Lim *et al. Plast Aesthet Res* 2023;10:43
DOI: 10.20517/2347-9264.2023.70

Plastic and Aesthetic Research

Case Report

Open Access

Evaluating the efficacy of major language models in providing guidance for hand trauma nerve laceration patients: a case study on Google's AI BARD, Bing AI, and ChatGPT

Bryan Lim[1,2], Ishith Seth[1,2,3], Gabriella Bulloch[3], Yi Xie[1], David J Hunter-Smith[1,2], Warren M Rozen[1,2]

[1]Department of Plastic Surgery, Peninsula Health, Melbourne 3199, Australia.
[2]Central Clinical School at Monash University, The Alfred Centre, Melbourne 3004, Australia.
[3]Faculty of Medicine and Surgery, The University of Melbourne 3053, Australia.

Correspondence to: Dr. Ishith Seth, Central Clinical School at Monash University, The Alfred Centre, 99 Commercial Rd, Melbourne 3004, Australia. E-mail: ishithseth1@gmail.com

How to cite this article: Lim B, Seth I, Bulloch G, Xie Y, Hunter-Smith DJ, Rozen WM. Evaluating the efficacy of major language models in providing guidance for hand trauma nerve laceration patients: a case study on Google's AI BARD, Bing AI, and ChatGPT. *Plast Aesthet Res* 2023;10:43. https://dx.doi.org/10.20517/2347-9264.2023.70

Received: 16 Jul 2023 **First Decision:** 3 Aug 2023 **Revised:** 5 Aug 2023 **Accepted:** 10 Aug 2023 **Published:** 17 Aug 2023

Academic Editor: Samuel O. Poore **Copy Editor:** Dan Zhang **Production Editor:** Dan Zhang

Abstract

This study evaluated three prominent Large Language Models (LLMs)-Google's AI BARD, Bing's AI, and ChatGPT-4 in providing patient advice for hand laceration. Five simulated patient inquiries on hand trauma were prompted to them. A panel of Board-certified plastic surgical residents evaluated the responses for accuracy, comprehensiveness, and appropriate sources. Responses were also compared against existing literature and guidelines. This study suggests that ChatGPT outperforms BARD and Bing AI in providing reliable, evidence-based clinical advice, but they still face limitations in depth and specificity. Healthcare professionals are essential in interpreting LLM recommendations, and future research should improve LLM performance by integrating specialized databases and human expertise to advance nerve injury management and optimize patient-centred care.

Keywords: Artificial intelligence, ChatGPT, BARD, Bings AI, large language model, nerve injury

INTRODUCTION

Exponential advancements in AI have catalyzed the emergence of LLMs, which are gaining interest for their abilities to synthesize large swaths of information and produce comprehensible answers to most queries[1,2].

ChatGPT has proven to answer complex medical queries, which could revolutionize healthcare for patients with limited access to medical professionals[1]. Hand trauma, specifically nerve lacerations, is common and can cause severe impairment if mismanaged[3]. Timely and accurate information is imperative but can be limited, particularly in rural and low-resource settings[4]. Opportunistically, AI-powered language models could fill this gap as potential medical guidance providers.

ChatGPT has dominated the mainstay of articles determining the suitability of LLMs in medical practice. With the launch of other AI tools, specifically Google's AI BARD and BingAI, their comparative performances should be evaluated. Therefore, this study evaluated Google's AI BARD, BingAI, and ChatGPT for providing accurate and relevant information to patients presenting with hand trauma nerve lacerations. Using objective and subjective metrics, we assessed contextual understanding and recommendation suitability. This comparison will highlight strengths and weaknesses, informing improvements in AI-driven guidance for hand trauma care.

CASE REPORT

The authors evaluated the suitability of three LLMs-Google's AI BARD, BingAI, and ChatGPT-4 by their capacity to interpret medical literature, extract relevant data, and produce precise, intelligible, and contextually suitable clinical advice. Their responses were also compared against established clinical guidelines. Additionally, the analysis will encompass the efficiency, dependability, potential biases, and ethical implications affiliated with each LLM within the realm of nerve injury management.

A set of simulated patient-perspective queries concerning digital nerve injury diagnosis and management were presented to ChatGPT, BARD, and Bing AI. Responses generated were compared to existing clinical guidelines and literature. Additionally, a panel of plastic surgery residents and Board-certified Plastic Surgeons with extensive peripheral nerve injury expertise evaluated the responses with a Likert scale. Assessment criteria included accuracy, comprehensiveness, and provision of relevant information sources.

To maintain consistency and precision, the same author (BL) documented the initial response each LLM provided for every question, avoiding further clarifications or alterations. Questions were crafted to avoid grammatical errors or ambiguity and were simultaneously inputted using separate accounts for OpenAI, Google, and Microsoft, granting access to ChatGPT-4, BARD, and Bing AI, respectively.

We compiled a dataset of hand trauma nerve laceration cases, encompassing diverse scenarios, symptoms, and treatment alternatives. Each language model's efficacy was assessed on a Likert scale [Table 1] based on its capacity to offer patient guidance accurately and effectively. Evaluation criteria comprised the subsequent elements:

Accuracy: the correctness and dependability of the information supplied by the language models.

Comprehensibility: the patients' ability to readily understand the provided information and directions.

Empathy and tone: the language models' capacity to convey empathy and sustain a suitable tone.

Table 1. Qualitative analysis results of large language models

Question	Language model	1-strongly disagree	2-disagree	3-neutral	4-agree	5-strongly agree
The language model provided accurate and reliable information on hand trauma nerve laceration	ChatGPT-4				×	
	Bing's AI			×		
	Google's BARD			×		
The information provided by the language model was easy to understand	ChatGPT-4				×	
	Bing's AI		×			
	Google's BARD				×	
The language model conveyed empathy and maintained an appropriate tone	ChatGPT-4				×	
	Bing's AI		×			
	Google's BARD					×
The language model provided relevant information quickly	ChatGPT-4			×		
	Bing's AI		×			
	Google's BARD				×	
Overall performance	ChatGPT-4				×	
	Bing's AI		×			
	Google's BARD				×	

Timeliness: the language models' speed in generating pertinent information.

Additionally, the readability and reliability of the LLMs' responses were evaluated using specific metrics. The Flesh Reading Ease Score, Flesch-Kincaid Grade Level, and the Coleman-Liau Index were employed for readability assessment, whereas the DISCERN score was utilized for reliability. The results are consolidated in Table 2 and subsequently subjected to a t-test [Table 3] for statistical significance appraisal.

Figure 1 introduced the first scenario to the LLMs, reading "Hi Large Language Model, I am a 23-year-old male who has cut his right index finger with a knife. I am right hand dominant and a professional guitar player. I do not have any sensation on my right index fingertips. What do you think has occurred and what treatment do I require?" ChatGPT's response, which highlighted the prompt in red and warned of content policy violations, immediately suggested consulting professionals for clinical advice[5]. Nonetheless, it identified the correct affected nerve and outlined damage mitigation steps[6]. Notably, ChatGPT considered the user's right-handedness and occupation, urging prompt expert help and personalized care. It concluded by discouraging risky activities and reiterating the need for professional assistance. Google's BARD presented a response between ChatGPT and Bing AI in quality, suggesting nerve injury and advising immediate physician consultation. It also recommended basic first-aid and concluded with insights into the injury's nature, prognosis, and potential therapies[7]. Bing AI proffered a comparable response, elucidating its primary diagnosis of nerve damage, advocating for professional consultation and delineating possible treatment methods. Unlike ChatGPT, Bing AI did not propose an intermediate care model.

Figure 2 aimed to evaluate the models' ability to follow up on previous queries (recall ability) with additional information. This read "Same patient as the previous question, what surgical procedure do you recommend? Do I require any diagnostic test prior to surgical intervention?" ChatGPT reemphasized its ill-suitedness to provide recommendations and encouraged users to consult a physician. Nonetheless, it enumerated diagnostic tests typically conducted for injury diagnosis and severity assessment, concluding its response by mentioning common treatment modalities[8,9]. BARD furnished a concise response, advocating for surgical repair and what it entailed, and listed identical diagnostic evaluations as ChatGPT did. It expounded on the repair prognosis being dependent on various factors, with earlier intervention often

Table 2. Readability and reliability scores comparing the three LLMs, including t-test analysis

		Readability		Reliability	
		Flesch reading ease score	**Flesch-Kincaid grade level**	**Coleman-Liau index**	**DISCERN score**
ChatGPT-4	No sensation on right fingertip	55.5	11.3	8	56
	Recommended surgical procedure	32.5	13.1	14	54
	Likely outcome of complete repair	37	12.4	14	45
	Timeframe of complete return	38.2	12.8	12	49
	Options aside from surgery	28.2	14.6	15	52
Mean		38.28	12.84	12.6	51.2
Standard deviation		10.41	1.20	2.79	4.32
Google's BARD	No sensation on right fingertip	51	10.5	9	47
	Recommended surgical procedure	35.2	13.2	12	40
	Likely outcome of complete repair	48.2	10.3	9	55
	Timeframe of complete return	41.6	13.7	10	48
	Options aside from surgery	41.2	12.2	13	55
Mean		43.44	11.98	10.60	49
Standard deviation		6.25	1.54	1.82	6.28
Bing's AI	No sensation on right fingertip	73	7	6	44
	Recommended surgical procedure	57.4	10	9	49
	Likely outcome of complete repair	NIL	NIL	NIL	NIL
	Timeframe of complete return	50.1	9.6	10	50
	Options aside from surgery	72.6	6.1	8	52
Mean		63.28	8.18	8.25	42.2
Standard deviation		11.40	1.92	1.71	14.9

Table 3. Student *T*-test analysis

***T*-test analysis**	**Flesch reading ease score**	**Flesch-Kincaid grade level**	**Coleman-Liau index**	**DISCERN score**
BARD *vs.* Bing AI	0.01	0.01	0.08	0.37
BARD *vs.* ChatGPT	0.01	0.002	0.03	0.23
Bing AI *vs.* ChatGPT	0.36	0.35	0.21	0.53

producing better outcomes. Augmenting ChatGPT's reply, BARD outlined postoperative rehabilitation strategies, encompassing orthotic support and physical therapy for functional recovery[10]. Unlike ChatGPT, BARD did not stress the importance of seeking expert counsel. Bing AI acknowledged its limitations by suggesting consultation with a professional and highlighting the importance of diagnostic assessments. However, it failed to delineate precise examinations and suitable therapeutic alternatives. Ultimately, it offered an indistinct summary compared to ChatGPT and BARD.

In Figure 3, the inquiry "If I have completely lacerated my digital nerve, what is the likely outcome of being completely repaired? Provide 5 high-level evidence studies to support your answer." sought to assess the models' capacity to supply pertinent references and predict surgical outcomes. ChatGPT delivered a vague

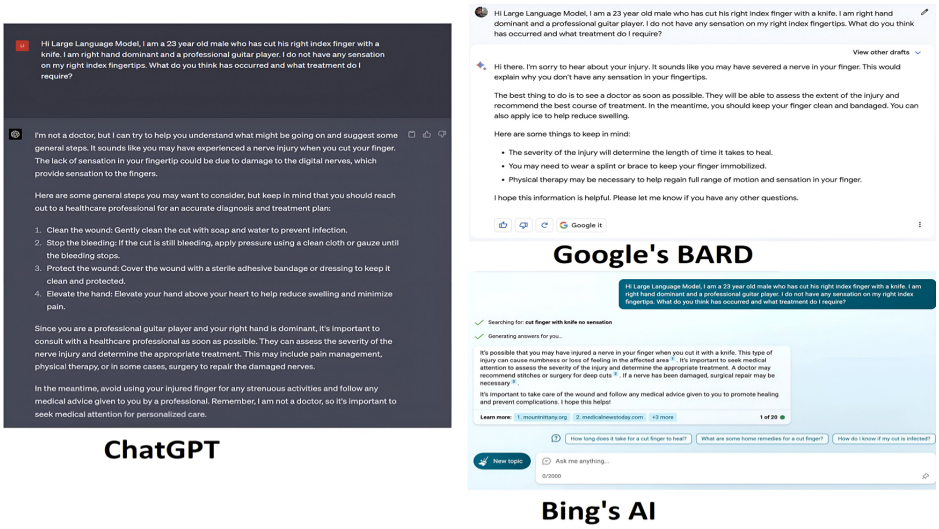

Figure 1. Hi Large Language Model, I am a 23-year-old male who has cut his right index finger with a knife. I am right hand dominant and a professional guitar player. I do not have any sensation on my right index fingertips. What do you think has occurred and what treatment do I require?

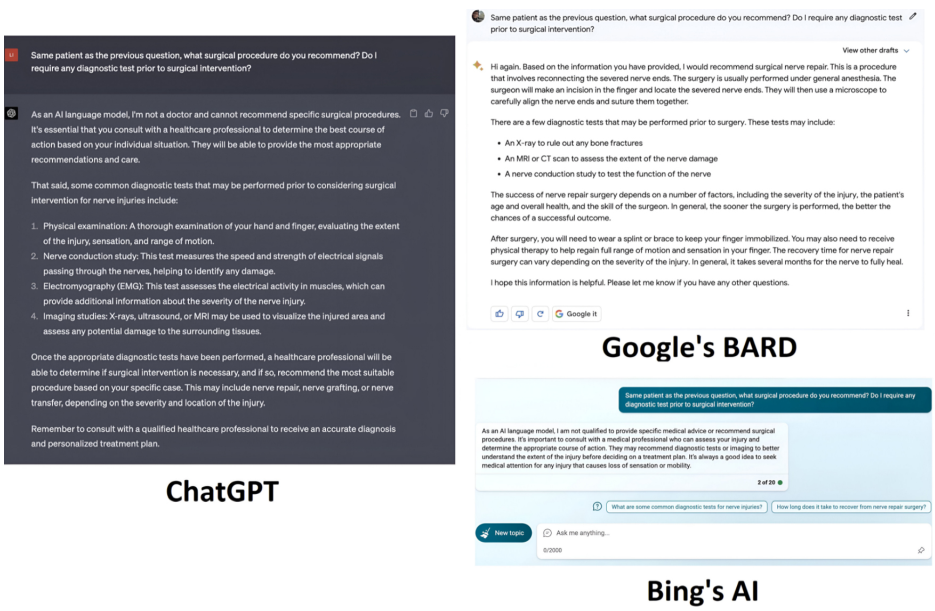

Figure 2. Same patient as the previous question, what surgical procedure do you recommend? Do I require any diagnostic test prior to surgical intervention?

response, stating that outcomes are generally favorable and contingent on factors such as injury severity, early surgical intervention and repair methodology[11]. The bulk of its reply addressed the query regarding high-level references, failing to provide five legitimate citations, and recommended users search reputable databases. It concluded by recommending consulting a healthcare professional. BARD merely asserted its inability to assist due to it being a language model, while Bing AI offered two systematic reviews and another article on recovery rates from nerve injuries, also advising for a healthcare consultation. The first two references, despite existing in the literature, share a primary author and cover similar subjects[12,13], while the third reference displayed the paper's title, omitting the authors and publication date. Furthermore, it

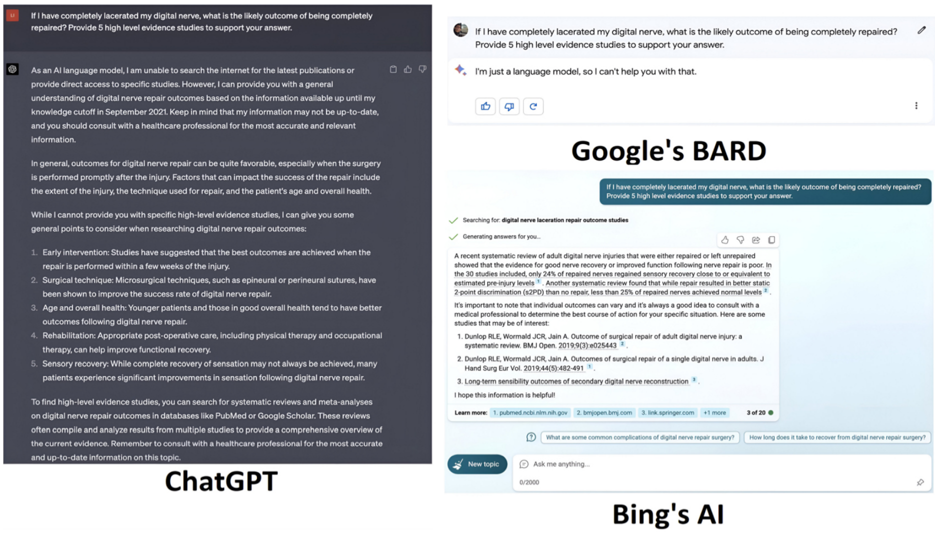

Figure 3. If I have completely lacerated my digital nerve, what is the likely outcome of being completely repaired? Provide 5 high-level evidence studies to support your answer.

failed to supply five references.

As shown in Figure 4, the prompt "Following digital nerve repair surgery, what is the timeframe for the complete return of sensation to the finger?" aimed to analyze the models' abilities to predict prognosis following a surgical procedure. ChatGPT commenced by noting the variable timeline is contingent on multiple factors. It delineated the average nerve regeneration rate, emphasizing the gradual process of regaining sensation, while also warning potential incomplete recovery[14,15]. ChatGPT finally advised adherence to healthcare professional guidance, encompassing hand therapy, sensory re-education, and pain management[16]. BARD generated a response akin to ChatGPT's, offering an approximate recovery timeframe of six months to two years, depending on injury severity and patient health, while similarly cautioning against incomplete sensation restoration[17,18]. It expounded on three factors influencing recovery duration and emphasized adhering to healthcare professional directives[19]. Bing AI proposed an alternative recovery timeframe of three to six months while cautioning that not all individuals achieve full recovery[3]. Interestingly, it cited a study to substantiate its assertion and furnished two additional links for users seeking further research.

The final input "Same patient, what are my options if I don't have surgery?" as shown in Figure 5, assessed the models' ability to offer a breadth of therapeutic options. Consistent with prior responses, ChatGPT advised seeking a healthcare professional's guidance. Initially, it elucidated potential complications arising from non-surgical approaches followed by enumerating alternative treatments[20]; however, it ultimately advocated for surgical management, attributing this to enhanced patient outcomes. BARD devised several recommendations for non-clinical settings, subsequently accentuating the potential inadequacy of these interventions and suggesting surgical measures should symptoms persist. The response culminated by listing the ramifications of forgoing surgery and urging users to consult their doctor for optimal management[20]. Bing AI presented and briefly explained distinct alternative therapies, underscoring the variability of patient outcomes and that not all nerve lacerations necessitate surgical intervention[21]. Bing AI also cited a systematic review to support its assertion and furnished at least four additional links for users to further study this topic. Unfortunately, the hyperlink to the systematic review was erroneous, raising concerns about Bing AI's credibility. It also failed to list possible complications of adopting non-surgical

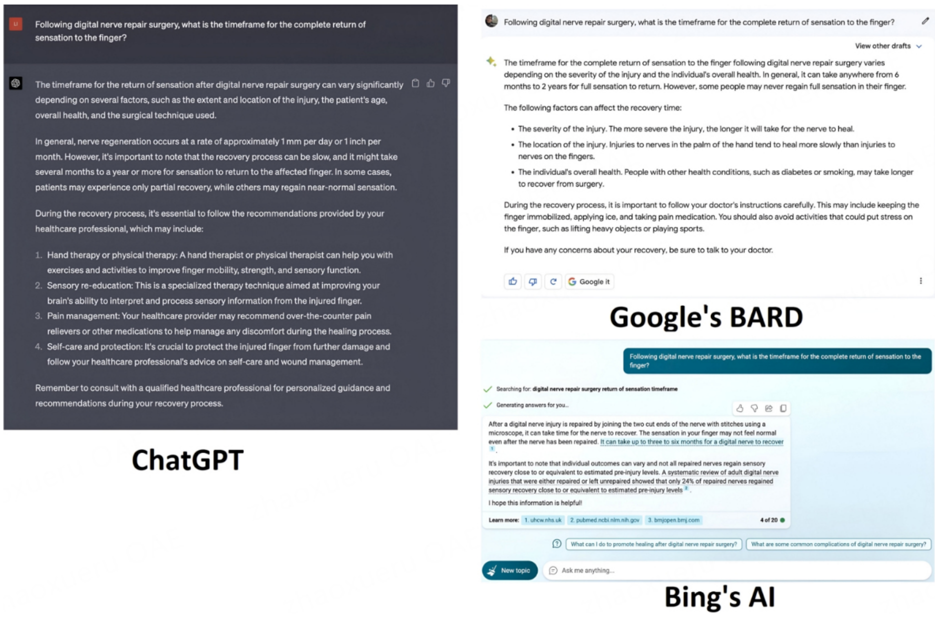

Figure 4. Following digital nerve repair surgery, what is the timeframe for the complete return of sensation to the finger?

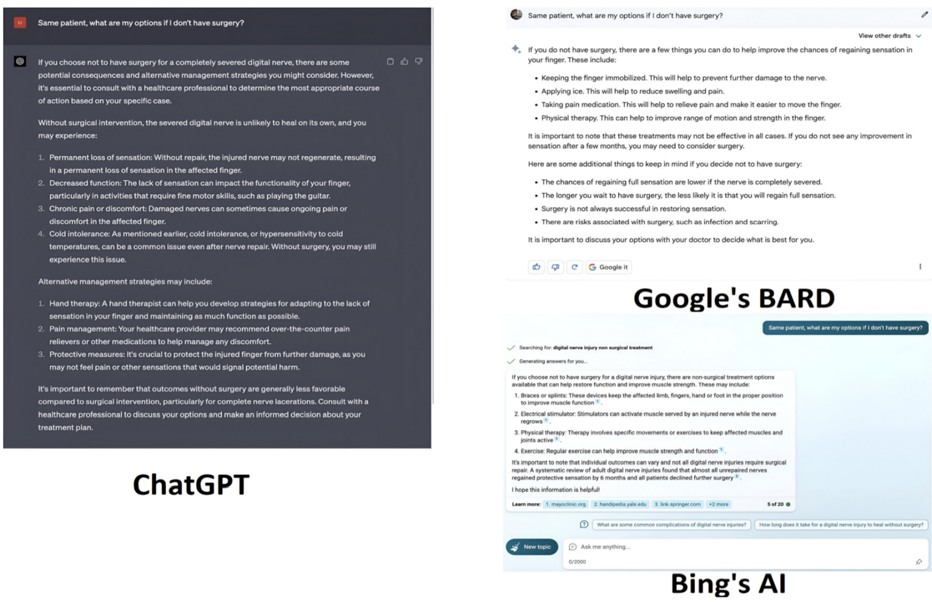

Figure 5. Same patient, what are my options if I don't have surgery?

management, resulting in a less comprehensive response.

In assessing readability, ChatGPT and BARD performed similarly [Table 2]. ChatGPT achieved the highest Flesch-Kincaid Grades and Coleman-Liau index scores, while BARD scored the highest Flesch Reading Ease Score. For reliability, ChatGPT marginally outperformed Bing AI, with BARD scoring the lowest. The t-test indicated that only six of the twelve comparisons were statistically significant. All the DISCERN score comparisons were statistically insignificant, meaning no LLM was necessarily more reliable than the other. Additionally, the readability for Bing AI *vs.* ChatGPT lacked statistical significance, suggesting that they exhibited similar comprehensibility.

DISCUSSION

Overall, ChatGPT provided the most comprehensive and comprehensible information, generating an extensive array of management solutions while incorporating contextual data, all in a reader-friendly manner. BARD's answers were similar in structure to ChatGPT's, albeit lacking the same level of detail and clarity. Bing AI struggled to attain similar levels of ChatGPT and BARD's comprehensibility and easiness of understanding, often utilizing technical language with frequent statistical references and lacking succinct summaries which came across as unempathetic and impersonal, as reflected in its Flesch Reading Ease Score. However, its consistent provision of sources compensates for its shortcomings, resulting in its comparable DISCERN score to ChatGPT's.

ChatGPT and BARD demonstrated superior comprehensibility and user-centricity to Bing's AI, making them more suitable for improving public comprehension of nerve injury management. Overall, the accuracy of the three LLMs was insufficient for use as automated diagnostic tools supporting healthcare professionals, as they neglected to reference key peripheral nerve injury guidelines or high-quality research[22]. Additionally, the LLMs omitted experimental treatments like stem cell therapy and photochemical tissue bonding, indicating a limitation in their algorithms' capacity for generating innovative solutions[23]. Despite this, as a clinical practice tool, they could ensure that patients are not misled or provided an unfeasible option.

In terms of information depth, ChatGPT provided supplementary data and enhanced its primary response when asked, outperforming BARD. Although BARD's comprehensiveness was comparable to ChatGPT, it failed to respond to the third prompt, severely impacting its readability and reliability scores. Moreover, ChatGPT presented the most thorough rationale for each suggestion, bolstering its credibility. BARD closely followed, but its explanations lacked ChatGPT's depth. Bing AI delivered the least detailed responses, which were sometimes not the gold standard. For example, it suggested exercise as an alternative treatment but neglected to specify that aerobic exercise is the most optimal form for addressing nerve injuries[24]. Nevertheless, Bing AI offered the most diverse range of treatment alternatives. Therefore, while Bing AI lacked depth, ChatGPT and BARD were limited in breadth. Notably, all three LLMs concentrated on layperson first aid, offering limited information for healthcare professionals and academics. They omitted management algorithms for nerve injuries, for example, when Bhandari (2019) recommends immediate surgery for penetrating trauma with neurological symptoms but conservative management for blunt trauma. The LLMs also neglect to address Seddon and Sunderland's categorizations of peripheral nerve injuries, which impact management[8,25,26]. This oversight may be attributed to the phrasing of the queries, as the LLMs could be presuming the authors are non-medical professionals, consequently yielding less scholarly and comprehensive replies. Considering this is the first study to compare these LLMs and on this topic, further research should seek to rectify these deficiencies.

Literature consistency was flagged by many studies investigating ChatGPT in the past, so the comparative performance in generating references was pertinent to this study. Bing AI demonstrated superior consistency compared to BARD and ChatGPT, often supplying relevant hyperlinks to fact-check its claims. Meanwhile, BARD failed to produce high-level references and ChatGPT only recommended databases for users to search or generated aberrant references. Despite this, Bing AI primarily cited health websites over scholarly articles and directed users to irrelevant web pages, resulting in a higher DISCERN score than BARD's but failed to surpass ChatGPT's. Despite these constraints, LLMs' rapid information retrieval and summarization capacity make them attractive for patients who are gathering information about emerging medical issues.

Regarding user-friendliness and comprehensibility, ChatGPT and BARD exhibited comparable success. ChatGPT's responses were succinct and devoid of technical jargon, rendering them advantageous for individuals with limited expertise, such as junior medical staff or patients. Additionally, ChatGPT acknowledged the user's occupation and hand dominance, personalizing the interaction. BARD displayed a similar syntactical approach and empathized with the user, offering greetings, and expressing sympathy for the user's injury. Although both models consistently advised consulting a doctor, BARD conveyed a warmer, more welcoming tone. Conversely, Bing AI manifested the least personable demeanor, utilizing cold, clinical language and frequent third-person pronouns. Although Table 2 shows Bing AI outperforming BARD in the Flesch-Kincaid Grade Level and Coleman-Liau Index, this is due to BARD's failure to answer prompt 3, adversely affecting its scores. Noting that half of the comparisons were statistically insignificant, the authors suggest further investigations to acquire more robust results.

Deploying AI chatbots in clinical settings engenders ethical concerns, encompassing potential patient confidentiality breaches and inaccuracies due to error-prone public health data. Such issues may protract diagnosis, compromise patient safety, and entail legal consequences. To establish reliable and accountable AI systems, prioritizing transparency, explainability, and adherence to regulations and privacy policies is imperative. Addressing data and algorithmic biases and ensuring ongoing monitoring, evaluation, and ethical guideline adherence is essential for developers and users.

This study's principal limitation was its reliance on a specialized cohort of certified Plastic Surgeons and trainees to assess LLMs' coherence, comprehensibility, and usability, which may impede generalizability and introduce subjectivity and biases. Nonetheless, this research constitutes a preliminary exploration, guiding future inquiries with varied clinician samples to appraise LLMs' utility in healthcare settings.

This study showed that ChatGPT consistently provided more reliable, evidence-based clinical advice than BARD and Bing AI. However, LLMs generally lack depth and specificity, limiting their use in individualized decision-making. Healthcare professionals are crucial in interpreting and contextualizing LLM responses, especially for complex cases requiring multidisciplinary input. Future research should enhance LLM performance by incorporating specialized databases and expert knowledge, ensuring traceability and credibility of AI-generated content, and integrating LLMs with human expertise to advance nerve injury management and support patient-centered care.

DECLARATIONS

Authors' contributions

Made substantial contributions to the conception and design of the study and performed data analysis and interpretation: Lim B, Seth I, Bulloch, Xie Y
Performed data acquisition, as well as provided administrative, technical, and material support: Lim B, Seth I
Supervision, validation: Hunter-Smith DJ, Rozen WM

Availability of data and materials

Not applicable.

Financial support and sponsorship
None.

Conflicts of interest
All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

REFERENCES

1. 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:online ahead of print. DOI PubMed
2. 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 PubMed
3. Menorca RM, Fussell TS, Elfar JC. Nerve physiology: mechanisms of injury and recovery. *Hand Clin* 2013;29:317-30. DOI
4. Ducic I, Fu R, Iorio ML. Innovative treatment of peripheral nerve injuries: combined reconstructive concepts. *Ann Plast Surg* 2012;68:180-7. DOI PubMed
5. Mhlanga D. Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. SSRN Journal 2023:online ahead of print.Available from: https://www.researchgate.net/profile/David-Mhlanga-2/publication/368476294_Open_AI_in_Education_the_Responsible_and_Ethical_Use_of_ChatGPT_Towards_Lifelong_Learning/links/63eb1c91bd7860764366f597/Open-AI-in-Education-the-Responsible-and-Ethical-Use-of-ChatGPT-Towards-Lifelong-Learning.pdf [Last accessed on 30 Aug 2023].
6. Newman RK, Mahdy H. Laceration. Treasure Island (FL): StatPearls Publishing; 2022. PubMed
7. Seddon HJ. A classification of nerve injuries. *Br Med J* 1942;2:237-9. DOI
8. Bhandari PS. Management of peripheral nerve injury. *J Clin Orthop Trauma* 2019;10:862-6. DOI PubMed PMC
9. Griffin MF, Malahias M, Hindocha S, Khan WS. Peripheral nerve injury: principles for repair and regeneration. *Open Orthop J* 2014;8:199-203. DOI
10. Campbell WW. Evaluation and management of peripheral nerve injury. *Clin Neurophysiol* 2008;119:1951-65. DOI PubMed
11. Dębski T, Złotorowicz M, Noszczyk BH. Long-term sensibility outcomes of secondary digital nerve reconstruction with sural nerve autografts: a retrospective study. *Eur J Trauma Emerg Surg* 2022;48:2341-7. DOI PubMed PMC
12. Dunlop RLE, Wormald JCR, Jain A. Outcome of surgical repair of adult digital nerve injury: a systematic review. *BMJ open* 2019;9:e025443. DOI PubMed PMC
13. Jain A, Dunlop R, Hems T, Tang JB. Outcomes of surgical repair of a single digital nerve in adults. *J Hand Surg Eur Vol* 2019;44:560-5. DOI PubMed
14. Höke A. A (heat) shock to the system promotes peripheral nerve regeneration. *J Clin Invest* 2011;121:4231-4. DOI PubMed PMC
15. Sulaiman W, Gordon T. Neurobiology of peripheral nerve injury, regeneration, and functional recovery: from bench top research to bedside application. *Ochsner J* 2013;13:100-8. PubMed PMC
16. Cheng AS, Hung L, Wong JM, Lau H, Chan J. A prospective study of early tactile stimulation after digital nerve repair. *Clin Orthop Relat Res* 2001:169-75. DOI
17. al-Ghazal SK, McKiernan M, Khan K, McCann J. Results of clinical assessment after primary digital nerve repair. *J Hand Surg Br* 1994;19:255-7. DOI PubMed
18. Rinker B, Liau JY. A prospective randomized study comparing woven polyglycolic acid and autogenous vein conduits for reconstruction of digital nerve gaps. *J Hand Surg Am* 2011;36:775-81. DOI PubMed
19. Cheung K, Hatchell A, Thoma A. Approach to traumatic hand injuries for primary care physicians. *Can Fam Physician* 2013;59:6148. PubMed PMC
20. Risitano G, Cavallaro G, Merrino T, Coppolino S, Ruggeri F. Clinical results and thoughts on sensory nerve repair by autologous vein graft in emergency hand reconstruction. *Chir Main* 2002;21:194-7. DOI PubMed
21. Wong JN, Olson JL, Morhart MJ, Chan KM. Electrical stimulation enhances sensory recovery: a randomized controlled trial. *Ann Neurol* 2015;77:996-1006. DOI PubMed
22. Modrak M, Talukder MAH, Gurgenashvili K, Noble M, Elfar JC. Peripheral nerve injury and myelination: potential therapeutic strategies. *J Neurosci Res* 2020;98:780-95. DOI PubMed PMC

23. Lopes B, Sousa P, Alvites R, et al. Peripheral nerve injury treatments and advances: one health perspective. *Int J Mol Sci* 2022;23:918. DOI PubMed PMC
24. Maugeri G, D'Agata V, Trovato B, et al. The role of exercise on peripheral nerve regeneration: from animal model to clinical application. *Heliyon* 2021;7:e08281. DOI PubMed PMC
25. Chhabra A, Ahlawat S, Belzberg A, Andreseik G. Peripheral nerve injury grading simplified on MR neurography: as referenced to Seddon and Sunderland classifications. *Indian J Radiol Imaging* 2014;24:217-24. DOI PubMed PMC
26. Kaya Y, Sarikcioglu L. Sir Herbert Seddon (1903-1977) and his classification scheme for peripheral nerve injury. *Childs Nerv Syst* 2015;31:177-80. DOI