. Some researchers asked *ChatGPT* the same set of neuroscience questions in different languages. In the Arabic version, the answers sometimes contained inaccurate or meaningless sentences that did not appear in the English and French answers [17].
Opportunities			
	-	-	ChatGPT, with the supervision of a human author, significantly shortened the manuscript writing time, reaching up to one hour for a complete manuscript draft [23,24,39,45]. Plagiarism detection could be performed using these tools [17,33,42,43]. ChatGPT assists in generating medical reports in multiple languages, making medical information more accessible to a wider audience which could potentially help overcome writer's block[20,23]. For researchers who are not native English speakers, the use of words always appears to be less authentic, and grammar or vocabulary errors may inadvertently occur. *ChatGPT* can be tremendously helpful [42].
Threats			
	Ethical concerns, including responsibility for inaccuracies and the potential for manipulation of study outcomes [23]. Dependency on AI tools like *ChatGPT* may stifle critical thinking among researchers. nuances in scientific content [35].	-	It can be concluded that the use of *ChatGPT* reduce the confidence of reviewers and readers in the article. This is contrary to the principles of openness, honesty, and equality advocated by scientists[17]. Possible unethical use of these tools (however, still not proven), such as its use to produce vast amounts of abstracts by organizations such as paper mills or even the production of fabricated research articles [17,20]. This can result in substantial expenses for non-native researchers, placing an additional burden on those from under-resourced countries [37]. This could make publication numbers less trustworthy as a parameter for promotion or hiring academics [17]. The possibility of research fraud associated with utilizing *ChatGPT* to produce academic publications, such as plagiarism, ghostwriting or falsified research, has become the primary concern [30].

Discussion

This review focused on identifying the impact of AI particularly ChatGPT in medical research. Such field is a dynamic and quickly evolving domain that interacts with all the discoveries around it and tries to profit from it [4]. This issue is clearly an urgent problem that has to be solved. The interest in this research came from the pressing need to establish a clear view about the use of AI in such a delicate domain [1]. The lack of information about the subject made it interesting to go for a synthetic study that can be a step in a whole path of discovering the optimized strategy of using AI in research [7]. All the articles included in this study were meta-analysis, systematic reviews and reviews. On the flip side, most of the published records about this subject are Editorials and their exclusion seems a loss of a wider range of information. However, the chosen type of records follows a predefined checklist and rigorous methodology. It's important to know that in this study, these reviews served as a valuable source for understanding the impact of ChatGPT. They offered a synthetic perspective from experts in the field, exhaustively consolidating and interpreting information presented in meta-analysis. The exclusion of studies in languages different from English or French could eliminate relevant articles, potentially introducing selection bias but the data loss is not significant. The analysis of the data revealed that some tasks are unanimously approved for execution using the chatbot within the medical community (consensus views). As previously explained, the consensus-driven uses of the AI model in terms of medical research will be labeled using the SWOT approach to maximize the efficiency of our results. Mainly ChatGPT is useful in collecting and extracting data. Even data analysis is efficient when given the needed prompts. Consequently, this tool can save energy and time allowing researchers to make more effort in other tasks. IT can also generate novel ideas and can help identify the research gap. On the flip side this model can hallucinate and misguide the whole research leading to a compromised scientific integrity. Furthermore, processing the different tasks of research without supervision might be a root of fraud and lack of creativity.

Strengths and Opportunities

The identification of research gaps is simply finding the areas that have to be explored or resolved. Recognizing these gaps creates great opportunities for novel investigations, enabling researchers, particularly those new to the discipline, to contribute original findings that enhance the understanding of a topic, address specific issues, or improve clinical and scientific practices. ChatGPT has demonstrated its potential in rapidly assimilating and explaining background information relevant to research questions. It provides support to medical researchers, especially novices, by helping in the selection of precise research topics. A recent survey published in 2025 that was conducted by 118 experts in research on 23 218 students from 109 countries and territories, offered a comprehensive global perspective on ChatGPT's initial acceptance and potential impact [10]. It stated that the generative AI improved researcher's specific knowledge and enhanced the access to information. This statement was endorsed by nearly 65% of the participants. Between 25% and 33% of the participants reported that the chatbot is mainly used for generating new ideas and brainstorming. In another recent study, this large language model generated 200 ideas for systematic reviews, with an overall accuracy rate of 57.5% in identifying novel research topics. The model achieved 39% accuracy for general subjects and 76% accuracy for more specialized topics [47].

ChatGPT's capability to rapidly and efficiently process big datasets and extract critical information makes various research activities easy, simplifying tasks that would otherwise be challenging. A survey where Students were aware of ChatGPT's capabilities, particularly in simplifying complex information (68%) and summarizing extensive content (67%) [10]. A study used a remarkably solid methodology in their preliminary study to establish evidence for the ability of ChatGPT's features to highlight key findings, explain statistical correlations, and provide clear summaries [63]. Another qualitative study (62 academic researchers from different fields in medicine) concluded that most participants (57/62) found literature analysis to be a complex and laborious process, often resulting in lack of critical information because of inevitable human error [48]. In such cases, the tool could be an irreplaceable assistant, as one participant in the study claimed [48] while 25% of students in another survey find it very convenient in terms of selecting articles for research.

ChatGPT has a remarkable performance in terms of article writing and editing and it is to be chosen by students in 22% of the cases [10]. The assistant demonstrated remarkable capabilities in improving medical writing productivity and enhancing the clarity and precision of scientific language [63]. Articles generated by the OpenAI's generative model scores better compared to those written by humans, mainly thanks to the highly standardized and structured and also logical writing provided by this tool, as evidenced by higher Likert scale ratings [6]. A review in the field of radiology confirmed that chatbots trained in medical writing could assist authors by drafting documents based on prompts and specifications [49]. Another research article stated that ChatGPT is better in writing argumentative student essays and outperforms the overall quality of essays written by humans [6]. The generated text can be well-executed and grammatically correct, but it may lack the creativity that is often notable in with human written text. This could lead to a homogenization of written content, and all text starts to be the same [50]. Furthermore, ChatGPT often produced paragraphs with a repetitive style and suboptimal grammar [51]. In a quantitative study results showed that the model had lower text perplexity relying on its competence in replicating commonly used terms and expressions [5]. Moreover, another investigation showed that human-written medical text size is obviously

larger than those of generated by the chatbot [52]. This suggests that medical texts produced by humans are more diverse [5].

ChatGPT's ability to analyze extensive datasets can boost the generation of research outcomes, allowing researchers to have more time to other critical aspects of the research process. Most of the recent studies have focused on utilizing open AI's tool to collect and synthesize complex datasets, including the generation of abstracts, conducting systematic reviews, and drafting essays. Furthermore, the assistant supports researchers in identifying novel connections and formulating new hypotheses within datasets, thus promoting scientific innovation [53]. A systematic review confirmed this notion, suggesting that it can propose new perspectives or underrepresented studies to enrich the literature review [27]. The platform also offers savings in terms of resources such as time and manpower, enabling researchers to allocate resources more effectively to other aspects of their work, potentially increasing overall productivity and research output [13]. This leads to more equitable access to publication opportunities, fosters diversity in research, and contributes to the democratization and dissemination of scientific knowledge. Its linguistic proficiency is especially beneficial for non-native speakers seeking to improve the quality of their medical writing [6].

Weaknesses and threats

The most concerning limitation of ChatGPT is its tendency to generate inaccurate or fabricated content, commonly referred to as "hallucinations" or "stochastic parroting". While being proficient in general medical writing, the tool may struggle with more complex or specialized topics requiring specific domain knowledge [63]. This issue is particularly concerning in specialized domains where erroneous information may be confidently produced without proper verification [54]. A recent investigation showed that students tend to deny that ChatGPT provides reliable information (41%) and replaces traditional classroom learning (41%) [10]. Different factors contribute to the higher incidence of errors and fabricated references, including the use of older versions of the language model (e.g., GPT-3), the absence of real-time internet access, and artificial hallucinations [55]. Additionally, performance degradation may appear over time due to concept drift, software errors, or the aging of the AI model itself [56]. An experimental study stated that inaccurate references are not negligible in the peer-reviewed literature, estimated between 4% and 8% of citations [57]. On the other hand, the inaccuracies in ChatGPT-generated references were more prevalent (93%) than those found in the peer-reviewed literature [57]. Further, these errors are more concerning since 47% of citations were invented [57]. Thus, these findings call into verifying the accuracy of any information provided by the chatbot [57].

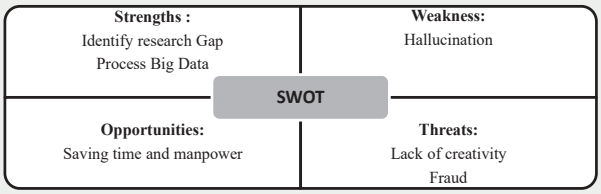
Another threat associated with the generative AI is the fact that researchers who depend excessively on AI models, such as ChatGPT, may lose the ability to innovate, potentially slowing down the acquisition of critical thinking, problem-solving, communicating, and social skills, particularly in healthcare students and educators. Although the AI model is valued for its technical efficiency in certain tasks, it is judged less capable of handling subjective content limiting its acceptance among students in fields that demand such capabilities notably medical research [10]. This reliance impairs the motivation to engage in any original thought [58]. Moreover, using the chatbot to generate research ideas could lead to superficial or trend-based suggestions rather than addressing medical knowledge gaps objectively [59]. The use of AI tools like ChatGPT inhibits the intellectual curiosity and critical thinking of researchers [48]. Another major concern is the generation of personalized patient-specific information, which could raise privacy concerns due to the need for access to protected health record [50]. On the other hand, Doyal and al. highlighted the issue of bias [50]. LLM are trained on extensive datasets, and any inherent biases within these datasets may be reflected in the text produced leading to the propagation of discriminatory or offensive language [50]. More concerning, such biases could affect the findings and conclusions of medical research. For instance, a model trained on datasets containing "fake news" may consistently generate inaccurate information [50]. Even writing and editing articles is not Guaranteed as ChatGPT's writing style can be repetitive, with similar syntax and sentence structures [63]. This agrees with the fact that ChatGPT-generated cover-letters lacked precision, were not fully informative, and seemed improperly structured to be submitted in a medical journal [64].

The rise of AI tools, especially ChatGPT, has triggered major discussions within the scientific community, making clear guidelines necessary. Even 66% of the inexperienced researchers believe that appropriate measures to protect personal information need to be taken and roughly 25% of them are concerned about privacy invasion [10]. While (ICMJE) does not ban the use of AI technologies, it requires full transparency regarding their use in studies and related research papers [8]. Specifically, the ICMJE does not consider ChatGPT or other AI tools as authors, and researchers must disclose the use of AI in their cover letter, acknowledgments, or methods section, depending on the case [8]. For example, if AI was involved in writing, editing, or proofreading, this should be mentioned in the acknowledgments. If AI was used for data collection, analysis, or creating figures, it must be reported in the methods section [8]. The European Commission follows the same guidelines as the ICMJE [60]. It also requires healthcare professionals to receive training on the use of Generative AI Systems to ensure that researchers are responsible for their texts and to protect privacy and confidential data [60]. A related debate centers on the integration of AI tools, such as ChatGPT, into research and education [60]. The Generative AI Product Tracker lists generative AI tools specifically available to be used or actively used by higher education instructors and students for teaching, learning, or research activities [60]. This tracker spans at least six domains in which AI tools have demonstrated efficiency and promise a new

era in these fields [60]. However, a research article also highlighted tensions between GPT users and GPT deniers, noting that a larger proportion of the first group had experimented with ChatGPT compared to the second one 58]. This warns about a potential generational divide between early adopters and more senior figures who dictate policies [58].

Given the potential for large LLM to be integrated into various contexts and the Gap between users and policy makers, it is reasonable to encourage different groups to explore and evaluate these technologies objectively. Only 40% of surveyed participants had used ChatGPT, with users demonstrating a more optimistic attitude towards it, while non-users expressed greater concerns about its widespread adoption [58]. Therefore, it is crucial to continue educating different contributors on such tools and their responsible use through practical demonstrations to familiarize non-users with the technology and mitigate concerns [58]. The Journal of the American Medical Association (JAMA) reacted to the rapid digitalization of health care and published a consensus-guided, adaptable framework in February 2025 [65]. This Framework came to conclusion that not only every medical school graduate has to be able to identify the different AI tools and LLM available in health care but also can describe how solutions based on these technologies may be used to achieve the best patient, population, and health system outcomes [65]. In the same vein, a master's program at the University of Porto's Faculty of Medicine was launched four years ago, offering a curriculum worth 360 credits following the recommendations of the European Credit Transfer and Accumulation System [61]. Incorporating ChatGPT into this program enables students to explore updated educational opportunities and engage with this technology in the safest way possible [61]. Moreover, it is important to promote institutional collaboration to help professors in integrating AI tools into the teaching process [61]. There is an ongoing discussion about the optimal and ethical integration of chatbots such as ChatGPT. Responsible use should be encouraged and future discussions must be continued to discover the limitations of this technology [58]. AI literacy, which is the ability to recognize, comprehend, utilize, and critically evaluate AI technologies and their implications, became needed especially in the medical field [62]. According to a recent global survey conducted in 2024, 53% of student participants viewed the chatbot as a valuable tool for enhancing AI literacy skills [10]. Another research paper showed that among 420 participants to their survey, only 40% of the respondents tried it [58]. Consequently, ongoing and intensive training session's should be organized for medical students, clinicians, and healthcare administrators who regularly engage with these systems [62]. Equipping medical contributors with new AI competencies is particularly crucial, as this will boost the development of an AI ecosystem that complements human knowledge [62]. Considering the emerging status of AI language models, the appearance of tools similar to ChatGPT and related research will multiply over time eventually leading future researchers to conduct more investigation on the topic and make sure that updates are provided whenever necessary [13].

Finally, ChatGPT is a promising tool currently revolutionizing medical research [53]. However, researchers are reporting risks and challenges that have to be kept in mind and more importantly need to be resolved [57]. This study is an overall view of the matter and more specific studies has to take every step of medical research and exploit its interaction with ChatGPT and AI tools in general in a comparative way. (**Box 5**) The collaboration between AI experts and medical researchers could be an effective way to cumulate knowledge and ultimately develop the optimal AI tool for the field in the future [39].



Box 5. Consensus-driven advantages and disadvantages of using ChatGPT in medical research

References

1. Sanmarchi F, Bucci A, Nuzzolese AG, Carullo G, Toscano F, Nante N, et al. A step-by-step researcher's guide to the use of an AI-based transformer in epidemiology: an exploratory analysis of ChatGPT using the STROBE checklist for observational studies. J Public Health (Berl). 2024;32(9):1761–96.
2. Ben Abdelaziz A, Sakly N, Melki S, Nouira S, Ben Abdelaziz A, Babba O, et al. The 5x5 approach in scientific biomedical writing. Tunis Med. 2021;99(6):585–600.
3. Dergaa I, Chamari K, Zmijewski P, Ben Saad H. From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. Biol Sport. 2023;40(2):615–22.
4. Chandra A, Dasgupta S. Impact of ChatGPT on Medical Research Article Writing and Publication. Sultan Qaboos Univ Med J. 2023;23(4):429.
5. Liao W, Liu Z, Dai H, Xu S, Wu Z, Zhang Y, et al. Differentiating ChatGPT-Generated and Human-Written Medical Texts: Quantitative Study. JMIR Med Educ. 2023;9(1):e48904.
6. Herbold S, Hautli-Janisz A, Heuer U, Kikteva Z, Trautsch A. A large-scale comparison of human-written versus ChatGPT-generated essays. Sci Rep. 2023;13(1):18617.
7. Gomes WJ, Evora PRB, Guizilini S. Artificial Intelligence is Irreversibly Bound to Academic Publishing — ChatGPT is Cleared for Scientific Writing and Peer Review. Braz J Cardiovasc Surg. 2023;38(4):e202309[65].
8. International Committee of Medical Journal Editors. Updated recommendations January 2024 [Internet]. 2024 [cited 2024 Oct 24]. Available from: https://www.icmje.org/news-and-editorials/updated_recommendations_jan2024.html.
9. Normile D. .Chinese firm's Large Language Model makes a splash. Science. 2025 Jan 17;387(6731):238. doi: 10.1126/science.adv9836. Epub 2025 Jan 16.
10. Ravšelj D, Keržič D, Tomaževič N, Umek L, Brezovar N, Iahad NA, et al. Higher education students' perceptions of ChatGPT: A global study of early reactions. PLoS One. 2025;20(2):e0315011. doi:10.1371/journal.pone.0315011.
11. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;n71.

12. Chebil D, Nouira S, Mkacher H, Yahia F, Ben Abdelaziz A, Barhoumi T, et al. How to write your "Research Protocol" in health sciences. Tunis Med. 2020 Jun;98(6):456-65.
13. Shorey S, Mattar C, Pereira TL-B, Choolani M. A scoping review of ChatGPT's role in healthcare education and research. Nurse Educ Today. 2024;135:106121.
14. Bagde H, Dhopte A, Alam MK, Basri R. A systematic review and meta-analysis on ChatGPT and its utilization in medical and dental research. Heliyon. 2023;9(12):e23050.
15. Wang D, Feng L, Ye J, Zou J, Zheng Y. Accelerating the integration of CHATGPT and other large-scale AI models into Biomedical Research and Healthcare. MedComm – Future Medicine. 2023 May 17;2(2). doi:10.1002/mef2.43
16. Semrl N, Feigl S, Taumberger N, Bracic T, Fluhr H, Blockeel C, et al. AI language models in human reproduction research: exploring ChatGPT's potential to assist academic writing. Hum Reprod. 2023;38(12):2281–8.
17. Li M, Zhang Y, Sun Y, Cui L, Wang S. AI-based ChatGPT Impact on Medical Writing and Publication. Adv Ultrasound Diagn Ther. 2023;7(2):188–92.
18. Khalifa AA, Ibrahim MA. Artificial Intelligence (AI) and CHATGPT involvement in scientific and medical writing, a new concern for researchers. A scoping review. Arab Gulf Journal Sci Res 2024 Jan 4;42(4):1770–87. doi:10.1108/agjsr-09-2023-0423
19. Sharma H, Ruikar M. Artificial intelligence at the pen's edge: Exploring the ethical quagmires in using artificial intelligence models like ChatGPT for assisted writing in biomedical research. Perspect Clin Res. 2024;15(3):108–15.
20. Arshad HB, Butt SA, Khan SU, Javed Z, Nasir K. ChatGPT and Artificial Intelligence in Hospital Level Research: Potential, Precautions, and Prospects. Methodist Debakey Cardiovasc J. 2023;19(5):77–84.
21. Chatterjee S, Bhattacharya M, Pal S, Lee S-S, Chakraborty C. ChatGPT and Large Language Models in orthopedics: from education and surgery to research. J Exp Orthop. 2023;10(1):128.
22. Keykha A, Behravesh S, Ghaemi F. ChatGPT and Medical Research: A Meta-Synthesis of Opportunities and Challenges. J Adv Med Educ Prof. 2024;12(3):135–47.
23. Awal SS, Awal SS. ChatGPT and the healthcare industry: a comprehensive analysis of its impact on medical writing. J Public Health [Internet]. 2023; Available from: https://doi.org/10.1007/s10389-023-02170-2.
24. Fatani B. ChatGPT for Future Medical and Dental Research. Cureus. 2023;15(4):e37285.
25. Mondal H, Mondal S. ChatGPT in academic writing: Maximizing its benefits and minimizing the risks. Indian J Ophthalmol. 2023;71(12):3600–6.
26. Jeyaraman M, Ramasubramanian S, Balaji S, Jeyaraman N, Nallakumarasamy A, Sharma S. ChatGPT in action: Harnessing Artificial Intelligence potential and addressing ethical challenges in medicine, education, and scientific research. World J Methodol. 2023;13(4):170–8.
27. Jeyaraman M, K SP, Jeyaraman N, Nallakumarasamy A, Yadav S, Bondili SK. ChatGPT in Medical Education and Research: A Boon or a Bane? Cureus. 2023;15(8):e44316.
28. Bhargava DC, Jadav D, Meshram VP, Kanchan T. ChatGPT in medical research: challenging time ahead. Med Leg J. 2023;91(4):223–5.
29. Fatima A, Shafique MA, Alam K, Fadlalla Ahmed TK, Mustafa MS. ChatGPT in medicine: A cross-disciplinary systematic review of ChatGPT's (artificial intelligence) role in research, clinical practice, education, and patient interaction. Medicine (Baltimore). 2024;103(32):e39250.
30. Tan S, Xin X, Wu D. ChatGPT in medicine: prospects and challenges: a review article. Int J Surg. 2024;110(6):3701–6.
31. Sahu PK, Benjamin LA, Singh Aswal G, Williams-Persad A. ChatGPT in research and health professions education: challenges, opportunities, and future directions. Postgrad Med J. 2023;100(1179):50–5.
32. Baig MI, Yadegaridehkordi E. ChatGPT in the higher education: A systematic literature review and research challenges. Int J Educ Res 2024;127:102411
33. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare (Basel). 2023;11(6):887.
34. Gupta R, Park JB, Bisht C, Herzog I, Weisberger J, Chao J, et al. Expanding Cosmetic Plastic Surgery Research With ChatGPT. Aesthet Surg J. 2023;43(8):930–7.
35. Garg RK, Urs VL, Agarwal AA, Chaudhary SK, Paliwal V, Kar SK. Exploring the role of ChatGPT in patient care (diagnosis and treatment) and medical research: A systematic review. Health PromotPerspect. 2023;13(3):183–91.
36. Alshami A, Elsayed M, Ali E, Eltoukhy A, Zayed T. Harnessing the Power of ChatGPT for Automating Systematic Review Process: Methodology, Case Study, Limitations, and Future Directions. Syst J. 2023;11(7).
37. Hwang SI, Lim JS, Lee RW, Matsui Y, Iguchi T, Hiraki T, et al. Is ChatGPT a "Fire of Prometheus" for Non-Native English-Speaking Researchers in Academic Writing? Korean J Radiol. 2023;24(10):952.
38. Zangrossi P, Martini M, Guerrini F, de Bonis P, Spena G. Large Language Model, AI and scientific research: why ChatGPT is only the beginning. J Neurosurg Sci. 2024;68(2):216–24.
39. Patel CR, Pandya SK, Sojitra BM. Perspectives of ChatGPT in Pharmacology Education, and Research in Health Care: A Narrative Review. J Pharmacol Pharmather. 2023;14(3):171–7.
40. Kim S. Research ethics and issues regarding the use of ChatGPT-like Artificial Intelligence platforms by authors and reviewers: a narrative review. Sci Ed. 2024;11(2).
41. Mojadeddi ZM, Rosenberg J. The impact of AI and ChatGPT on research reporting. New Zealand Med J. 2023;136(1575):60–4.
42. Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific review articles. Am J Cancer Res. 2023;13(4):1148–54.
43. Margetts TJ, Karnik SJ, Wang HS, Plotkin LI, Oblak AL, Fehrenbacher JC, et al. Use of AI Language Engine ChatGPT 4.0 to Write a Scientific Review Article Examining the Intersection of Alzheimer's Disease and Bone. Curr Osteoporos Rep. 2024;22(1):177–81.
44. Ruksakulpiwat S, Kumar A, Ajibade A. Using ChatGPT in Medical Research: Current Status and Future Directions. J.Multidiscip. Healthc. 2023;16:1513–20.
45. Lee S-W, Choi W-J. Utilizing ChatGPT in clinical research related to anesthesiology: a comprehensive review of opportunities and limitations. Anesth Pain Med (Seoul). 2023;18(3):244–51.
46. Omar M, Ullanat V, Loda M, Marchionni L, Umeton R. ChatGPT for digital pathology research. Lancet Digit Health. 2024;6(8):e595–600.
47. Zhang A, Dimock E, Gupta R, Chen K. The new frontier: utilizing ChatGPT to expand craniofacial research. Arch Craniofac Surg. 2024;25(3):116–22.
48. Alsadhan A, Al-Anezi F, Almohanna A, Alnaim N, Alzahrani H, Shinawi R, et al. The opportunities and challenges of adopting ChatGPT in medical research. Front Med (Lausanne). 2023;10:1259640.
49. Biswas S .ChatGPT and the Future of Medical Writing. Radiology. 2023 Apr;307(2):e223312. doi: 10.1148/radiol.223312. Epub 2023 Feb 2. PMID: 36728748.
50. Doyal AS, Sender D, Nanda M, Serrano RA. ChatGPT and Artificial Intelligence in Medical Writing: Concerns and Ethical Considerations. Cureus. 2023;15(8):e43292.
51. Cunningham AR, Behm HE, Ju A, Peach MS. Long-Term Survival of Patients With Glioblastoma of the Pineal Gland: A ChatGPT-Assisted, Updated Case of a Multimodal Treatment Strategy Resulting in Extremely Long Overall Survival at a Site With Historically Poor Outcomes. Cureus. 2023;15(3):e36590.
52. Hwang T, Aggarwal N, Khan PZ, Roberts T, Mahmood A, Griffiths MM, et al. Can ChatGPT assist authors with abstract writing in medical journals? Evaluating the quality of scientific abstracts generated by ChatGPT and original abstracts. PLoS One. 2024;19(2):e0297701.
53. Wu J, Ma Y, Wang J, Xiao M. The Application of ChatGPT in Medicine: A Scoping Review and Bibliometric Analysis. J MultidiscipHealthc. 2024;17:1681–92.
54. Sharun K, Banu SA, Pawde AM, Kumar R, Akash S, Dhama K, et al. ChatGPT and artificial hallucinations in stem cell research: assessing

the accuracy of generated references – a preliminary study. Ann Med Surg. 2023;85(10):5275–8.

55. Zimmerman A. A Ghostwriter for the Masses: ChatGPT and the Future of Writing. Ann Surg Oncol. 2023;30(6):3170–3.
56. Athaluri SA, Manthena SV, Kesapragada VSRKM, Yarlagadda V, Dave T, Duddumpudi RTS. Exploring the Boundaries of Reality: Investigating the Phenomenon of Artificial Intelligence Hallucination in Scientific Writing Through ChatGPT References. Cureus. 2023;15(4):e37432.
57. Bhattacharyya M, Miller VM, Bhattacharyya D, Miller LE. High Rates of Fabricated and Inaccurate References in ChatGPT-Generated Medical Content. Cureus2023;15(5):e39238.
58. Hosseini M, Gao CA, Liebovitz D, Carvalho A, Luo Y, MacDonald N, et al.An exploratory survey about using ChatGPT in education, healthcare, and research.PLoS One. 2023;18(10):e0292216.
59. Guckenberger M, Andratschke N, Ahmadsei M, Christ SM, Heusel AE, Kamal S, et al. Potential of ChatGPT in facilitating research in radiation oncology? Radiother Oncol. 2023;188:109894.
60. Generative AI Product Tracker [Internet]. Ithaka S+R. 2024 [Last visit: 2024 Oct 24]. Available from: https://sr.ithaka.org/our-work/generative-ai-product-tracker/
61. Araujo SM, Cruz-Correia R. Incorporating ChatGPT in Medical Informatics Education: Mixed Methods Study on Student Perceptions and Experiential Integration Proposals. JMIR Med Educ. 2024;10(1):e51151.
62. UygunIlikhan S, Özer M, Tanberkan H, Bozkurt V. How to mitigate the risks of deployment of Artificial Intelligence in medicine? Turk J Med Sci 2024;54(3):483–92.
63. Dergaa I, Ben Saad H, Glenn JM, Ben Aissa M, Taheri M, Swed S, et al. A thorough examination of ChatGPT-3.5 potential applications in medical writing: A preliminary study. Medicine. 2024;103(40):e39757. doi:10.1097/MD.0000000000039757.
64. Hachfi H, Kechida M, Kaddoussi R, Rejeb H, Alaya W, Hidouri S, et al. The cover letter in the era of Artificial Intelligence (ChatGPT as an example). Tunis Med. 2024 Dec 5;102(12):985-987. doi:10.62438/tunismed.v102i12.5439.
65. Car J, Ong QC, Erlikh Fox T, Leightley D, Kemp SJ, Švab I, et al. The Digital Health Competencies in Medical Education Framework: An International Consensus Statement Based on a Delphi Study. JAMA Netw Open. 2025 Jan 2;8(1):e2453131. doi:10.1001/jamanetworkopen.2024.53131.