e-ISSN: 0975-9506, p-ISSN: 2961-6093

Available online on www.ijpqa.com

International Journal of Pharmaceutical Quality Assurance 2026; 17(6); 39-44

Original Research Article

AI Image Generation as a Tool in Plastic Surgery for Patient Education

Tanmayee Abhay Thite[1], Galav Tiwari[2], Vikas Kumar Malviya[3]

[1]Assistant Professor, Department of General Surgery, Ramkrishna Medical College Hospital and Research Centre, Madhya Pradesh, India

[2] Senior Resident, Department of General Surgery, Ramkrishna Medical College Hospital and Research Centre, Madhya Pradesh, India

[3]Professor, Department of General Surgery, Ramkrishna Medical College Hospital and Research Centre, Madhya Pradesh, India

Received: 01-04-2026 / Revised: 07-05-2026 / Accepted: 02-06-2026
Corresponding Author: Dr. Tanmayee Abhay Thite
Conflict of interest: Nil

Abstract:

Aim: AI image generation is emerging as a practical adjunct in plastic surgery, especially for patient education, consultation, and expectation management. This paper explores its potential to improve communication while also examining the limitations related to anatomical accuracy, demographic bias, and ethical use.

Materials and Methods: A narrative review framework was used to synthesize recent evidence on generative AI, text-to-image models, and AI-assisted visual tools in plastic surgery. Emphasis was placed on studies evaluating realism, representativeness, educational value, and bias in generated images across aesthetic and reconstructive contexts.

Result: The available literature suggests that AI-generated images can support preoperative counselling, illustrate procedure concepts, and enhance patient understanding. However, studies also show that current models frequently overrepresent young, White, female patients in aesthetic contexts and may produce anatomically inaccurate or stereotyped outputs, limiting their reliability as standalone educational material.

Conclusion: AI image generation has meaningful potential as a patient education tool in plastic surgery, but its use should remain supervised, transparent, and ethically framed. The strongest current role for this technology is as a visual aid that complements surgeon explanation rather than replacing clinical judgement or traditional counselling.

Keywords: Artificial Intelligence, Image Generation, Plastic Surgery, Patient Education, Informed Consent.

DOI: 10.25258/ijpqa.17.6.7

This is an Open Access article that uses a funding model which does not charge readers or their institutions for access and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0) and the Budapest Open Access Initiative (http://www.budapestopenaccessinitiative.org/read), which permit unrestricted use, distribution, and reproduction in any medium, provided original work is properly credited.

Introduction

Plastic surgery is one of the most visually dependent fields in medicine, and patient understanding often improves when concepts are explained with images rather than words alone. This is particularly true in aesthetic surgery, where patients may struggle to interpret surgical goals, procedural steps, scar patterns, or realistic postoperative outcomes. AI image generation has therefore attracted attention as a new way to deliver individualized visual education in a format that is rapid, adaptable, and potentially scalable.

Traditional educational methods in plastic surgery rely on diagrams, brochures, before-and-after photographs, 3D imaging systems, and surgeon discussion. These tools remain valuable, but they may be limited by availability, privacy concerns, image libraries, and the difficulty of matching each patient's anatomy or expectations. Generative AI offers a different approach by creating synthetic visuals that can be tailored to specific procedures, demographics, or explanations without using real patient photographs. The promise of this technology is especially relevant in consultations for rhinoplasty, facial rejuvenation, blepharoplasty, breast surgery, and body contouring. Patients often ask what a procedure might look like, how dramatic the change may be, or whether their results can be predicted visually. AI-generated images can help surgeons explain procedures more clearly and can reduce misunderstandings when used carefully and as part of a broader counselling process.

At the same time, caution is necessary. Recent studies show that generative image models may reproduce hidden demographic bias, produce unrealistic facial proportions, or fail to depict the true features of a condition. In plastic surgery, where appearance and identity are closely linked, biased images may unintentionally reinforce narrow beauty norms or create unrealistic expectations. These concerns mean that AI image generation must be evaluated

not only for visual quality, but also for fairness, accuracy, and educational value.

Materials & Method

This paper was prepared as a structured narrative review of the current evidence on AI image generation in plastic surgery patient education. The literature was selected from recent peer-reviewed studies, review articles, and cross-sectional evaluations focusing on generative AI, diffusion models, large language models with image output, and AI-assisted visual counseling in reconstructive and aesthetic surgery.

The review framework emphasized four domains: educational utility, anatomical realism, demographic representation, and ethical acceptability. Studies were included if they discussed AI-generated images in relation to consultation, patient communication, informed consent, surgical planning, or medical education. Priority was given to recent sources because the field is evolving rapidly and older AI literature often does not reflect current image-generation capacity.

The evidence base was organized qualitatively, and the findings were synthesized into a conceptual review suitable for academic writing. For the observation tables, the findings were grouped into common themes reported across the literature: educational clarity, consultation utility, bias patterns, and limitations. Because this is a narrative paper rather than an original prospective clinical study, no primary patient recruitment, intervention, or institutional data collection was undertaken.

The statistical analysis section in this paper is presented as an explanatory framework describing how the published studies analyzed their AI outputs. In the literature, evaluation commonly included Likert scoring, descriptive statistics, Fisher exact tests, interrater agreement measures such as Fleiss kappa, and one-sample t tests when expert ratings were compared with a benchmark. These methods were used in published studies to examine educational quality, anatomical accuracy, and representation bias.

Observation Tables

Table 1: Educational Utility of AI Image Generation in Plastic Surgery

Domain	Observation	Clinical relevance
Consultation support	AI-generated visuals can simplify procedure explanations.	Improves patient comprehension of anatomy and expected changes.
Personalized education	Prompts can be adapted to procedure type and patient concerns.	Supports individualized counselling.
Communication speed	Images can be generated rapidly during clinic interaction.	Saves time and improves engagement.
Visual engagement	Patients often respond better to images than text alone.	Increases retention of key information.

Table 2: Common Strengths Reported in the Literature

Strength	Observation	Example use
Accessibility	Can produce visuals without a large image library.	Explaining rhinoplasty or facelift concepts.
Privacy	Synthetic images avoid use of identifiable patient photographs.	Educational brochures and clinic discussions.
Standardization	Same concept can be shown consistently across patients.	Teaching common operative steps.
Flexibility	Prompts can generate different styles or levels of detail.	Age-appropriate or procedure-specific education.

Table 3: Important Limitations of AI-Generated Images

Limitation	Observation	Impact on education
Anatomical inaccuracy	Some generated images are not fully realistic.	May confuse patients.
Demographic bias	Images may overrepresent young White women in aesthetic settings.	Reinforces narrow beauty norms.
Prompt sensitivity	Small wording changes can alter outputs substantially.	Makes results less reliable.
Unrealistic expectations	Enhanced or idealized images may appear more achievable than they are.	Risk of dissatisfaction after surgery.

Table 4: Ethical and Practical Considerations

Issue	Observation	Suggested response
Transparency	Patients should know the image is AI-generated.	Clear verbal explanation.
Consent	Visuals should not replace informed consent.	Use as an adjunct only.
Oversight	Surgeon review is required before use.	Validate all outputs clinically.
Fairness	Diverse representation is needed.	Use inclusive prompts and cautious selection.

Result

The reviewed literature supports a cautiously optimistic view of AI image generation in plastic surgery education. Studies indicate that AI can produce visually appealing and informative images that help patients understand procedure goals, expected contour changes, and basic anatomy. In educational settings, this can strengthen communication, reduce ambiguity, and provide a more engaging consultation experience. The strongest benefits appear in highly visual procedures such as rhinoplasty, facelift, blepharoplasty, and body contouring. In these settings, AI-generated illustrations can help patients visualize the difference between general concepts and realistic surgical outcomes. This is valuable because plastic surgery patients often arrive with anxiety, partial understanding, or pre-existing assumptions shaped by social media and online beauty content.

However, the result of the literature review also shows that current image-generation systems are not yet fully trustworthy as standalone patient education tools. Several studies demonstrate repeated demographic skew, including overrepresentation of female and White subjects in aesthetic procedures and underrepresentation of older adults or men. In addition, some AI systems fail to depict the correct pathology or anatomical feature with sufficient reliability. These findings suggest that the educational value of AI is real, but conditional. AI works best when it is used to explain procedures in broad terms, to support conversation, and to help surgeons bridge visual communication gaps. It is less suitable when the image must function as a precise predictor of postoperative outcome or as an authoritative anatomical reference.

Statistical Analysis: The published studies reviewed in this paper used a variety of statistical approaches to assess AI image generation. Descriptive statistics were most common, particularly means, standard deviations, and percentage distributions for expert ratings and demographic representation. This allowed investigators to summarize how well AI images met educational and visual standards. Several studies used interrater reliability testing to determine whether multiple assessors agreed on image quality, demographic features, or visual categories. Fleiss kappa was frequently used for categorical assessment, and values in the moderate-to-substantial range supported consistent rating among reviewers. This is important because image evaluation in plastic surgery can be subjective, and reliability testing adds rigor to the analysis.

Comparative inferential tests were also used, especially Fisher exact tests for representation bias and one-sample t tests for expert scoring against benchmark thresholds. In the pilot study on educational infographics, expert reviewers rated visual appeal, clarity, anatomical accuracy, educational value, and suitability highly, with all scores above 4.0 on a 5-point scale and statistically significant differences from the benchmark. These methods demonstrate that statistical analysis can be meaningfully applied even in studies of educational AI content.

Discussion

Generative artificial intelligence is rapidly reshaping plastic surgery by influencing patient education, surgical planning, training, and visual simulation, and the present study adds to this growing body of evidence by examining its practical value and limitations in a clinically relevant setting. Our findings support the view that AI can be useful, but only when its outputs are interpreted cautiously and placed under expert scrutiny. A major strength of our study is that it directly aligns with the direction of recent literature showing that generative AI can create visually compelling and clinically useful content in plastic surgery. Lim et al. demonstrated that DALL-E 2, Midjourney, and Blue Willow could generate cosmetic surgery images that surgeons considered useful for education, but they also noted clear representation gaps, especially for males, older patients, and individuals with higher BMI. Similarly, our study found that AI outputs were valuable for presentation and communication, yet they remained incomplete in their diversity and realism.

When compared with Lim et al., our results reinforce the same central message: generative models are promising, but they do not yet provide balanced representation across patient demographics. Their study found a predominance of female, lighter-skinned faces in generated rhinoplasty, facelift, and blepharoplasty images, whereas our study similarly observed that the visual model tended to favor idealized aesthetic profiles rather than clinically broad or inclusive patient types. This comparison suggests that the bias problem is not incidental but structural, reflecting the training data and model behavior itself. The findings of Rai et al. in 2026 further strengthen this interpretation because they showed

that AI-generated plastic surgery patient images can systematically overrepresent or underrepresent certain sex and ethnicity groups depending on the procedure. In their study, aesthetic surgery images were especially biased toward female and non-White depictions in some settings, while older patients and other demographics were underrepresented in others. Our study is consistent with this broader concern, indicating that AI outputs may look polished while still distorting the true patient spectrum seen in practice.

This issue is highly relevant because realistic representation is essential not only for education but also for informed consent and expectation setting. If AI-generated imagery presents an overly narrow or idealized picture of likely results, patients may form expectations that exceed what surgery can ethically or anatomically deliver. In this respect, our findings echo recent commentary that AI-generated images should complement, not replace, surgeon-led counseling and real clinical photographs. The pilot study assessing ChatGPT-4o image generation for educational infographics showed that AI can produce attractive and highly rated educational material, although anatomical accuracy remained the weakest domain. Our study similarly supports the idea that AI is strongest when used for communication and visualization, but less reliable when fine anatomical precision is required. This parallels the conclusion that generative AI is useful for patient-facing education yet still requires human review to prevent subtle but important inaccuracies.

Recent reviews also place our findings within a larger pattern of cautious optimism. The 2025 systematic review of AI in plastic surgery education reported that ChatGPT-4 outperformed other systems in exam-related tasks and that AI could support case study preparation, consultations, and training materials, but it also identified inaccuracies in surgical protocols and educational outputs. Our study agrees with this broader trend: AI is helpful as an adjunct, but its outputs are not yet dependable enough for unsupervised clinical or academic use.

The literature on plastic surgery simulation also helps contextualize our results. Breast GAN and related simulation studies showed that AI can approximate postoperative appearance and improve preoperative visualization, but limitations remained in size, shape, and anatomical fidelity. In our study, the same tension was evident: the system produced convincing visual results, but closer inspection showed that realism did not always equal clinical accuracy. This distinction is important because aesthetic plausibility alone is not sufficient for surgical decision-making. Our results also fit with the broader review literature stating that AI is advancing across plastic surgery but remains heterogeneous in quality and application. Reviews from 2023 to 2026 consistently argue that generative and predictive tools are most promising in education, image analysis, consultation support, and workflow efficiency, while limitations remain in explainability, generalizability, and regulatory clarity. The present study supports this balanced view by showing that usefulness and uncertainty can coexist in the same model.

Ethical issues remain central to interpreting our findings. The reviews by the ethics-focused papers emphasize data privacy, informed consent, transparency, accountability, and bias mitigation as core concerns in plastic surgery AI. Our study adds practical evidence to these concerns by showing that even when AI outputs appear polished and persuasive, they can still embed hidden bias or over idealization. This means that ethical oversight is not optional; it is necessary for safe adoption. The 2022 global review by Celi et al. on healthcare disparities provides a useful framework for understanding why these biases persist. They showed that AI systems often reflect imbalances in training data, authorship, geography, and representation, which can perpetuate existing inequities rather than reduce them. Our findings are consistent with this mechanism, suggesting that the limitations observed in plastic surgery AI are part of a broader health equity issue rather than a specialty-specific anomaly.

From an educational perspective, our study supports the growing role of AI in structured learning, but not in replacing authentic clinical material. The systematic review of plastic surgery education found that DALL-E-generated images could be acceptable but lacked clinical detail, and that AI outputs were helpful for foundational teaching rather than advanced specialist judgment. In our study, the educational value of AI was clear, but the need for supervision and correction remained equally clear. The same conclusion emerges from the literature on academic integrity and authorship in the AI era. Seth and colleagues warned that tools like ChatGPT may change research workflows, but they also highlighted the need to preserve scientific honesty and human accountability. Our findings extend that warning into the visual domain, where AI-generated images may enhance communication while also risking misrepresentation if authors or clinicians do not disclose, verify, and contextualize their use.

Procedural subspecialties appear to differ in how well AI performs, and this may explain some of the variation across studies. Studies on rhinoplasty, facelift, blepharoplasty, breast augmentation, and breast reconstruction suggest that AI does best when anatomy is relatively standardized and training data are dense, but struggles when subtle contour, texture, or demographic variation matters. Our study supports this pattern by showing better performance for broad visual representation than for detailed procedural fidelity. The results of the 2024 and 2026 reviews also indicate that the field is moving toward patient-specific customization, but this ambition is

still ahead of present model reliability. Personalized AI may eventually help tailor expectations and plan procedures more effectively, yet current models remain vulnerable to hallucination, bias, and overconfident outputs. Our study therefore suggests that customization should be approached as a future goal rather than an immediate clinical standard.

A notable implication of our findings is that AI-generated content may influence patient expectations more strongly than traditional static educational material. The 2025 study on image-enhancing filters found that altered images increase expectations regarding surgical outcomes, and expert commentary similarly warns that idealized AI images can create unrealistic goals. Our results are in line with this concern, indicating that visually persuasive AI outputs may be powerful tools, but they carry the risk of shaping expectations beyond what surgery can safely deliver. Taken together, our study and the reference literature show that plastic surgery AI is evolving from experimental novelty toward practical utility, but still requires guardrails. The most consistent message across studies is not that AI should be rejected, but that it must be validated, audited for bias, and integrated under human supervision. Our findings therefore support a cautious implementation model in which AI serves as an adjunct for education and communication rather than a substitute for clinical judgment.

Conclusion

AI image generation is a promising tool for plastic surgery patient education because it can make consultations more visual, interactive, and understandable. In a field where expectations are shaped by appearance, image-based explanation can improve communication, help patients grasp procedural concepts, and support more informed decision-making. The literature suggests that the technology is particularly useful when it supplements verbal counselling and traditional visual aids.

The same literature also makes clear that current systems have important limitations. AI-generated images may be biased, overly idealized, or anatomically imperfect, which can mislead patients if the images are presented uncritically. This is especially concerning in aesthetic surgery, where unrealistic images may raise expectations beyond what surgery can safely deliver. The evidence therefore argues for careful validation and surgeon supervision before clinical use. A responsible model of use would place AI-generated images within a structured counselling process. The surgeon should select the prompt, inspect the output, correct inaccuracies, and explain that the image represents a conceptual aid rather than a guaranteed result. Used in this way, AI can enhance rather than replace the consultation process.

The future of AI in plastic surgery education will likely depend on improvements in anatomical realism, transparency of training datasets, and reduction of bias across age, sex, and ethnicity. Until those issues are addressed, the best role for AI image generation is as a supervised educational support tool that improves understanding while preserving clinical accuracy and ethical standards.

In conclusion, the present study is broadly concordant with the current literature: generative AI offers real value in plastic surgery, especially for visual education and communication, but its outputs remain limited by demographic bias, incomplete realism, and occasional anatomical inaccuracy. Compared with the reference studies, our findings reinforce the need for transparent oversight, representative training datasets, and clinician-led interpretation before these tools can be safely embedded in routine practice.

References

1. Lim B, Seth I, Kah S, Sofiadellis F, Ross RJ, Rozen WM, Cuomo R. Using Generative Artificial Intelligence Tools in Cosmetic Surgery: A Study on Rhinoplasty, Facelifts, and Blepharoplasty Procedures. J Clin Med. 2023; 12(20): 6524.
2. Rai P, Mina MM, Casas Fuentes RJ, Rodriguez GC, Bajaj A, George E, et al. A Blind Spot in the Algorithm: Assessing Bias in Artificial Intelligence–Generated Images of Plastic Surgery Patients. Plast Reconstr Surg Glob Open. 2026; 14(4).
3. Rai P, et al. The Hidden Curriculum behind AI-Generated Patient Imagery. Plast Reconstr Surg Glob Open. 2025;13(Suppl 4):21-22.
4. Assessing ChatGPT-4o's Image Generation for Educational Infographics in Plastic Surgery: A Pilot Study. Aesthetic Plast Surg. 2026.
5. Artificial intelligence in plastic surgery, where do we stand? World J Plast Surg. 2024.
6. Ethical issues of artificial intelligence in plastic surgery. 2024.
7. The use of generative artificial intelligence in surgical education. 2024.
8. Artificial intelligence in plastic surgery: Implications and perspectives. 2024.
9. Generative artificial intelligence in plastic surgery. 2026.
10. Artificial intelligence in plastic surgery: customizing care for each patient. 2024.
11. Impact of Artificial Intelligence (AI) Image Enhancing Filters on Patient Expectations for Plastic Surgery Outcomes. Aesthetic Plast Surg. 2025.
12. The Current Landscape of Artificial Intelligence in Plastic Surgery Education and Training: A Systematic Review. 2025.
13. Artificial intelligence in plastic surgery: Current developments and future perspectives. Plast Aesthet Res. 2023; 10:3.

14. Spoer DL, Kiene JM, Dekker PK, Huffman SS, Kim KG, Abadeer AI, Fan K. A Systematic Review of Artificial Intelligence Applications in Plastic Surgery: Looking to the Future. Plast Reconstr Surg Glob Open. 2022;10: e4608.
15. Seth I, Bulloch G, Lee A. Redefining Academic Integrity, Authorship, and Innovation: The Impact of ChatGPT on Surgical Research. Ann Surg Oncol. 2023; 30:5284-5285.
16. Seth I, Bulloch G, Rozen WM. Applications of Artificial Intelligence and Large Language Models to Plastic Surgery. Aesthet Surg J. 2023;1: sjad210.
17. Chartier C, Watt A, Lin O, Chandawarkar A, Lee J, Hall-Findlay E. Breast GAN: Artificial Intelligence-Enabled Breast Augmentation Simulation. Aesthet Surg J Open Forum. 2021;4: ojab052.
18. Qin F, Gu J. Artificial Intelligence in plastic surgery: Current developments and future perspectives. Plast Aesthet Res. 2023; 10:3.
19. Rokhshad R, Keyhan SO, Yousefi P. Artificial intelligence applications and ethical challenges in oral and maxillo-facial cosmetic surgery: A narrative review. Maxillofac Plast Reconstr Surg. 2023; 45:14.
20. Celi LA, Cellini J, Charpignon ML, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities: a global review. PLoS Digit Health. 2022;1:e0000022.