Use of GenAI to Obtain Public Information on Plastic and Reconstructive Surgery Procedures: A Focus on Migraine Surgery

Edoardo Raposio[1,2] and Elisa Bertulla[1,2]
[1]Department of Surgical Sciences and Integrated Diagnostics (DISC), University of Genova, Italy
[2]Plastic and Reconstructive Surgery Division, IRCCS Ospedale Policlinico San Martino, Genova, Italy

edoardo.raposio@unige.it (corresponding author)
elisa.bertulla@gmail.com

Abstract: Background: In recent years, Generative Artificial Intelligence (GenAI) technologies have seen rapid development and widespread public release. These tools are now accessible to both professional researchers and the general public, offering new ways to obtain information across a wide range of disciplines, including medicine. Given their growing presence in clinical and academic environments, it is important to understand the reliability and scientific rigor of the information these platforms provide. The aim of this study was to evaluate the quality, accuracy, depth and readability of responses generated by nine widely available GenAI tools when asked to describe the indications, outcomes, potential complications of migraine surgery and the available alternatives. **Methods:** Nine most prominent and widely used GenAI platforms—ChatGPT, Gemini, Perplexity, Elicit, Scispace, Consensus, PaperPal, Julius, and Mistral AI—were prompted with the same standardized questions: *"Detail the outcomes and complications of migraine surgery."*, "*What are the indications for migraine surgery?" and "What are the alternatives to migraine surgery?*". The responses were then assessed in terms of scientific credibility, clarity, readability, depth of information, and the presence of references to peer-reviewed literature or established medical knowledge. **Results:** Overall, the responses received were highly satisfactory. All tools delivered prompt replies that were logical and scientifically credible. However, the level of detail, specificity, and accuracy varied across platforms. The most comprehensive and detailed answers were provided, in order, by Mistral, Julius, Scispace and Consensus, while the most readable ones were given by Mistral, ChatGPT and Elicit. Readability analysis indicated that content generally required a college-level education. **Conclusions:** Across all platforms, the core principles of migraine surgery were accurately outlined, including the indications, the expected outcomes, the incidence and severity of possible complications and the possible alternatives. In terms of results, the consistently high success rate of the procedure was clearly emphasized and conveyed an accurate overview of its clinical relevance.

Keywords: Migraine surgery, Artificial intelligence, Headache surgery, Generative AI tools

1. Introduction

Generative Artificial Intelligence (GenAI) is a branch of artificial intelligence that utilizes generative models to produce text, images, videos, and other types of data. These models are trained to recognize and learn patterns and structures from large datasets, allowing them to generate new content based on user input—usually in the form of natural language prompts. In simple terms, these tools process specific user questions (prompts) and respond using a combination of freely accessible online information (e.g., from Wikipedia or open-source texts).

In recent years, many GenAI tools have become publicly accessible, serving both research purposes and public use. These tools are increasingly employed by healthcare professionals for research and institutional tasks, as well as by non-experts seeking information on the diagnosis and treatment of medical conditions. The rise of GenAI technology has significantly improved the efficiency and accessibility of medical knowledge acquisition, processing, and dissemination.

The use of GenAI in surgical and healthcare settings covers the following areas:

1. Enhancing Medical Knowledge and Education

- GenAI can condense vast amounts of medical literature, guidelines, and research into concise, digestible summaries.
- It supports medical training by generating case studies, describing surgical techniques, and simulating patient interactions.
- Through virtual tutor functions, GenAI helps medical students and surgeons stay updated on the latest advancements in the field.

2. Supporting Clinical Decision-Making

- Virtual assistants provide personalized medical guidance to patients and healthcare providers by analyzing medical databases.• Surgeons can leverage AI for best-practice reviews, complication risk assessments, and postoperative checklist evaluations to aid in sound decision-making.

3. Patient Education and Engagement

- GenAI helps translate complex surgical information into understandable language for patients.
- Chatbots and virtual assistants can answer frequently asked questions about preoperative preparation, recovery, and follow-up care.
- Multimedia content generated by AI, such as videos and interactive diagrams, improves patient understanding, informed consent, and treatment compliance.

4. Research and Systematic Reviews

- GenAI facilitates faster literature searches, assists in conducting meta-analyses, and identifies relevant studies efficiently.
- Global studies are analyzed using generative artificial intelligence (GenAI), with a focus on the evolution over time of surgical outcomes, complication rates, and innovative techniques.

5. Surgical Planning and Simulation

- AI-powered tools support preoperative planning by converting imaging data (e.g., MRI and CT scans) into 3D models, helping design optimal surgical strategies.
- Surgeons can use virtual simulations to practice and refine techniques before performing actual procedures.

Since Guyuron first proposed surgical treatment for migraines in 2000, this approach has gained growing scientific validation and clinical support. Despite its development over the years, the procedure remains relatively unknown both in academic circles and among the public. Consequently, the use of GenAI tools to obtain more detailed and accessible information on the topic is becoming increasingly relevant—for both medical professionals and laypeople.

We aimed to assess what type of information could be retrieved from publicly accessible GenAI tools when asked about the expected outcomes and potential complications of migraine surgery.

2. Materials and Methods

It was consulted nine of the most well-known and widely used GenAI tools: ChatGPT (OpenAI, San Francisco, CA, USA), Gemini (Google DeepMind, London, UK), Perplexity (Perplexity AI Inc, San Francisco, CA, USA), Elicit

(Elicit Research PBT, Oakland, CA, USA), Scispace (Polydora Technologies Private Ltd, Bangalore, India), Consensus (American Fork, UT, USA), PaperPal (Cactus Communications Services Pte Ltd, Singapore), Julius (Triller, New York, NY, USA), and Mistral AI (Paris, France). The software tools were selected based on three criteria: relevance to the field of biomedical research (excluding those primarily designed for other purposes such as music, image or video generation, or programming), their popularity (at the end of July 2024) and free accessibility. Each tool was posed the same questions: "Detail the outcomes and complications of migraine surgery.", “What are the indications for migraine surgery?” and “What are the alternatives to migraine surgery?”. The research was carried out during the month of August 2025. The responses, along with the bibliographic references (when available) on which the answers were based, were collected and analyzed. The accuracy and specificity of the responses have been assessed by the senior author (ER - who has specific clinical and scientific expertise in the field for over 20 years) in accordance with the guidelines provided by the American Society of Plastic Surgeons (https://www.plasticsurgery.org/reconstructive-procedures/migraine-surgery), with a score from 0 to 30 where 0 is insufficient and 30 is high.

3. Results

All the software provided logically valid answers, with response times consistently short and comparable (under 4 seconds) except for Elicit and Spispace, which took about 3 minutes.

The full responses from each tool are available in Additional Material I.

Our initial evaluation and comparison of the responses were conducted based on the quantity and quality of the bibliographic references cited (Table I-II-III).

Table 1: Degree of detail of the answers

SOFTWARE	References	Degree of detail of the answers	Flesch-Kincaid Score
ChatGPT	3	30	11.80
Gemini	0	25	12.35
Perplexity	4	28	16.62
Elicit	4	27	15.55
SciSpace	5	28	15.14
Consensus	10	26	17.06
PaperPal	8	25	15.62
Julius	4	27	17.17
Mistral AI	0	30	11.84

Summary data of the analyses conducted on the first prompt “Detail the outcomes and complications of migraine surgery”, using the various softwares covered by the study.

Degree of detail of the answers from 0 to 30 where 0 is insufficient and 30 is high.

Table 2: Completeness and appropriateness of the answers

SOFTWARE	References	Completeness and appropriateness of the answers	Flesch-Kincaid Score
ChatGPT	2	24	14.35
Gemini	7	30	15.22
Perplexity	8	27	17.63
Elicit	9	26	11.70
SciSpace	15	26	23.37
Consensus	19	30	22.03
PaperPal	9	25	17.90
Julius	0	27	21.60
Mistral AI	0	30	13.65

Summary data of the analyses conducted on the second prompt: “What are the indications for migraine surgery?”

Degree of detail of the answers from 0 to 30 where 0 is insufficient and 30 is high.

Table 3: Completeness and appropriateness of the answers

SOFTWARE	References	Completeness and appropriateness of the answers	Flesch-Kincaid Score
ChatGPT	0	28	15.50
Gemini	8	28	18.28
Perplexity	7	26	17.43
Elicit	5	0	14.74
SciSpace	17	30	25.60

SOFTWARE	References	Completeness and appropriateness of the answers	Flesch-Kincaid Score
Consensus	6	28	19.37
PaperPal	7	26	22.64
Julius	0	30	27.95
Mistral AI	0	28	15.80

Summary data of the analyses conducted on the third prompt: "What are alternatives to migraine surgery?"

Degree of detail of the answers from 0 to 30 where 0 is insufficient and 30 is high.

Of all the tools, only Mistral AI failed to provide any references.

Based on the average, the AI tools are ranked by the number of references provided as follows: Spispace (12.3), Consensus (11.6), PaperAI (8), Perplexity (6.3), Elicit (6), Gemini (5), ChatGPT (1.3), and Julius (1.3).

The average scores for all systems are presented in Table IV.

Table 4: Completeness and appropriateness of the answers

SOFTWARE	References	Completeness and appropriateness of the answers	Flesch-Kincaid Score
ChatGPT	1.6	27.3	13.88
Gemini	5	27.6	15.26
Perplexity	6.3	27	17.22
Elicit	6	17.6	13.99
SciSpace	12.3	28	21.37
Consensus	11.6	28	19.48
PaperPal	8	25.3	18.72
Julius	1.3	28	22.23
Mistral AI	0	29.3	13.76

Mean scores for each AI software, based on the combined results of the three administered questions.

For the sake of completeness, it's worth mentioning that ChatGPT provided some inaccurate references.

Most of the references provided by Gemini and Perplexity were from websites rather than scientific publications.

In addition, Elicit provided an irrelevant answer (hallucination) to the third question.

The tools that delivered the most comprehensive and detailed responses were, in order: Mistral, Julius, Scispace and Consensus, while the most readable ones were given by Mistral, ChatGPT and Elicit.

All the software tools accurately presented the universally accepted principles of migraine surgery, covering the appropriate indications for the surgical treatment, the expected outcomes, the frequency and severity of potential complications, and the available alternatives to surgery. The high success rate of the procedure was consistently emphasized, while complications were uniformly described as minor, infrequent, and typically temporary.

Some platforms went further by including additional, potentially overly technical, details for a general audience. For example, Consensus discussed binary outcomes and distinctions between endoscopic and non-endoscopic techniques. Scispace provided success rates for different trigger sites. Perplexity organized outcomes by migraine intensity, frequency, and resolution rates. ChatGPT offered insights into various aspects, such as postoperative quality of life, types of complications, trigger site differences, and criteria for patient selection.

Specific percentages related to success rates or complication rates were rarely mentioned. Nevertheless, even in these cases, the answers remained sensible, accurate, and appropriate in tone.

To evaluate the clarity and comprehensibility of the responses, we used the reference software Text Inspector (WebLingua Ltd, UK). Developed by Professor Stephen Bax, an expert in Applied Linguistics, Text Inspector received the British Council ELTons Digital Innovation Award in 2017 and is academically supported by the Centre for Research in English Language Learning and Assessment (CRELLA) at the University of Bedfordshire. The tool assesses the educational level required to fully understand a given text, generating a score aligned with various educational benchmarks (Flesch-Kincaid Score, see Table V). The results showed consistently high scores across all responses, ranging from 11.70 to 27.95 (see Table I-II-III). While these high scores reflect the credibility and sophistication of the sources, they may also represent a barrier to understanding for individuals with lower levels of education.

Table 5: Flesch-Kincaid Score

Flesch-Kincaid Score	Corresponding US Educational Level	Approximate UK Equivalent	Typical Usage
1.0 – 3.0	Early Elementary School	Primary School (KS1-2)	Children's books, beginner reading materials
4.0 – 6.0	Upper Elementary/Middle School	Primary School (KS2) / Lower Secondary (KS3)	Standard classroom texts, general fiction
7.0 – 9.0	Middle School/Early High School	Lower Secondary (KS3)	Most adult fiction, general non-fiction, newspapers
10.0 – 12.0	High School	Upper Secondary (KS4-5)	Academic texts, some technical writing
13.0+	College/University	Higher Education(University or degree-level studies)	Advanced academic papers, complex technical manuals

Flesch-Kincaid Score

4. Discussion

To the best of our knowledge, only one other recent study—conducted by Chartier et al. —has explored the application of Generative AI in the context of migraine surgery. However, that research specifically examined the use of patient drawings as predictors of surgical outcomes. To date, no studies have investigated the information provided by various GenAI tools regarding headache surgery, particularly in relation to procedural success rates and potential complications.

Overall, the results of our analysis were highly satisfactory. All the tools evaluated delivered responses that were not only impressively fast but also generally sensible and scientifically credible, though with varying levels of accuracy.

The language used across all responses was consistently appropriate and correct. While the content was always comprehensible, it generally required a higher level of education—typically at the college or university level—for full understanding (see Table V).

The performance of a generative Artificial Intelligence system largely depends on the modality and nature of its training dataset. Most of the tools evaluated in this study generated responses based on information drawn from accredited scientific literature.

This is a particularly significant finding, as it reinforces the reliability and scientific validity of the information provided. In contrast, sourcing data from general websites—as opposed to peer-reviewed scientific publications—may increase the risk of inaccuracies or misinformation.

In our analysis three GenAI tools appeared to incorporate information from websites: ChatGPT, Perplexity and Gemini, and not exclusively from established and reputable scientific institutions.

It is important to highlight that—even in the absence of citations—the responses from the tools that did not reference scientific literature were still logical, coherent, and reasonable.

This aspect is particularly significant given the current absence of regulations or legal standards governing the accuracy of content generated by GenAI tools. In this context, the internal consistency and plausibility of the responses take on added importance.

It is crucial to clarify that our intention was not to establish a "ranking" of the tools evaluated. The applications of Generative AI are vast, spanning text generation, video creation, sound design, and image production. Each of the tools used in this study offers unique strengths and specialized use cases. Our objective was solely to assess the overall reliability of the responses provided, particularly in relation to a niche topic within the field of plastic surgery—migraine surgery—which continues to evolve.

This study focused on assessing the plausibility and factual accuracy of responses to a specific question, especially in the context of non-professional users seeking medical information. In our view, future research could explore the utility and precision of GenAI tools for medical professionals with a specific interest in migraine surgery—for example, as a resource in preparing reviews or conducting systematic research on the topic.

The use of GenAI in healthcare is growing rapidly, but it raises ethical concerns of equal importance, including:

- Protection of Data and Security: GenAI systems often rely on large datasets, some of which may contain sensitive patient information. Strict adherence to data privacy regulations such as HIPAA and GDPR is essential to protect patient confidentiality. Unregulated access to personal health data highlights the urgent need for robust AI security measures in most medical applications.
- Biased Algorithms and Lack of Transparency: Many AI systems function as "black boxes," where the decision-making process is not clear to users. In medicine, understanding how AI reaches certain conclusions is crucial. Poorly trained AI algorithms can introduce bias, potentially leading to unequal treatment of patients and adversely impacting marginalized groups. Researchers and developers must take active steps to identify and mitigate such biases.
- Likelihood of Misinformation: GenAI can generate authoritative-sounding but inaccurate or misleading content, which may negatively influence medical decisions. Ensuring that AI-generated medical information follows established guidelines and is evidence-based is vital.
- Lack of Benchmarking Against Expert Responses: This study did not compare GenAI responses with those from migraine surgery experts. A comparison with answers from board-certified neurologists or headache surgeons would provide a valuable reference to evaluate how closely AI-generated responses align with expert consensus.

The use of GenAI in medical and surgical fields brings to mind the classic computer science adage: "garbage in, garbage out." Most surgical and medical journals are behind paywalls, limiting access to high-quality medical information. This challenge is especially notable in plastic surgery, where physicians must adhere to strict ethical standards to protect patient confidentiality, particularly regarding clinical photographs. While GenAI can retrieve article abstracts from databases such as PubMed, the information it accesses tends to be limited to freely available content. Open-access materials often present information in a predominantly positive light and lack the critical evaluation provided by expert reviewers. Consequently, GenAI is unable to deliver the detailed, nuanced, and critical analysis necessary for informed clinical decision-making. Furthermore, dissenting or minority viewpoints are often overlooked in AI-generated content.

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Conflict of Interests: The authors have no relevant financial or non-financial interests to disclose. None of the authors has a financial interest in any of the products, devices, or drugs mentioned in this manuscript.

Author Contributions: ER developed the concept, checked the results and wrote the manuscript. EB checked the results and wrote the manuscript.

Ethics approval: Not required.

AI declaration: AI tools were not used.

Availability of data and materials: The data analyzed in this study are available from the corresponding author upon request.

References

Alrahbeni, T., Mahal, A., Alkhouri, A. (2024) 'Surgical interventions for intractable migraine: a systematic review and meta-analysis', *Int J Surg*, 110, pp. 6306–13.

Baldelli I, Mangialardi ML, Salgarello M, Raposio E. Nummular headache and its surgical treatment. Plast Reconstr Surg Glob Open. 2020;8(7):e2989.

Baldelli I, Mangialardi ML, Salgarello M, Raposio E. Peripheral occipital nerve decompression surgery in migraine headache. Plast Reconstr Surg Glob Open. 2020;8:e3019.

Baldelli I, Mangialardi ML, Raposio E. Site V surgery for temporal migraine headaches. Plast Reconstr Surg Glob Open. 2020;8(6):e2886.

Caruana G, Bertozzi N, Boschi E, Pio Grieco M, Grignaffini E, Raposio E. Endoscopic forehead surgery for migraine therapy Personal technique. Ann Ital Chir. 2014;85:583-6.

Caruana G, Grignaffini E, Raposio E. Endoscopic forehead muscle resection for nerve decompression: a modified procedure. Plast Reconstr Surg Glob Open. 2015 Apr 7;3(3):e342.

Chang I.A. et al. (2023) 'Nonpharmacologic Treatments for Chronic and Episodic Migraine: A Systematic Review and Meta-Analysis', *Plast Reconstr Surg*, 152, pp. 1087–98.

Chartier C. et al. (2023) 'Artificial Intelligence–Enabled Evaluation of Pain Sketches to Predict Outcomes in Headache Surgery', *Plast Reconstr Surg*, 151(2), pp. 405–11.

Egan K.G. et al. (2016) 'Evaluation of Migraine Surgery Outcomes through Social Media', *Plast Reconstr Surg - Global Open*, 4(10).

Elhawary H. et al. (2021) 'Efficacy and Safety of Migraine Surgery', *Ann Surg*, 275:e315-e323.

Evans A.G. et al. (2023) 'Outcomes of Surgical Treatment of Migraines: A Systematic Review & Meta-Analysis', in *Plast Surg (Oakv)*

Faizo E. et al. (2024) 'The Efficacy of Trigger Site Surgery in the Elimination of Chronic Migraine Headache: An Update in the Rate of Success and Failure', *Cureus*, 20;16(2):e54504.

Gfrerer L. et al. (2014) 'Nonendoscopic Deactivation of Nerve Triggers in Migraine Headache Patients: Surgical Technique and Outcomes', *Plast Reconstr Surg*, 134, pp. 771–8.

Gfrerer L. et al. (2018) 'Migraine Surgery: An All or Nothing Phenomenon? Prospective Evaluation of Surgical Outcomes', *Ann Surg*, 269, pp. 994–9.

Gfrerer L. et al. (2019) 'Migraine Surgery: An All or Nothing Phenomenon? Prospective Evaluation of Surgical Outcomes', *Ann Surg*, 269(5), pp. 994–9.

Guyuron B. et al. (2000) 'Corrugator supercilii muscle resection and migraine headaches', *Plast Reconstr Surg*, 106(2), pp. 429–34.

Guyuron B. et al. (2011) 'Five-Year Outcome of Surgical Treatment of Migraine Headaches', *Plast Reconstr Surg*, 127(6038).

Hatef D.A. et al. (2020) 'A Comprehensive Review of Surgical Treatment of Migraine Surgery Safety and Efficacy', *Plast Reconstr Surg*, 146(187).

https://utswmed.org/conditions-treatments/migraine/migraine-surgery/migraine-qualityoutcomes/.

https://www.plasticsurgery.org/Documents/Health-Policy/Positions/ASPS-Statement_Migraine-Headache-Surgery.pdf

Huayllani, M.T., Janis, J.E. (2023) 'Migraine Surgery and Determination of Success over Time by Trigger Site: A Systematic Review of the Literature', *Plast Reconstr Surg*, 151(1).

Israel JS et al. (no date) 'Outcomes in Migraine Surgery. 2020:183-193', in A. Afifi, Z. Peled, and J. Janis (eds) *Surgical Treatment of Chronic Headaches and Migraines*. Cham: Springer.

Janis, J.E., Hehr, J., Huayllani, M.T. (2023) 'Functional outcomes between headache surgery and targeted botox injections: A prospective multicenter pilot study', *JPRAS Open*, 38, pp. 152–62.

Jose, A., Nagori, S.A., Roychoudhury, A. (2017) 'Surgical Management of Migraine Headache', *J Craniofac Surg*, 29, p. 106108.

Kurlander D.E. et al. (2016) 'In-Depth Review of Symptoms, Triggers, and Surgical Deactivation of Frontal Migraine Headaches (Site I', *Plast Reconstr Surg*, 138, pp. 681–88.

Mangialardi ML, Baldelli I, Salgarello M, Raposio E. Decompression surgery for frontal migraine treatment. Plast Reconstr Surg Glob Open. 2020;8(10):e3084.

Nagori, S., Jose, A., Roychoudhury, A. (2019) 'Surgical Management of Migraine Headaches: A Systematic Review and Meta-analysis', *Ann Plast Surg*, 83(2), pp. 232–40.

Omranifard M. et al. (2016) 'A comparison of outcome of medical and surgical treatment of migraine headache: In 1 year follow-up', *Adv Biomed Res*, 5(1).

Peled Z.M. et al. (2023) 'Anatomic Anomalies of the Nerves Treated during Headache Surgery', *Plast Reconstr Surg Glob Open* [Preprint].

Raposio E, Caruana G. Tips for the surgical treatment of occipital nerve-triggered headaches. Eur J Plast Surg. 2017.40(3):177-82.

Raposio E, Simonacci F. Frontal trigger site deactivation for migraine surgical therapy. Plast Reconstr Surg Glob Open. 2020;8(4):e2813.

Raposio G, Antonini A, Gualdi A, Raposio E. Frontal site surgery for chronic migraine therapy. Acta Biomed, 2023; 94(6): e2023253.

Raposio, E., Bertozzi, N. (2019) 'Trigger site inactivation for the surgical therapy of occipital migraine and tension-type headache: Our experience and review of the literature', *Plast Reconstr Surg Glob Open*, 7:e2507.

Raposio, E., Caruana, G. (2015) 'Frontal endoscopic myotomies for chronic headache', *J Craniofac Surg*, 26;3:e2013.

Raposio, E., Raposio, G. (2023) 'Surgical therapy of migraine: a 12-year single-center experience', *Eur J Plast Surg*, 46(5), pp. 699–705.

Raposio, G., Raposio, E. (2022a) 'Principles and techniques of migraine surgery', *Eur Rev Med Pharmacol Sci*, 26, pp. 6110–3.

Raposio, G., Raposio, E. (2022b) 'Temporal surgery for chronic migraine treatment: A minimally-invasive perspective', *Ann Med Surg*, 76, p. 103578.

Shikowitz-Behr, L., Austen, W.J. (2020) 'Postoperative Management After Headache and Migraine Surgery', in A. Afifi, Z. Peled, and J. Janis (eds) *Surgical Treatment of Chronic Headaches and Migraines*. Cham: Springer. Available at: https://doi.org/10.1007/978-3-030-36794-7_11.

Simonacci F Lago G, Bertozzi N, Raposio E. Surgical deactivation of occipital migraine trigger site. Chirurgia. 2020;33(3):143-.9.

Urhan N. et al. (2023) 'Long-term results of migraine surgery and the relationship between anatomical variations and pain', *JPRAS*, 82, pp. 284–90.

Yang Y. et al. (2015) 'Occipital Nerve Stimulation for Migraine: A Systematic Review', *Pain Practice : The Official Journal of World Institute of Pain*, 16(4), p. 509517.