

Exploring the application of CHATGPT in plastic surgery: a comprehensive systematic review

Fizzah Arif¹, Mohammad Khozema Safri², Zenab Shahzad³, Sobia Yasmeen⁴, Mohammad Fazlur Rahman⁵, Safdar Ali Shaikh⁶

Abstract

Objective: To determine the impact of ChatGPT in plastic surgery research and assess the authenticity of such contributions.

Methods: The study conducted a literature search in Sep'23 from databases like Pubmed, Google Scholar, SCOPUS, and OVID Medline. The following keywords 'ChatGPT', 'chatbot', 'reconstruction', 'aesthetic' and 'plastic surgery' were used. 32 papers were included from the initial 131 results of articles. English language articles from November 2022 to July 2023 discussing ChatGPT's role in plastic and aesthetic surgery were included whereas non-English documents, irrelevant content, and non-academic sources were excluded from the study.

Results: The manuscripts included in the systematic review had a diverse range, including original research articles, case reports, letters to the editor, and editorials. Among the included studies, there were 9 original research articles, 1 case report, 23 letters to the editor, and 2 editorials. Most publications originated from the United States (18) and Australia (7). Analysis suggested concerns, such as inaccuracies, plagiarism, outdated knowledge, and lack of personalized advice. Various authors recommend using ChatGPT as a supplementary tool rather than a replacement for human decision-making in medicine.

Conclusion: ChatGPT shows potential in plastic surgery research, concerns about inaccuracies and outdated knowledge may provide deceiving information and it always requires human input and verification.

Keywords: Plastic, ChatGPT, Reconstructive Surgery, Aesthetic Surgery, Chatbot.

DOI: <https://doi.org/10.47391/JPMA.AKU-9S-04>

^{1,4-6}Department of Plastic Surgery, Aga Khan University Hospital, ²Medical Student, Aga Khan University Hospital and Medical College, ³Medical Student, Liaquat National Medical College, Karachi, Pakistan.

Correspondence: Sobia Yasmeen **Email:** sobia.yasmeen@aku.edu

ORCID ID. 0009-0008-0761-9320

Introduction

With advancing technology, artificial intelligence (AI) models have started to revolutionize various aspects of our life, including scientific and clinical research. The introduction of ChatGPT (Generative Pre-trained Transformer) in November 2022, by OpenAI captured public attention world-wide.¹ ChatGPT is a type of chatbot designed to provide natural language-processing capabilities for a wide range of applications.² It is a large language model (LLM) that provides human like text content using statistical patterns from an enormous database of textbooks, articles and websites.³

Looking at the recent developments in the field of medical research alongside AI, there has been immense progress and expansion. Use of open AI based models has gained significant attention since its launch, but the role of ChatGPT in medical research especially could raise questions about its authenticity. Plastic Surgery is an innovative field that relies on research and evidence-based outcomes which is necessary for its progression⁴.

Some researchers are believed to keep chatbots as their research assistants helping them organize their thoughts, provide relevant feedback and even for conducting literature reviews, as it helps locate relevant articles and summarize the information precisely⁵.

Identifying medical conditions from radiologic images in the field of radiology have proved to be remarkable using AI systems. GPT-4, unlike GPT-3.5, can process images alongside deep learning and big data. This aids in more precise medical image analysis, like X-rays and Computed Tomography (CT) scans. It can identify and separate structures like organs, tumours, and vessels, as well as calculate parameters like tumour size and vascular constriction, enhancing surgical procedures, improving diagnostics, planning, efficiency, safety, and postoperative care in other surgical fields.^{6,7}

Moreover, it has been shown that research abstracts that are written by ChatGPT can pass plagiarism detectors making it absolutely difficult for scientists to identify if the abstract is an original work or is just generated by

ChatGPT⁸. This issue has raised important questions about authorship and publication ethics. The recent publications made using AI have been acknowledged for their effectiveness, however they have also received criticism for the potential flaws. One study has reported its lack of knowledge about basic anatomy demonstrated as inaccurate answers to the questions⁹. Moreover, the potential biases in the algorithms and the spread of inaccurate information are also a major aspect that needs to be considered^{4,8}.

This review article aims to determine the use and extent to which ChatGPT can contribute to ideation, teaching methodologies, and practical advancements in the field of Plastic Surgery.

Methods

The present study was carried out in Sep'23, through a literature search of scientific articles available on Pubmed,

Google Scholar, SCOPUS and OVID Medline published between November 2022 to July 2023. The keywords used included ChatGPT, chatbot, reconstruction, aesthetic, plastic surgery, with the "and" filler was also included during the keywords search. The PRISMA declaration methodology for systematic reviews was used¹⁰.

We initially identified 131 papers through our literature search using above mentioned keywords, and after excluding 82 similar or duplicate papers, we were left with 49 distinct articles. These were further narrowed down based on reading titles and abstracts, excluding those that did not meet the inclusion criteria. On reviewing, abstracts and titles, eight more articles were excluded. Thirty-two papers were selected, for which full text was managed and were finally selected for systemic review. (Figure 1).

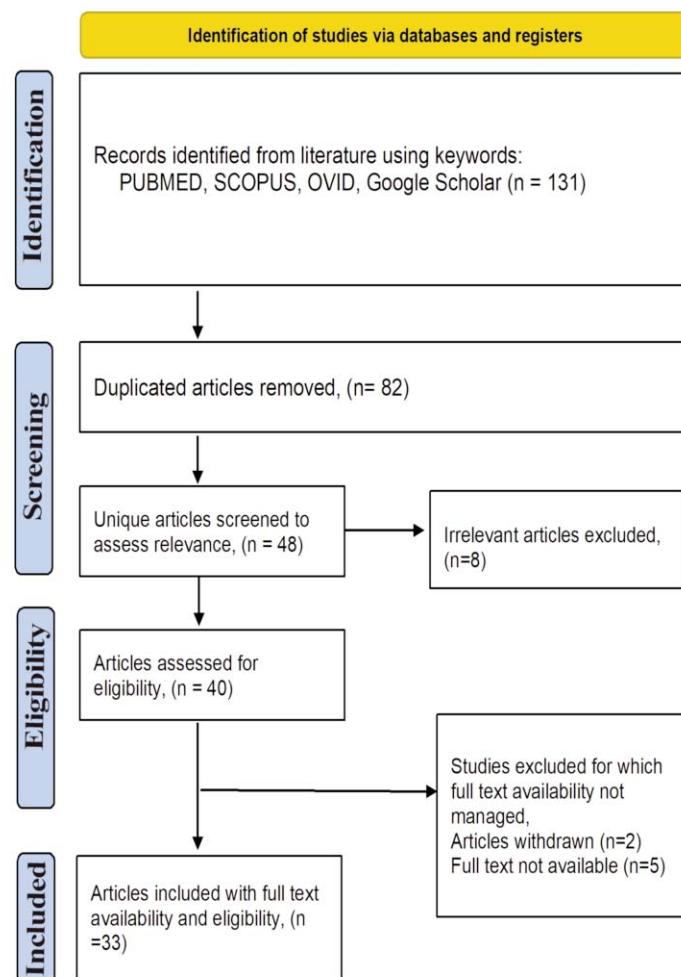


Figure-1: PRISMA flow diagram of the study selection process.

Inclusion criteria and exclusion criteria

The following were inclusion requirements:

- (1) Articles containing above keywords, published between November 2022 and July 31, 2023
- (2) Types of Articles included were articles, reviews, communications, editorials, and other types of preprints that explicitly describe or report the role of ChatGPT in plastic and aesthetic surgery.

The following were exclusion requirements:

- (1) manuscripts not composed in English;
- (2) manuscripts not directly pertinent to plastic or reconstructive surgery;
- (3) publications from non-academic sources, such as websites, newspapers, and magazines.
- (4) publications for which the full text was not available.

Results

Characteristics of Manuscripts

The systematic review included a variety of manuscripts, such as original research articles, review papers, case reports, editorials, and letters to editors. Some of these manuscripts explored the utilization of ChatGPT in clinical practices related to plastic and aesthetic surgery, while

Table-1: List of countries of first authors of all publications.

Countries of first author	N
America	16
Australia	7
China	3
Cambodia	2
Italy	2
UK	1
Canada	1
India	1
Total no. of Articles	33

others delved into its role in plastic surgery research. Additionally, some manuscripts examined ChatGPT's involvement in algorithm design through code generation and its relevance in academic contexts.

In total, we retrieved 33 manuscripts, among which 8 were original articles, 23 were letters to the editor, and 2 editorials. Majority of the publications were originated from the United States of America 16 (48.5%), followed by Australia 7 (21.2%), China 3 (9.1%), Cambodia 2 (6.1%), and Italy 2 (6.1%), UK, 1 (3%), Canada 1 (3%) and India 1 (3%). (Table no.1) (Fig. 2)

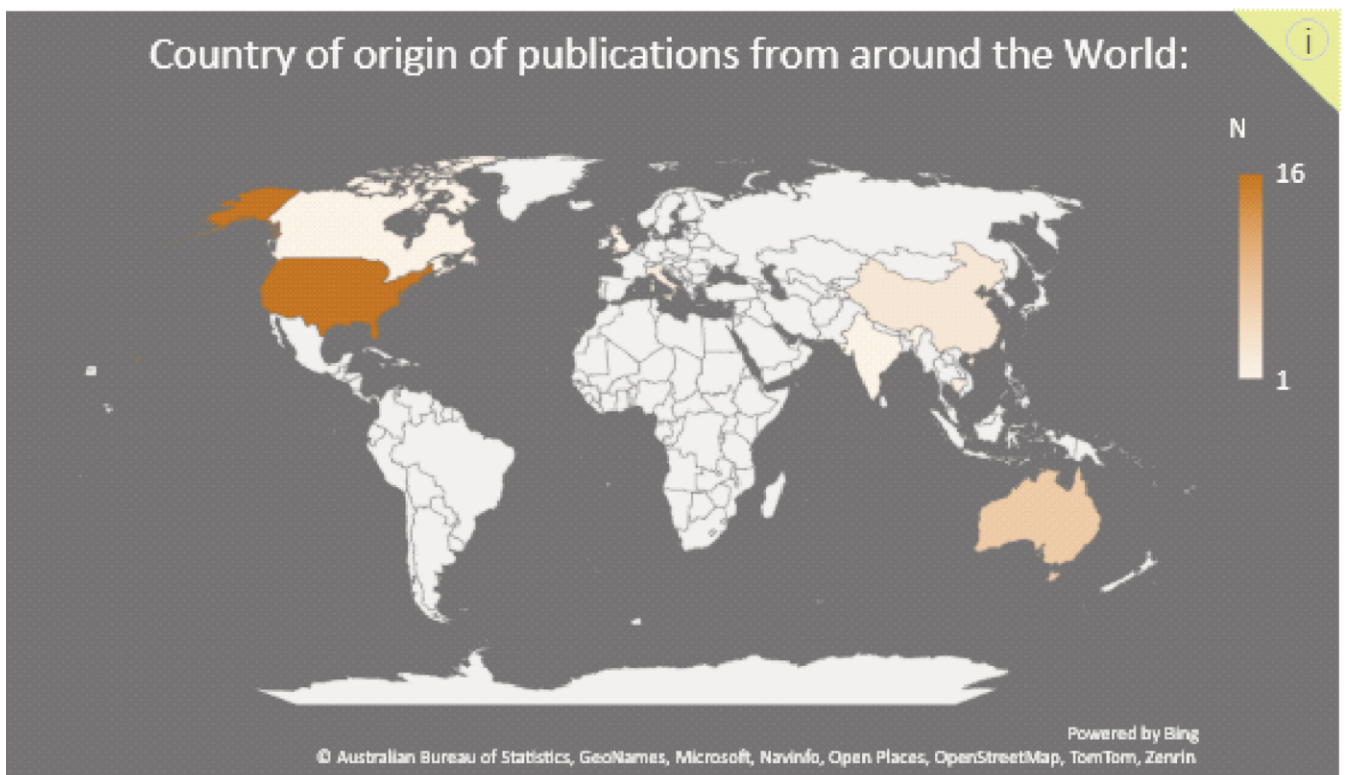


Figure-2: World Map showing country of origin of publications:

Table- 2a:Characteristics of Articles:

Sr no	Author year	No. of authors	Citations	Country	Aim	Methodology	No. of Prompt	Conclusion:	Name of Journal
1.	Gupta et al ⁴ . (2023).	8	21	USA	To assess the uniqueness of ChatGPT in generating systematic review ideas about cosmetic surgery.	ChatGPT was prompted to generate 20 unique ideas for systematic review for 12 topics of cosmetic surgery. These topics generated as a response by ChatGPT were assessed for their accuracy and uniqueness by conducting a literature review with PubMed, CINAHL, EMBASE, and Cochrane.	-	-ChatGPT generated 55% accurate novel ideas for systematic review, therefore can be used as a helping tool for innovation in research.	Aesthetic surgery journal
2.	Xie et al ¹¹ . (2023)	6	11	Australia	To evaluate ChatGPT responses to hypothetical questions in simulation of an initial rhinoplasty consultation.	Four Plastic Surgeons assessed the responses of ChatGPT to nine questions published by the American Society of Plastic Surgeons on rhinoplasty.	9	-ChatGPT provided simple answers to the health-specific queries.—Unable to provide detailed or personalized advice, based on individual scenario.	Aesthetic Plastic Surgery
3.	Ali MJ ¹² . (2023)	1	5	India	This study aimed to evaluate the knowledge of ChatGPT about lacrimal drainage disorders.	Three lacrimal Surgeons evaluated the responses of ChatGPT on lacrimal drainage disorders, qualitatively and graded into three categories: correct, partially correct, and factually incorrect.	21	-Inaccurate medical fact statement by ChatGPT, without any evidence.- Overall average knowledge about lacrimal disorders.- Need to be retrained for different medical subspeciality.	Ophthalmic plastic and reconstructive surgery
4.	Freedman et al ¹³ . (2023)	2	2	USA	The aim of this study was to determine GPT-4's proficiency in the field of plastic and reconstructive surgery using PSITE MCQs from 2 years.	"Few-shot learning," method was used, in which, for each question from the PSITE exam, GPT-3.5 was given a prompt containing five sample PSITE questions with their multiple options and correct answers with	Not specified	-GPT-4 has shown improved competency and accuracy over GPT-3.5, in answering PSITE MCQs exam.	ArXiv

Continued on next page...

Continued from previous page...

						explanations. Followed by this, the original PSITE exam question and its multiple-choice options, without the correct answer was prompted to language model and in last follow-up prompt was given to explain its answer to the question.			
5.	Seth et al ¹⁴ . (2023)	6	-	Australia	The aim of this study was to evaluate the accuracy and quality of information provided by ChatGPT's for the purpose of research writing in the plastic surgery field.	Three Plastic Surgeons scrutinized the responses of ChatGPT to five questions on 'base of thumb arthritis.'	5	- ChatGPT-3 provided a brief summarized answers of the asked questions.-Minimal grammatical error and good sentence structure.-Generated fake and incorrect references.	Plastic and reconstructive surgery. Global open
6.	Topsakal et al ¹⁵ . (2023)	4	-	USA	The aim to develop unique software tools with suitable algorithms to perform thorough analysis using area and volume measurements of face subsection. Evaluation of ChatGPT in generating the code required for calculating the area and volume on a 3D face model was also done.	In this study, two algorithms were introduced along with their freely available source code, designed to calculate the surface area and volume of a specific part within a 3D model. Additionally, ChatGPT was used to generate code, and the code generated was based on the approach of leveraging vertex data from the 3D model file for calculating area and volume, similar to second algorithm.	>1 (Not specified)	-ChatGPT was able to generate code to compute the area and volume of regions on a 3D face model, but the results were incorrect on application of this code.	IEEE Access
7.	Seth et al ¹⁶ . (2023).	4	-	Australia	This study evaluated ChatGPT's potential to contribute in future oculoplastic research.	Four physicians with experience in oculoplastic research reviewed the	3	-ChatGPT provided relevant superficial responses for complex questions.-	Annals of Ophthalmology and visual Sciences

Continued on next page...

Continued from previous page...

						responses of ChatGPT about future innovations in oculoplastic research.		Source of innovation and new ideas in field of oculoplasty for the general public.	
8.	Xie et al ¹⁷ . (2023)	4	-	Australia	The authors aim to determine ChatGPT aptitude about medical information on breast reconstruction options, and its ability to determine research ideas.	Two Plastic Surgeons qualitatively assessed the responses of ChatGPT to six questions on post mastectomy breast reconstruction, on Likert framework structured from 1 (strongly disagree) to 5 (strongly agree).	6	-ChatGPT provided relevant, accurate but superficial answers in response.-Generated fake and incorrect references.	Aesthetic Plastic Surgery

Arranged in descending order of citation numbers.

Table- 2b: Characteristics of miscellaneous types of articles:

Sr no	Authors, year	No. of authors	Citation	Country	Comments	Name of Journal
1.	Gupta et al ¹⁸ . (2023)	7	19	USA		Aesthetic surgery journal
2.	Najafali et al ¹⁹ . (2023)	5	11	USA		Aesthetic surgery journal
3.	Gupta et al ²⁰ . (2023)	6	9	USA		Journal of plastic, reconstructive & aesthetic surgery
4.	Najafali et al ²¹ . (2023)	6	7	USA		Aesthetic surgery journal
5.	Cox et al ²² . (2023)	5	6	Australia		Aesthetic surgery journal
6.	Gupta et al ²³ . (2023)	3	4	USA	Response to LTE ²¹	Aesthetic surgery journal
7.	Najafali et al ²⁴ . (2023)	6	4	USA		Aesthetic surgery journal
8.	Gupta et al ²⁵ . (2023)	6	3	USA		Aesthetic surgery journal
9.	Najafali et al ²⁶ . (2023)	6	2	USA		Aesthetic surgery journal
10.	Sun et al ²⁷ . (2023)	5	1	China		Aesthetic surgery journal
11.	Zhou et al ²⁸ . (2023)	4	1	China		Aesthetic surgery journal
12.	Kleebayoon et al ²⁹ . (2023)	2	1	Cambodia		Aesthetic surgery journal
13.	Buzzaccarini et al ³⁰ . (2023)	3	1	Italy		Aesthetic surgery journal
14.	Xie et al ³¹ . (2023)	5	-	Australia	Response to LTE ²⁹	Aesthetic surgery journal
15.	Seth et al ³² . (2023)	3	-	Australia	Response to LTE ³⁰	Aesthetic surgery journal
16.	Kleebayoon et al ³³ . (2023)	2	-	Cambodia	Response to LTE ²⁰	Journal of plastic, reconstructive & aesthetic surgery
17.	Tay et al ³⁴ . (2023)	1	-	UK		European Journal of Plastic Surgery
18.	ElHawary et al ³⁵ . (2023)	3	-	USA		Plastic and reconstructive surgery. Global open
19.	Kang et al ³⁶ . (2023)	3	-	China		Aesthetic surgery journal
20.	Lanzano et al ³⁷ . (2023)	1	-	Italy		Aesthetic surgery journal
21.	Sue et al ³⁸ . (2023)	1	-	USA		Plastic and reconstructive surgery.
22.	Abi-Rafeh et al ³⁹ . (2023)	3	-	Canada		Aesthetic surgery journal
23.	William et al ⁴⁰ . (2023)	3	-	USA		Journal of reconstructive microsurgery
24.	Cress et al ⁴¹ . (2023)	1	-	USA	Editorial	Aesthetic surgery journal
25.	Weidman et al ⁴² . (2023)	4	-	USA	Editorial	Plastic and reconstructive surgery.

LTE=Letter to the Editor. Arranged in descending order of citation numbers..

Characteristics of included studies are shown in table no. 2a and table no.2b. The range of citations for original articles was 2-21, whereas for letters to the editor, it was 1-19. No citations were noted for 4 original articles and 12 letters to the editors/editorials. On average, there were

4.4 authors per original manuscript and 3.8 authors per letter to the editor/editorial. Among the letters to the editors, four were responses to previous letters (Table no. 2b). The manuscript with the highest number of 8 co-authors and the highest number of citations was related

Table- 3: Top articles with highest no. of citations

Sr no	Authors, year	Title of publication	Type	No. of citations
1.	Gupta et al 4. (2023).	Expanding Cosmetic Plastic Surgery Research Using ChatGPT.	Article	21
2	Gupta et al 18 . (2023).	Application of ChatGPT in Cosmetic Plastic Surgery: Ally or Antagonist?	LTE	19
3	Xie et al 11. (2023).	Aesthetic Surgery Advice and Counselling from Artificial Intelligence: A Rhinoplasty Consultation with ChatGPT.	Article	11
4	Najafali et al 19 . (2023)	Let's Chat About Chatbots: Additional Thoughts on ChatGPT and Its Role in Plastic Surgery Along With Its Ability to Perform Systematic Review.	LTE	11

LTE=Letter to the Editor.

to aesthetic surgery research, cited 21 times, was same⁴. (Table no. 3)

Chat GPT in Scientific Writing, Academia, and Plastic Surgery Research:

Gupta et al, found that ChatGPT can generate novel systematic review ideas, can give relevant responses to specific prompts.⁴ In another study, Chat GPT-4 demonstrated notable enhancements and superior performance in comparison to GPT-3.5 as use of ChatGPT-4 for plastic surgery exam, led to a significant boost in scores, elevating them from the 8th to the 88th percentile and from the 3rd to the 99th percentile for the years 2021 and 2022, respectively.¹³

ChatGPT can provide summarized response with minimal grammatical or structural errors and only 1% of plagiarism was detected in the responses¹⁴. Topsakal et al; demonstrated that ChatGPT is capable of generating code to compute the area and volume of regions on a 3D face model, using multiple prompts.¹⁵ ChatGPT has understanding for oculoplastic research and can conduct thorough literature reviews, proposing novel experiments, improving surgical results through personalized surgical planning and patient communication within this domain.¹⁶

Highlighting the implications of ChatGPT in plastic surgery research, Najafali et al.²¹ reported its ability to perform PRISMA systematic review and meta-analyses from its own dataset. Similarly, other researchers^{18,20,21,25,3} mentioned that it can only generate ideas for systematic reviews at this stage. In another letter, Najafali²⁴ discussed that it may help researchers in generating ideas for their grant application and can expedite review process of grant applications. According to few researchers^{26,40}, the accessibility and open-access aspect of ChatGPT make it an attractive information and educational resource for both patients and healthcare professionals, whereas other authors^{35,40} referred it to as an assistive tool in manuscript writing allowing researchers to focus more on hypotheses and methodology part. It can also be utilized as an English language assistant for authors whose primary language is

not English. Zhou et al.²⁸ reported a unique use of extracting keywords of medical literature from ChatGPT, that were unique from those, found after bibliometric analysis, and were more representative and operational for scientific research.

Chat GPT in Plastic Surgery clinical practice:

ChatGPT can provide us with personalized consultations without having inpatient visits and can also generate educational and marketing materials like social media posts and blogs. Furthermore, it can also remember previous conversations and can be customized to cater to different age groups within the population⁴. Xie et al; performed a study to assess ChatGPT role in clinical settings and consultation for rhinoplasty and reported that this AI model provided concise basic medical and surgical information, gives a disclaimer to be examined by the surgeon, provides patients with information about the surgeon's credentials and qualification, and can assist in pre-operative planning. Furthermore, it has the potential to enhance healthcare in regions where algorithmic decision-making plays a pivotal role, including fields like diagnostics, resource allocation, and data collection.¹¹

Ali MJ; reported that it acknowledges errors and rectifies them when confronted with opposing viewpoints.¹² Another study, reported that ChatGPT can provide accurate basic information, excels in summarizing established knowledge, demonstrates understanding of standard post-mastectomy reconstructive procedures, and refrains from employing medical terminology.¹⁷

Buzzaccarini et al.³⁰ reported that ChatGPT has precise, dependable and accessible information of various aesthetic procedures. Its artificial intelligence algorithms can provide virtual consultations, preoperative planning, patient education, and postoperative care, tailored recommendations for treatment plans by reviewing patient data, including medical records. It can be used as an adjunct to clinical decision support to patients and healthcare providers, as indicated by various authors.^{18,20,31} In general, ChatGPT-4 swiftly offers safe medical guidance which can complement consultation

and encourages users to engage in discussions with their plastic surgeons as well.²² Sun et al, reported GPT-4 responses correct and good enough for the purpose of aesthetic consultation.²⁷ Kang et al. highlighted that besides providing basic answers to the patient queries and improving doctor-patient communication, it provided timely management about situations like burns and wounds, in less developed areas.³⁶ Lanzano et al. conducted an assessment and found that ChatGPT provided professional, comprehensive, and scientific communication with patients about breast reconstruction.³⁷ Sue et al. discovered that it can enhance diagnostic accuracy, facilitate surgical planning, improve patient outcomes based on algorithms, and can enable personalized treatments.³⁸

Limitations of ChatGPT

Authors have raised several ethical concerns related to the use of ChatGPT in academic, research, and clinical settings. These limitations encompass:

Inaccurate and incorrect responses to certain prompts,^{4,25,29,40} a high likelihood of generating plagiarized content,^{1,19,29,34,40} lack of references or fake/fabricated references,^{3,4,11,19,26} and inability to draft comprehensive manuscripts or conduct thorough research²¹. It also leads to breach of data privacy and security^{30,32}, lack of accountability for the content it generates and failure to meet the International Committee of Medical Journal Editors (ICMJE)⁵⁶ criteria for authorship.^{4,19,23,29} Furthermore, ChatGPT lacks the capability to compose aim pages for grant proposals.²⁴ AI dataset biases may potentially impact on its effectiveness in clinical applications by response generated on biased AI algorithms.^{4,8,26,32,34,35,38,40}

Additionally, there is an increased risk of losing the personal connection between patients and healthcare providers, resulting in an inability to convey empathy and demonstrate psychological competence.^{27,30,43} ChatGPT is incapable of providing real-time data and knowledge is outdated, extending only up to September 2021. Consequently, its responses may not align with the latest consensus or clinical guidelines.^{1,3,4,8,22,26,27,40} The dataset used for training this model does not include privately restricted data.²⁷ Most of its medical responses are general recommendations rather than personalized advice for the users^{43,22,27} and are not evidence based.⁴ Furthermore, GPT-4's responses were found to be superficial and inadequate when addressing complex questions.^{11,25,27,39,40}

Authors have labelled that this language model is not specialized for medicine as it responds to harmful

commands and tends to produce repetitive content⁴, and they strongly recommend that every information should be verified by a professional plastic surgeon.^{4,26,27,37,40} Therefore, ChatGPT can only serve as a supplementary tool for cosmetic consultations^{27,43} rather than a replacement for human decision-making⁸ or as an author for research paper writing.^{3,16} Additionally, one study discovered that while it can generate code for task execution, the results achieved using this generated code were incorrect when compared to other mesh (vertex and polygon) algorithm processing methods.¹⁵

Discussion

ChatGPT, is a part of the GPT family of language-based models, with the first model (GPT-1) initially being released in 2018.⁴³ It is an emerging language model, based on both supervised and reinforcement learning techniques, using human trainers for both approaches.

The above results reveal that majority of the article types that were published were letters to the editor. Despite AI technologies being globally accessible, their adoption and prominence differs significantly between Eastern and Western countries. This may be due to local needs, policies, and preferences and it is reflected in this review that abundance of studies originated from developed countries compared to the East. Furthermore, it is noteworthy that three out of the top four articles with the highest citation counts pertain to aesthetic surgery and cosmetic procedures. This underscores the high interest of the users in aesthetic surgeries and willingness of plastic surgeons to integrate artificial intelligence into their clinical practice.

This study elaborates the role of ChatGPT in the field of plastic surgery research, academia, and clinical practice by reviewing the available literature, along with the evaluation of both the strengths and limitations of utilizing this Large Language Model (LLM) in healthcare settings. Since the release of ChatGPT, several types of research papers documenting its implications in different fields of medicine and surgery, have been published, including original studies, case reviews, letters to the editors, and editorials. This study is distinctive in that it analysed all kinds of manuscripts related to Plastic Surgery.

There were notable achievements reported by the authors in the application of ChatGPT for plastic and aesthetic surgery academia and scientific writing. Observations reported by the authors included its capability to suggest unique review ideas, perform literature review, keywords extraction, data analyses, manuscript writing in English with no grammatical error

and avoidance of plagiarism. One case report claimed that text generated as a response in ChatGPT is unique and free of plagiarism.⁴⁴ Interestingly, other authors have also asserted its effectiveness in preventing plagiarism, all while maintaining a high level of grammatical and structural accuracy.^{5,45} Authors from non-English speaking countries often use language editing services, but ChatGPT offers free, quick, and refined editing. This is a blessing for researchers from developing countries whose first language is not English as ChatGPT can be solely used for scientific article language editing without issue after mentioning in acknowledgement³⁵. This helps in significant reduction of added financial burden related to scientific publications. Moreover, these findings suggest that ChatGPT can play a crucial role in assisting a researcher from the beginning to the end i.e. from generating ideas, keywords to final proofread.

Various studies in other medical specialties have been conducted regarding ChatGPT's utility in educational writing and scientific research, only to find similar observations.^{46,47,48} ChatGPT responses were tailored to the provided prompts and offered fundamental medical information while skilfully avoiding the use of intricate medical terminology. This capability renders it valuable for creating educational materials and promoting social awareness and again without adding any cost, it is free to use.

Its potential within academic settings, extends to assisting in the generation of code for studying 3D facial models, as well as aiding in preparation for plastic surgery examinations. Few authors reported that correct code generation required human prompting various times¹⁵. Other than plastic surgery examination, ChatGPT's abilities have been previously tested against exams like USMLE⁴⁹, in which it cleared the exam but did not score a high grade. Another study reported, it scores below the threshold for American Heart Association (AHA), Basic Life Support (BLS) and Advanced Cardiovascular Life Support (ACLS) exams. However, alternatively author referred it as a great self-learning tool for the preparation of life support exams⁵⁰. Similarly, this has been reported for numerous other academic examinations^{51,52}. This highlights that it can be used as an academic coach for preparation of medical exams, to practice, learn and clear confusions by medical trainees.

ChatGPT is available 24/7 around the clock and could help surgeons collect and analyse patients' medical history, provide virtual consultation, laboratory findings, and imaging features to provide a personalized surgical plan. The implementation of this feature can reduce the clinic burden on outpatients, by offering flexible consultation

timings in the comfort of their homes, and to people worldwide across different time zones. Also, it can be used to streamline the process of guiding patients to the appropriate department, whether it be for surgery or medicine, based on the investigation results, advised after consultation using chatbot. Additionally, it has the potential to save patients valuable time by reducing the need to wait on lengthy consultation queues.

Park explored the potential role of ChatGPT in the field of clinical otolaryngology, and discussed its potential clinical applications, such as in clinical decision-making, delivering personalized care, and performing numerous administrative tasks such as insurance communication, billing, and ICD coding. However, certain limitations he mentioned includes the inability to conduct physical examinations, ethical considerations regarding transparency, potential bias, and accountability for the Different authors of included manuscripts have articulated these limitations, such as inaccurate responses along with a high susceptibility to copyright plagiarism, repetitive content and authorship.^{1,53,54} Furthermore, as discussed in result section, ChatGPT is known to generate fake references, incorrect citations, and on some occasions respond to harmful commands.⁵⁴ Similar findings were reported by the authors of the studies included in our review. These can be misleading and may delay the diagnosis of disease/ condition among the users. Also, it can become a source of wrong medical information because of fake references, that can be used by fraudulent companies to promote unverified treatments. Given that this program isn't specifically designed for medical use, and only provides basic level of medical advice, it necessitates verification of all information by a professional healthcare provider.

Major limitation lies in its deficiency in psychological competency, rendering it incapable of generating empathy or human emotions and outdated knowledge, with the most recent information extending only up to September 2021, thus lacking access to updated consensus or recent clinical guidelines for management of diseases. Similar deficiencies have been reported by researchers of other specialties too.

When discussing the clinical capabilities of ChatGPT, ChatGPT's clinical application raises ethical concerns related to beneficence, nonmaleficence, physician integrity, and justice. It can predict diagnoses and advise treatments, reducing human error and knowledge gaps in medicine, however reliance on it without clinical judgment may jeopardize patient well-being, compromising beneficence and non-maleficence. Patients may view it as a replacement for physicians,

undermining physician integrity. However, regarding justice, ChatGPT's free accessibility could improve medical information access for patients.⁵⁵ Major concern is that it may not always deliver patient-specific, medically accurate responses, possibly leading to harm for patients.

The preceding discussion highlights several important points. Firstly, it is evident that the willingness of the researchers to incorporate this language model into clinical practice hinges on the establishment of a universally accepted code of conduct or guideline about use of Chatbots.^{29,31,33,34,35,36} On the other hand, some authors express reservations about the use of artificial intelligence, fearing that it may diminish human creativity, which forms the foundation of plastic and reconstructive surgery.^{14,34,39,46}

Conclusion

In summary, ChatGPT can provide assistance in plastic surgery research and clinic practice, but there are concerns and limitations. Most of these studies originated from developed countries, raising concerns like bias related to publications from developed countries. Moreover, ethical issues regarding its clinical use and its potential impact on human creativity in surgery, is still debatable. Also, this highlights the need of universal guidelines for using language models like ChatGPT in healthcare.

Limitation: This systematic review was not registered in PROSPERO. Additionally, the review is confined to English-language articles and there might be omission of important information from the articles that were excluded due to unavailability of full text or were in language other than English..

Disclaimer: None.

Conflict of Interest: None.

Funding Disclosure: None.

References

- Macdonald C, Adeloje D, Sheikh A, Rudan I. Can ChatGPT draft a research article? An example of population-level vaccine effectiveness analysis. *J Glob Health* 2023;13:01003. doi: 10.7189/jogh.13.01003.
- Altmäe S, Sola-Leyva A, Salumets A. Artificial intelligence in scientific writing: a friend or a foe? *Reprod Biomed Online* 2023;47:3-9. doi: 10.1016/j.rbmo.2023.04.009.
- Stokel-Walker C. AI bot ChatGPT writes smart essays - should professors worry? *Nature* 2022. doi: 10.1038/d41586-022-04397-7. [ahead of print]
- 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;43:930-7. doi: 10.1093/asj/sjad069.
- Hutson M. Could AI help you to write your next paper? *Nature* 2022;611:192-3. doi: 10.1038/d41586-022-03479-w.
- Cheng K, Sun Z, He Y, Gu S, Wu H. The potential impact of ChatGPT/GPT-4 on surgery: will it topple the profession of surgeons? *Int J Surg* 2023;109:1545-7. doi: 10.1097/JS9.0000000000000388.
- Park I, Joshi AS, Javan R. Potential role of ChatGPT in clinical otolaryngology explained by ChatGPT. *Am J Otolaryngol* 2023;44:103873. doi: 10.1016/j.amjoto.2023.103873.
- Else H. Abstracts written by ChatGPT fool scientists. *Nature* 2023;613:423. doi: 10.1038/d41586-023-00056-7.
- Mogali SR. Initial impressions of ChatGPT for anatomy education. *Anat Sci Educ* 2024;17:444-7. doi: 10.1002/ase.2261.
- PRISMA. The PRISMA declaration methodology for systematic reviews. [Online] 2023 [Cited 2023 September 22]. Available from URL: <http://prisma-statement.org/prismastatement/flowdiagram.aspx>.
- Xie Y, Seth I, Hunter-Smith DJ, Rozen WM, Ross R, Lee M. Aesthetic Surgery Advice and Counseling from Artificial Intelligence: A Rhinoplasty Consultation with ChatGPT. *Aesthetic Plast Surg* 2023;47:1985-93. doi: 10.1007/s00266-023-03338-7.
- Ali MJ. ChatGPT and Lacrimal Drainage Disorders: Performance and Scope of Improvement. *Ophthalmic Plast Reconstr Surg* 2023;39:221-5. doi: 10.1097/IOP.0000000000002418.
- Freedman JD, Nappier IA. GPT-4 to GPT-3.5: 'Hold My Scalpel' -- A Look at the Competency of OpenAI's GPT on the Plastic Surgery In-Service Training Exam. [Online] 2023 [Cited 2023 September 22]. Available from URL: <https://arxiv.org/abs/2304.01503v1>
- Seth I, Sinkjær Kenney P, Bulloch G, Hunter-Smith DJ, Bo Thomsen J, Rozen WM. Artificial or Augmented Authorship? A Conversation with a Chatbot on Base of Thumb Arthritis. *Plast Reconstr Surg Glob Open* 2023;11:e4999. doi: 10.1097/GOX.0000000000004999.
- Topsakal O, Sawyer P, Akinci TC, Celikoyar MM. Algorithms to measure area and volume on 3D face models for facial surgeries. *IEEE Access* 2023;11:39577-85. Doi: 10.1109/ACCESS.2023.3268174.
- Seth I, Bulloch G, Xie Y, Zhu Z. Exploring the potential of ChatGPT for advancing oculoplastic surgical research. *Ann Ophthalmol Vis Sci* 2023;6:1038.
- Xie Y, Seth I, Rozen WM, Hunter-Smith DJ. Evaluation of the Artificial Intelligence Chatbot on Breast Reconstruction and Its Efficacy in Surgical Research: A Case Study. *Aesthetic Plast Surg* 2023;47:2360-9. doi: 10.1007/s00266-023-03443-7.
- Gupta R, Pande P, Herzog I, Weisberger J, Chao J, Chaiyasate K, et al. Application of ChatGPT in Cosmetic Plastic Surgery: Ally or Antagonist? *Aesthet Surg J* 2023;43:NP587-90. doi: 10.1093/asj/sjad042.
- Najafali D, Reiche E, Camacho JM, Morrison SD, Dorafshar AH. Let's Chat About Chatbots: Additional Thoughts on ChatGPT and Its Role in Plastic Surgery Along With Its Ability to Perform Systematic Reviews. *Aesthet Surg J* 2023;43:NP591-2. doi: 10.1093/asj/sjad056.
- Gupta R, Herzog I, Weisberger J, Chao J, Chaiyasate K, Lee ES. Utilization of CHATGPT for plastic surgery research: Friend or foe? *J Plast Reconstr Aesthet. Surg* 2023;80:145-7. doi:10.1016/j.bjps.2023.03.004
- Najafali D, Camacho JM, Reiche E, Galbraith LG, Morrison SD, Dorafshar AH. Truth or Lies? The Pitfalls and Limitations of ChatGPT in Systematic Review Creation. *Aesthet Surg J* 2023;43:NP654-5. doi: 10.1093/asj/sjad093.
- Cox A, Seth I, Xie Y, Hunter-Smith DJ, Rozen WM. Utilizing ChatGPT-4 for Providing Medical Information on Blepharoplasties to Patients. *Aesthet Surg J* 2023;43:NP658-62. doi: 10.1093/asj/sjad096.
- Gupta R, Bagdady K, Mailey BA. Response to: Truth or Lies? The Pitfalls and Limitations of ChatGPT in Systematic Review Creation. *Aesthet Surg J* 2023;43:NP656-7. doi: 10.1093/asj/sjad108.

24. Najafali D, Hinson C, Camacho JM, Galbraith LG, Gupta R, Reid CM. Can Chatbots Assist With Grant Writing in Plastic Surgery? Utilizing ChatGPT to Start an R01 Grant. *Aesthet Surg J* 2023;43:NP663-5. doi: 10.1093/asj/sjad116.
25. Gupta R, Herzog I, Najafali D, Firouzbakht P, Weisberger J, Mailey BA. Application of GPT-4 in Cosmetic Plastic Surgery: Does Updated Mean Better? *Aesthet Surg J* 2023;43:NP666-9. doi: 10.1093/asj/sjad132.
26. Najafali D, Camacho JM, Galbraith LG, Reiche E, Dorafshar AH, Morrison SD. Ask and You Shall Receive: OpenAI ChatGPT Writes Us an Editorial on Using Chatbots in Gender Affirmation Surgery and Strategies to Increase Widespread Adoption. *Aesthet Surg J* 2023;43:NP715-7. doi: 10.1093/asj/sjad119.
27. Sun YX, Li ZM, Huang JZ, Yu NZ, Long X. GPT-4: The Future of Cosmetic Procedure Consultation? *Aesthet Surg J* 2023;43:NP670-2. doi: 10.1093/asj/sjad134.
28. Zhou J, Jia Y, Qiu Y, Lin L. The Potential of Applying ChatGPT to Extract Keywords of Medical Literature in Plastic Surgery. *Aesthet Surg J* 2023;43:NP720-3. doi: 10.1093/asj/sjad158.
29. Kleebayoon A, Wiwanitkit V. Rhinoplasty Consultation with ChatGPT. *Aesthetic Plast Surg* 2023;47:2203-4. doi: 10.1007/s00266-023-03394-z.
30. Buzzaccarini G, Degliuomini RS, Borin M. The Artificial Intelligence application in Aesthetic Medicine: How ChatGPT can Revolutionize the Aesthetic World. *Aesthetic Plast Surg* 2023;47:2211-2. doi: 10.1007/s00266-023-03416-w.
31. Xie Y, Seth I, Hunter-Smith DJ, Rozen WM, Ross RJ, Lee M. Response to "Rhinoplasty Consultation with ChatGPT" by Kleebayoon et al. *Aesthetic Plast Surg* 2023;47:2205-6. doi: 10.1007/s00266-023-03417-9.
32. Seth I, Xie Y, Rozen WM. Response to "The Artificial Intelligence Application in Aesthetic Medicine: How ChatGPT can Revolutionize the Aesthetic World" by Buzzaccarini et al. *Aesthetic Plast Surg* 2023;47:2213-4. doi: 10.1007/s00266-023-03439-3.
33. Kleebayoon A, Wiwanitkit V. Utilization of ChatGPT for plastic surgery research: Comment. *J Plast Reconstr Aesthet Surg* 2023;82:275. doi: 10.1016/j.bjps.2023.04.084.
34. Tay JQ. ChatGPT and the future of plastic surgery research: evolutionary tool or revolutionary force in academic publishing?. *Eur J Plast Surg* 2023;46:643-4. doi: 10.1007/s00238-023-02081-1.
35. ElHawary H, Gorgy A, Janis JE. Large Language Models in Academic Plastic Surgery: The Way Forward. *Plast Reconstr Surg Glob Open* 2023;11:e4949. doi: 10.1097/GOX.0000000000004949.
36. Kang Y, Xia Z, Zhu L. When ChatGPT Meets Plastic Surgeons. *Aesthetic Plast Surg* 2023;47:2190-3. doi: 10.1007/s00266-023-03372-5.
37. Lanzano G. Harnessing the Potential of ChatGPT in Breast Reconstruction: A Revolution in Patient Communication and Education. *Aesthetic Plast Surg* 2023;47:2215-6. doi: 10.1007/s00266-023-03427-7.
38. Sue GR. Artificial Intelligence for Plastic Surgeons. *Plast Reconstr Surg Glob Open* 2023;11:e5057. doi: 10.1097/GOX.0000000000005057.
39. Abi-Rafeh J, Xu HH, Kazan R. Preservation of Human Creativity in Plastic Surgery Research on ChatGPT. *Aesthet Surg J* 2023;43:NP726-7. doi: 10.1093/asj/sjad162.
40. Tian WM, Sergesketter AR, Hollenbeck ST. The Role of ChatGPT in Microsurgery: Assessing Content Quality and Potential Applications. *J Reconstr Microsurg* 2024;40:e1-2. doi: 10.1055/a-2098-6509.
41. Cress PE. ChatGPT and the Honor System. *Aesthet Surg J* 2023;43:938-40. doi: 10.1093/asj/sjad163.
42. Weidman AA, Valentine L, Chung KC, Lin SJ. OpenAI's ChatGPT and Its Role in Plastic Surgery Research. *Plast Reconstr Surg* 2023;151:1111-3. doi: 10.1097/PRS.00000000000010342.
43. Kumar B. GPT-1, GPT-2 and GPT-3 models explained. [Online] 2023 [Cited 2024 February 15]. Available from URL: <https://360digitmg.com/blog/types-of-gpt-in-artificial-intelligence>
44. Rathore A, Reich DA, Qadir N, Harrison C. A Case Report of Culture-Negative Necrotizing Fasciitis With Low Laboratory Risk Indicator for Necrotizing Fasciitis (LRINEC) Score: A Diagnostic Dilemma. *Cureus* 2023;15:e37268. doi: 10.7759/cureus.37268.
45. Kim SG. Using ChatGPT for language editing in scientific articles. *Maxillofac Plast Reconstr Surg* 2023;45:13. doi: 10.1186/s40902-023-00381-x.
46. Borger JG, Ng AP, Anderton H, Ashdown GW, Auld M, Blewitt ME, et al. Artificial intelligence takes center stage: exploring the capabilities and implications of ChatGPT and other AI-assisted technologies in scientific research and education. *Immunol Cell Biol* 2023;101:923-35. doi: 10.1111/imcb.12689.
47. Lee SW, Choi WJ. Utilizing ChatGPT in clinical research related to anesthesiology: a comprehensive review of opportunities and limitations. *Anesth Pain Med (Seoul)* 2023;18:244-51. doi: 10.17085/apm.23056.
48. Khlaif ZN, Mousa A, Hattab MK, Itmazi J, Hassan AA, Sanmugam M, et al. The Potential and Concerns of Using AI in Scientific Research: ChatGPT Performance Evaluation. *JMIR Med Educ* 2023;9:e47049. doi: 10.2196/47049.
49. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Med Educ* 2023;9:e45312. doi: 10.2196/45312.
50. Fijačko N, Gosak L, Štiglic G, Picard CT, John Douma M. Can ChatGPT pass the life support exams without entering the American heart association course? *Resuscitation* 2023;185:109732. doi: 10.1016/j.resuscitation.2023.109732.
51. Wang X, Gong Z, Wang G, Jia J, Xu Y, Zhao J, et al. ChatGPT Performs on the Chinese National Medical Licensing Examination. *J Med Syst* 2023;47:86. doi: 10.1007/s10916-023-01961-0.
52. Huynh LM, Bonebrake BT, Schultis K, Quach A, Deibert CM. New Artificial Intelligence ChatGPT Performs Poorly on the 2022 Self-assessment Study Program for Urology. *Urol Pract* 2023;10:409-15. doi: 10.1097/UPJ.0000000000000406.
53. Liebreiz M, Schleifer R, Buadze A, Bhugra D, Smith A. Generating scholarly content with ChatGPT: ethical challenges for medical publishing. *Lancet Digit Health* 2023;5:e105-6. doi: 10.1016/S2589-7500(23)00019-5.
54. Curtis N. To ChatGPT or not to ChatGPT? The impact of artificial intelligence on academic publishing. *Pediatr Infect Dis* 2023;42:275. DOI: 10.1097/INF.0000000000003852.
55. Beltrami EJ, Grant-Kels JM. Consulting ChatGPT: Ethical dilemmas in language model artificial intelligence. *J Am Acad Dermatol*

- 2024;90:879-80. doi: 10.1016/j.jaad.2023.02.052.
56. International Committee of Medical Journal Editors (ICMJE). Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. [Online] 2024

[Cited 2024 February 26]. Available from URL: <https://icmje.org/icmje-recommendations.pdf>

DISCLAIMER

Articles published in JPMA and Supplements of JPMA do not represent the views of the editor or editorial Board. Authors are solely responsible for the opinions expressed and the accuracy of the data.
The contribution of each author towards the published research included in this supplement is the responsibility of the authors and their institutions. It is expected to be in accordance and compliance with the ICMJE Guidelines.
Any questions/queries raised by readers should be directed to the corresponding author.