



Applications of Artificial Intelligence for Facial Aesthetic Surgery: a Systematic Review

Sarasukma Maharani*, Ni Putu Gita Raditya Sanjiwani, I Gusti Agung Ngurah Widya Pramana

Rumah Sakit Tentara Udayana, Indonesia

Email: ayas.maharani@gmail.com*, Sanjiwani60@gmail.com,
agungwidyapramana@gmail.com

Keywords:

artificial intelligence; facial aesthetic surgery; plastic surgery; generative artificial intelligence; computer vision.

ABSTRACT

Artificial intelligence (AI) offers opportunities to improve assessment, visualization, and communication in facial aesthetic surgery; however, its clinical reliability remains uncertain. This study aimed to synthesize contemporary evidence on AI applications, performance, methodological quality, and clinical relevance in facial aesthetic surgery. A systematic review conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines searched PubMed/MEDLINE, SAGE Journals, ScienceDirect, EBSCO, Scopus, and Web of Science for English-language original studies published between January 2022 and August 2026. Eleven studies met the eligibility criteria and were included in the qualitative synthesis and risk-of-bias assessment using an adapted framework informed by the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool. AI applications included apparent-age estimation, automated facial morphometry, rhinoplasty simulation, patient counseling, and social media analysis. AI-based systems quantified postoperative changes in perceived age and facial morphology, generated realistic surgical visualizations, and provided comprehensible patient education responses. However, visual realism did not establish patient-specific predictive accuracy, and improvements in computational performance did not necessarily translate into greater patient satisfaction or psychosocial benefits. Six studies raised some concerns regarding risk of bias, while five were judged to have a high risk of bias. Methodological heterogeneity, limited external validation, proprietary algorithms, and potential demographic biases limited the interpretation and generalizability of the findings. AI represents a promising adjunct to facial aesthetic surgery; however, prospective multicenter validation, transparent reporting, demographically diverse datasets, and patient-centered outcome assessment are needed before routine clinical integration can be recommended with confidence.

INTRODUCTION

Artificial intelligence (AI) is a field of computer science focused on developing algorithms and systems capable of performing complex tasks traditionally associated with human cognitive abilities, including learning, reasoning, decision-making, and pattern recognition (Bajwa et al., 2021; Bekbolatova et al., 2024). Since the COVID-19 pandemic, AI has become increasingly integrated into various fields and aspects of everyday life, including

healthcare. In healthcare, AI is rapidly emerging as a component of clinical practice, diagnostic procedures, continuous patient monitoring, and risk assessment across a growing range of health conditions and interventions (Bajwa et al., 2021; Shen et al., 2019).

In recent years, plastic surgery has emerged as a field increasingly influenced by technological innovation (Copeland-Halperin, 2026; Duong et al., 2024). The visual and data-intensive nature of plastic surgery makes it particularly suitable for AI techniques, including artificial neural networks (ANNs), support vector machines (SVMs), decision trees (DTs), k-nearest neighbors (k-NN), and deep learning architectures such as convolutional neural networks (CNNs). Artificial neural networks and CNNs designed for image analysis have demonstrated strong performance in medical image analysis and diagnostic classification (Han et al., 2018; Litjens et al., 2017; Shen et al., 2017), and their potential applications in plastic surgery include risk assessment, surgical planning, and outcome simulation (Qin & Gu, 2023; Tokgöz & Carro, 2023).

The emergence of AI has also introduced novel applications in medicine, particularly in facial aesthetic surgery (Choi et al., 2023). Facial appearance can substantially influence psychosocial well-being and may represent an important component of overall health. AI offers opportunities to enhance the precision, predictability, and personalization of care (Atiyeh et al., 2023; Bekbolatova et al., 2024). Rhinoplasty, one of the most technically demanding facial aesthetic procedures, has increasingly incorporated deep learning-based approaches for quantifying nasal morphology and symmetry (Bhowmik et al., 2023; Borsting et al., 2020). AI-based age-estimation models have further expanded the potential role of AI by providing quantitative measures of perceived facial rejuvenation (Elliott et al., 2023; Goodyear et al., 2023).

Given the requirements for precision and personalization in facial aesthetic surgery, AI may represent a valuable adjunct to contemporary surgical practice (Al-Dhubaibi et al., 2025; Dumitrascu et al., 2026). With continued advances in AI, researchers have explored its application across various aspects of facial aesthetic surgery, with improving patient education, surgical planning, assessment, and outcomes as common objectives (Bhowmik et al., 2023; Flynn et al., 2021; Fuzi et al., 2023). This systematic review therefore investigates the development, implementation, methodological quality, and clinical applications of AI in facial aesthetic surgery.

Facial aesthetic surgery faces a global challenge in delivering individualized and predictable outcomes while assessing changes in appearance using reproducible and clinically meaningful measures. The scale of this challenge is reflected in the International Society of Aesthetic Plastic Surgery (ISAPS) Global Survey for 2024, released in 2025, which reported more than 17.4 million surgical and approximately 20.5 million nonsurgical aesthetic procedures performed by plastic surgeons worldwide. Overall procedure volume increased by 42.5% over four years, indicating the expanding global reach of aesthetic interventions. This substantial clinical activity reinforces the need for reliable assessment, effective patient communication, and evidence-based surgical planning. Within this context, AI has attracted attention as a potential tool for supporting clinical tasks that depend heavily on visual interpretation and individualized patient expectations.

The relevance of these developments is particularly evident in facial aesthetic practice. According to ISAPS, more than 7.4 million facial and head surgical procedures were performed

worldwide in 2024, representing a 4.3% increase from the previous year. Eyelid surgery became the most frequently performed aesthetic surgical procedure, exceeding 2.1 million procedures, while rhinoplasty accounted for approximately one million procedures and facial fat grafting for approximately 0.9 million. These figures demonstrate substantial demand for interventions targeting facial appearance, although trends vary among procedures. They also underscore the practical importance of tools capable of documenting anatomical changes, supporting realistic discussions of anticipated outcomes, and facilitating consistent postoperative assessment across diverse patient populations and surgical techniques.

Despite this demand, evaluating facial aesthetic outcomes remains challenging because anatomical measurements, perceived age, attractiveness, and patient satisfaction represent distinct dimensions of surgical outcomes. Recent research has explored AI-based facial analysis as a means of quantifying selected outcomes, particularly postoperative changes in apparent age. However, numerical improvements in image-derived measurements cannot automatically be interpreted as corresponding improvements in patients' experiences or satisfaction. The specific challenge is therefore to determine whether AI-generated outputs provide valid and clinically useful information for facial aesthetic care. Addressing this issue requires consideration of what each algorithm measures, how its performance has been validated, and whether the resulting information is appropriate for the intended clinical application.

Previous empirical studies have provided evidence that AI can quantify certain aspects of facial rejuvenation. Goodyear et al. (2023) developed an apparent-age estimation model using 47,394 facial images and evaluated its application in 103 patients undergoing blepharoplasty. The model achieved a validation mean absolute error of 1.38 years, and the authors interpreted the postoperative findings as indicating an approximately two-year rejuvenating effect. Similarly, Chiou et al. (2024) examined 150 patients undergoing lower blepharoplasty-related procedures using four age-estimation models and reported an average postoperative reduction in estimated age of 1.68 years. Nevertheless, predictive performance differed among platforms. Collectively, these findings support the feasibility of computational outcome assessment while demonstrating why apparent-age estimates should be interpreted in relation to the specific model, dataset, and evaluation setting.

Beyond postoperative assessment, generative AI has been investigated for preoperative visualization and patient communication (Landau et al., 2025). Knoedler et al. (2024) evaluated a generative adversarial network (GAN) for rhinoplasty simulation and found that 101 participants correctly identified AI-generated images with a mean accuracy of 52.5%, suggesting that the generated simulations were often visually difficult to distinguish from genuine postoperative photographs. However, this finding primarily demonstrates visual realism and does not independently establish accurate prediction of an individual patient's eventual surgical outcome. In a complementary application, Xie et al. (2023) evaluated ChatGPT responses to nine rhinoplasty consultation questions and found that the responses were understandable and coherent but limited in detail and personalization. These studies highlight promising applications for visualization and patient communication while leaving important questions regarding patient-specific clinical reliability unresolved.

The potential benefits of these technologies must also be considered alongside concerns regarding demographic representation and algorithmic or aesthetic bias. Lim et al. (2023)

examined generative AI outputs related to rhinoplasty, facelift, and blepharoplasty and identified a predominance of female faces with lighter skin tones, together with limited representation of older individuals. Their findings suggest that AI-generated representations of desirable facial appearance may reflect restricted aesthetic assumptions rather than the diversity encountered in clinical practice. In facial aesthetic surgery, this represents an important concern because treatment planning should account for individual anatomy, identity, preferences, and aesthetic goals. Consequently, evaluating AI requires more than assessing image quality or response fluency; it also requires examining which facial characteristics are represented in training and generated data and whether model outputs remain appropriate across diverse patient populations.

Existing evidence syntheses confirm that AI in facial aesthetic surgery is already an established area of research (Abdali et al., 2025; Nogueira et al., 2025; Shoham et al., 2026). Stephanian et al. (2024), in a systematic review incorporating a Scopus search, included 66 studies from 494 initially identified records and examined applications in diagnostic assessment, outcome prediction, patient counseling, and treatment planning. Therefore, the research gap cannot reasonably be characterized as a complete absence of previous systematic reviews. Rather, the studies discussed above indicate an ongoing need to relate reported technical performance to validation quality and clinical relevance. An updated appraisal should examine how the evidence differs across measurement, visualization, and educational applications while determining whether individual study findings adequately support the clinical claims made for each technology.

This need is particularly important because convincing AI-generated outputs may encourage clinical adoption before their limitations are sufficiently understood. Realistic surgical simulations may influence patient expectations even when patient-specific predictive accuracy remains uncertain, while accessible chatbot responses may be interpreted as personalized clinical guidance despite limited patient-specific assessment. Similarly, demographic biases in AI-generated facial images could influence discussions of aesthetic goals. These concerns make rigorous evidence appraisal particularly relevant to elective aesthetic procedures, in which communication, informed decision-making, and expectation management are central to patient care. A contemporary systematic review can help clarify which AI applications have demonstrated measurement or communication capabilities, which require further validation, and which claims extend beyond the outcomes actually evaluated in the available studies.

Against this background, the contribution of the present review is an integrated assessment of contemporary AI applications in facial aesthetic surgery, with particular attention to the relationship between technical performance and clinical usefulness. Its novelty lies in providing an updated, application-specific appraisal of original research published from January 2022 through August 2026. By considering postoperative assessment, automated facial morphometry, preoperative visualization, patient education, and digital information analysis within a single framework, the review seeks to clarify how different AI technologies may support different stages of facial aesthetic care. Examining study design, validation methods, reported outcomes, and risk of bias alongside these applications is intended to provide a clearer assessment of the evidence supporting their use.

Accordingly, this systematic review aims to synthesize and critically assess empirical evidence concerning AI applications in facial aesthetic surgery. Its specific objectives are to identify the procedures and clinical tasks addressed, characterize the AI technologies and input data used, examine reported accuracy and comparative performance, and evaluate methodological limitations affecting interpretation and generalizability. The review also seeks to identify priorities for future research, including external validation, demographic representation, standardized outcome assessment, and alignment with patient-reported outcomes. Its anticipated scientific contribution is a structured assessment of the strengths and limitations of contemporary evidence, while its practical value lies in supporting informed decisions by clinicians and researchers regarding how AI may assist facial aesthetic care and what additional evidence is required for reliable clinical integration.

METHOD

Study Design

This systematic review was conducted in accordance with the PRISMA 2020 guidelines to synthesize and critically assess empirical evidence on the applications and performance of artificial intelligence (AI) in facial aesthetic surgery, with particular emphasis on algorithmic accuracy, predictive validity, clinical utility, and comparative performance against conventional or expert-based assessments.

Eligibility Criteria

a) Population

Patients of any age undergoing or considering facial aesthetic surgery (rhinoplasty, rhytidectomy, blepharoplasty, otoplasty, facial rejuvenation, browplasty, and lip/chin/cheek augmentation).

b) Intervention/Exposure

AI technologies, including machine learning, deep learning, computer vision, preoperative 3D simulation, or automated tools used to plan, perform, or assess procedures.

c) Comparators

Not mandatory (several eligible studies were exploratory/proof-of-concept); where available, comparators included chronological age, manual measurements, surgeon assessment, or non-AI approaches.

d) Outcomes

accuracy of preoperative planning/outcome simulation, surgical precision, objective assessment of aesthetic results/symmetry, and patient satisfaction/workflow efficiency.

e) Study Design, Language, and Period

Cross-sectional, retrospective, prospective, observational, or case-control studies with original data (reviews, meta-analyses, abstracts, and grey literature excluded); English-language publications only; published 1 January 2022-31 August 2026, reflecting the contemporary phase of generative AI and deep-learning-based facial analysis.

Search Strategy and Study Selection

A structured search was conducted across PubMed/MEDLINE, SAGE Journals, ScienceDirect, EBSCO, Scopus, and Web of Science (searched August 2026) using combinations of controlled vocabulary and free-text terms Artificial Intelligence, Machine

Learning, Deep Learning, Computer Vision, Generative AI, LLM combined via Boolean operators with Facial Aesthetic Surgery, Rhinoplasty, Blepharoplasty, Facelift and related terms, supplemented by manual reference tracking. Titles/abstracts were independently screened by two blinded reviewers, with full texts subsequently assessed against eligibility criteria and disagreements resolved by discussion or a third reviewer. After applying all inclusion/exclusion criteria, 11 studies were included in the final qualitative synthesis (PRISMA flow diagram provided in the original manuscript).

Data Extraction and Risk of Bias Assessment

A standardized, pilot-tested extraction form was used to record author/year/country, study design, clinical application, population/dataset, AI modality/model, input data, primary AI application, outcomes/validation measures, main findings, and proposed clinical role, independently by two reviewers and verified by a third. Methodological quality was assessed using an adapted ROBINS-I-informed framework across five domains (selection bias, confounding, AI model/algorithm-related bias, outcome measurement, and selective reporting), performed independently by one reviewer using predefined criteria applied uniformly across all 11 included studies.

RESULTS AND DISCUSSION

Across the 11 included studies, AI was applied across several domains of facial aesthetic practice - postoperative outcome assessment, periocular morphometry, preoperative simulation, generative visualization, patient counseling, and digital information analysis - spanning blepharoplasty, facelift/facial rejuvenation, rhinoplasty, brow lifting, and fat grafting. Study designs and sample sizes were highly heterogeneous, reflecting the breadth of current AI applications but limiting direct cross-study comparison. Study characteristics and AI applications are summarized in Table 1 below.

Table 1. Characteristics of Included Studies and Artificial Intelligence Applications

Study	Country	Design	Application	AI Modality/Model
Goodyear et al., 2023	United States	Retrospective + AI model development/validation	Cosmetic blepharoplasty	Deep CNN (ResNet-34) for apparent-age estimation
Elliott et al., 2023	United States	Retrospective clinical study	Facial rejuvenation/facelift	CNN-based age estimation (FaceX)
Xie et al., 2023	Australia	Observational study	Aesthetic rhinoplasty	ChatGPT LLM for patient counseling
Hebel et al., 2023	United States	Retrospective comparative study	Facial rejuvenation/facelift	FaceReader computer-vision expression analysis
Lim et al., 2023	Australia/Italy	Observational study	Rhinoplasty, blepharoplasty, facelift	ChatGPT-4, Google Bard; DALL-E 2, Midjourney, BlueWillow

Study	Country	Design	Application	AI Modality/Model
Chiou et al., 2024	Taiwan	Retrospective clinical study	Lower ± upper blepharoplasty/fat grafting	Custom FaceAge (Face++, Azure Face, Amazon Rekognition + 1 more CNN)
Knoedler et al., 2024	—	—	Rhinoplasty (GAN-generated postoperative images)	GAN-based image generation
Kreh et al., 2025	—	—	Periorbital surgery (brow lift and related procedures)	4 AI age-estimation platforms
Er Bilir & Eroğul, 2026	Türkiye	Retrospective clinical study	Upper blepharoplasty (dermatochalasis)	MediaPipe Face Mesh + OpenCV (automated landmark detection)
Kambham et al., 2026	Multinational (US-led)	Cross-sectional observational study	Blepharoplasty-related digital information/education	ChatGPT-4.5-turbo; Evernote AI Transcribe
Segura-Bermudez et al., 2026	United States	Retrospective clinical study	Rhinoplasty/ethnicity-specific nasal aesthetics	Text-to-image diffusion model (AlbedoBase XL/SDXL)

Key quantitative outcomes, validation measures, and principal findings for each study are summarized in Table 2 below.

Table 2. Key Outcomes, Validation Measures, and Principal Findings of Included Studies

Study	Key Quantitative Finding
Goodyear et al., 2023	Validation accuracy ~75%, MAE 1.38 years, Pearson r=0.92; patients predicted ~2.52 years younger postoperatively (p<0.01).
Elliott et al., 2023	Apparent-age reduction of ~3.5 years at 3 months and ~4.9 years at 12 months (p<0.001); no technique consistently superior.
Xie et al., 2023	ChatGPT responses to rhinoplasty questions were coherent and understandable but lacked technical nuance/individualization.
Hebel et al., 2023	High-SMAS group: detected happiness increased 1.03%→13.17% (p=0.008); anger decreased 14.66%→0.63% (p=0.032).
Lim et al., 2023	Generative AI images were inconsistent/anatomically unrealistic and disproportionately depicted younger, female, lighter-skinned, low-BMI faces (demographic bias).
Chiou et al., 2024	Face++ performed best (ICC 0.781, MAE 4.82 yrs); mean apparent-age reduction ~1.68 yrs (p<0.0001), greater in men (3.41 vs 1.30 yrs; p=0.0133).
Knoedler et al., 2024	Evaluators correctly identified AI-generated postoperative images only 52.5±14.3% of the time (near chance); ~56±11.8 ms generation time per image set.
Kreh et al., 2025	Overall apparent-age reduction ~1.03 years (p<0.001); brow lift independently associated with ~1.43-year reduction (p=0.031).

Study	Key Quantitative Finding
Er Bilir & Eroğul, 2026	Framework accuracy ~97.5%, F1 ~92.9%, IoU ~0.82, ICC ~0.93; measurement variability ~68% lower than manual assessment; MRD-1, PFH, TPS, BFS all improved postoperatively ($p<0.001$).
Kambham et al., 2026	Overall informational quality low (median GQS ~2, mDISCERN 1-2); AI-human sentiment agreement 82.21% (Cohen's κ 0.694); 42% positive/15% negative/43% neutral comments.
Segura-Bermudez et al., 2026	AI-generated nasal metrics differed significantly across six ethnic groups; only select groups conformed to traditional “ideal” ratios.

Methodological quality, assessed using an adapted ROBINS-I framework across five domains, is summarized in Table 3 below.

Table 3. Risk of Bias Assessment Using an Adapted ROBINS-I Framework

Study	Selection Bias	Confounding	AI/Algorithm Bias	Outcome Measurement	Selective Reporting	Overall
Goodyear et al., 2023	Some concerns	Some concerns	Some concerns	Low	Low	Some concerns
Elliott et al., 2023	Some concerns	High	Some concerns	Low	Low	Some concerns
Xie et al., 2023	High	Some concerns	Some concerns	High	Low	High
Hebel et al., 2023	High	High	Some concerns	Some concerns	Low	High
Lim et al., 2023	Some concerns	Some concerns	High	High	Some concerns	High
Chiou et al., 2024	Some concerns	Some concerns	Some concerns	Low	Low	Some concerns
Knoedler et al., 2024	Some concerns	Low	Some concerns	Low	Low	Some concerns
Kreh et al., 2025	Some concerns	Some concerns	Some concerns	Low	Low	Some concerns
Er Bilir & Eroğul, 2026	Some concerns	Some concerns	Low	Low	Low	Some concerns
Kambham et al., 2026	High	Some concerns	Some concerns	Some concerns	Low	High
Segura-Bermudez et al., 2026	High	Some concerns	High	Low	Low	High

Overall, 6 of 11 studies were rated “some concerns” and 5 were rated “high” risk of bias, with selection bias and AI/algorithm-related bias the most frequently flagged domains; no study achieved an overall “low risk” rating.

Discussion

From Subjective Aesthetic Judgment to Quantifiable Digital Phenotyping

A primary conclusion is AI's growing role in converting subjective aesthetic evaluation into measurable digital metrics. AI-driven age-estimation and computer-vision systems reliably detected postoperative decreases in perceived age and changes in periocular measurements/facial action units, offering objective value where traditional measures struggle (Chiou et al., 2024; Elliott et al., 2023; Goodyear et al., 2023; Hebel et al., 2023; Kreh et al., 2025). However, AI-derived numerical changes require careful interpretation, as measurable improvements do not inherently guarantee higher patient satisfaction, attractiveness, or psychosocial benefit; future research should validate these computational endpoints against recognized outcome tools and patient-reported measures.

Apparent-Age Estimation as an Emerging Surrogate for Rejuvenation Outcomes

Automated age estimation consistently detected significant age reductions across facelift, blepharoplasty, brow lift, and periorbital procedures (Chiou et al., 2024; Elliott et al., 2023; Goodyear et al., 2023; Kreh et al., 2025), but apparent age remains a surrogate metric rather than a definitive aesthetic outcome, as commercial prediction systems are influenced by dataset demographics, lighting, and camera angle and lack standardized biological invariance. Its primary clinical value may lie in longitudinal within-patient assessment using uniform imaging, rather than cross-patient or cross-system comparison, pending enhanced external validation.

Automated Morphometry and Objective Anatomical Outcome Assessment

AI-based facial landmark identification and periocular morphometry (e.g., margin reflex distance, palpebral fissure height, tarsal platform show, bilateral symmetry) offer a stronger anatomical foundation than global age metrics, improving reproducibility and reducing observer variability - particularly valuable in blepharoplasty (Er Bilir & Eroğul, 2026). However, effectiveness depends heavily on image standardization (makeup, edema, lighting, head position can affect detection), and accurate measurement does not guarantee aesthetic desirability; automated morphometry functions best as a precise measurement tool requiring clinical interpretation and multicenter validation.

Generative AI and the Reconstruction of Aesthetic Expectations

Generative AI (GANs, diffusion models) shifts AI's role from postoperative assessment toward predictive visualization, producing postoperative simulations that viewers often cannot distinguish from real photographs - valuable for aligning patient expectations (Knoedler et al., 2024; Ozmen & Schwarz, 2025). However, visual realism does not equal predictive validity, as image generation struggles to account for soft-tissue behavior, scarring, skeletal structure, and healing variation. Generative AI should therefore be presented as a visualization/communication tool rather than an outcome-prediction system unless prospectively validated, with simulated images clearly labeled and uncertainties communicated.

Algorithmic Aesthetics, Ethnic Representation, and Embedded Beauty Norms

Studies reveal ethnic disparities and demographic biases in generative AI outputs, which disproportionately favor younger, female, lighter-skinned, lower-BMI phenotypes (Lim et al., 2023; Segura-Bermudez et al., 2026; Trabilsy et al., 2025; Yassa et al., 2025) - reflecting biased training data rather than universal beauty standards. This is especially significant in

facial cosmetic surgery, where treatment must be tailored to individual anatomy, identity, and cultural background; future research must evaluate models across diverse demographics, disclose training datasets, and ensure AI supports personalized rather than standardized aesthetics.

Large Language Models as Adjuncts to Patient Education

LLMs (ChatGPT, Gemini) generally produced understandable responses to common aesthetic-surgery questions (Kambham et al., 2026; Shiraishi et al., 2024; Sultan et al., 2026; Xie et al., 2023), potentially reducing informational barriers before formal consultation. However, clinical utility is limited by response variability, unverifiable sources, and inability to replicate the relational/diagnostic requirements of consultation (psychological readiness, anatomical suitability, motivation). LLMs should therefore serve as supplementary educational tools, with clinical deployment requiring strict boundaries against providing direct medical advice, transparent AI-use disclosure, and clear escalation pathways to clinicians.

AI Beyond the Consultation Room: Social Media and Digital Patient Behavior

AI-driven social-media sentiment analysis extends AI's utility to mapping the digital landscape surrounding aesthetic procedures, offering a scalable framework to detect misconceptions and health misinformation (Kambham et al., 2026; Marusza et al., 2026), given that engagement often correlates with popularity rather than scientific validity. However, automated content analysis remains constrained by dataset quality and difficulty parsing irony, sarcasm, and cultural nuance, and should complement rather than replace rigorous content-quality evaluation.

Methodological Heterogeneity and the Maturity of the Evidence Base

Substantial heterogeneity across study designs, sample sizes, AI architectures, and endpoints - combined with widespread reliance on proprietary models with opaque training data, small samples, single-center retrospective designs, and limited external validation - precludes direct comparison and definitive conclusions. The current literature should be regarded as an emerging evidence base demonstrating feasibility and technical performance, but not yet establishing that AI improves long-term clinical outcomes, patient satisfaction, safety, or healthcare efficiency.

Toward Clinically Responsible Integration of AI

Future research must shift from demonstrating technical feasibility to establishing clinical utility, through prospective multicenter studies across diverse populations, prioritized external validation, and standardized reporting of model architecture, training data, and performance metrics. Studies should incorporate patient-relevant outcomes (validated satisfaction measures, psychosocial outcomes, complication/revision rates, expectation concordance), evaluate generative models against actual postoperative images rather than visual realism alone, and assess LLMs for safety and consistency beyond readability. Ethical and regulatory considerations - patient privacy, image ownership, informed consent, algorithmic bias, explainability, and medicolegal responsibility - will become increasingly important as AI moves closer to clinical deployment (Rokhshad et al., 2023), particularly given the highly identifiable nature of facial imagery.

CONCLUSION

Artificial intelligence is emerging as a multifaceted adjunct in facial aesthetic surgery, with applications in outcome assessment, facial morphometric analysis, preoperative simulation, patient education, and digital data analysis. Across the reviewed studies, AI-based systems demonstrated potential to quantify postoperative changes in perceived age and facial morphology, facilitate preoperative visualization, and support the delivery and analysis of patient information. Collectively, these applications suggest that AI may improve the consistency and objectivity of selected assessment and communication processes in facial aesthetic practice.

However, clinical adoption remains constrained by methodological heterogeneity, small and predominantly retrospective cohorts, variability across AI platforms and algorithms, limited external validation, and potential demographic and algorithmic biases. Technical performance alone cannot substitute for clinical judgment or establish patient-specific clinical utility. Future research should prioritize prospective multicenter validation, transparent reporting of model development and validation procedures, evaluation across demographically diverse populations, and alignment of AI-generated outputs with clinically meaningful and patient-reported outcomes before widespread integration into routine facial aesthetic practice.

REFERENCES

- Abdali, H. M., Ayoub, A. Y., Omer, R. A., Alanazi, R. A., Alsaedi, B. S., Almutairi, L. B., Alhakami, B. A., Alfifa, A. Y., Alnasser, N. A., & Hakami, A. H. (2025). Applications of artificial intelligence in oral and maxillofacial cosmetic surgery: A systematic review of diagnostic, planning, and outcome assessment tools. *Cureus*, 17(9).
- Al-Dhubaibi, M. S., Mohammed, G. F., Atef, L. M., Bahaj, S. S., Al-Dhubaibi, A. M., & Bukhari, A. M. (2025). Artificial intelligence in aesthetic medicine: Applications, challenges, and future directions. *Journal of Cosmetic Dermatology*, 24(6), Article e70241.
- Atiyeh, B., Emsieh, S., Hakim, C., & Chalhoub, R. (2023). A narrative review of artificial intelligence (AI) for objective assessment of aesthetic endpoints in plastic surgery. *Aesthetic Plastic Surgery*. <https://doi.org/10.1007/s00266-023-03328-9>
- Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: Transforming the practice of medicine. *Future Healthcare Journal*, 8(2), e188–e194.
- Bekbolatova, M., Mayer, J., Ong, C. W., & Toma, M. (2024). Transformative potential of AI in healthcare: Definitions, applications, and navigating the ethical landscape and public perspectives. *Healthcare*, 12(2), Article 125. <https://doi.org/10.3390/healthcare12020125>
- Bhowmik, R. T., Kandathil, C. K., & Most, S. P. (2023). Automating the standardized cosmesis and health nasal outcomes survey classification with convolutional neural networks. *Facial Plastic Surgery & Aesthetic Medicine*, 25(6), 487–493.
- Borsting, E., et al. (2020). Applied deep learning in plastic surgery: Classifying rhinoplasty with a mobile application. *Plastic and Reconstructive Surgery*, 145(1), 194e–200e.
- Chiou, T.-W., Yen, C.-I., Hsiao, Y.-C., & Chen, H.-C. (2024). AI prediction for post-lower blepharoplasty age reduction. *Aesthetic Surgery Journal*, 44(12), NP922–NP930. <https://doi.org/10.1093/asj/sjae182>
- Choi, E., Leonard, K. W., Jassal, J. S., et al. (2023). Artificial intelligence in facial plastic surgery: A review of current applications, future applications, and ethical considerations. *Facial Plastic Surgery*, 39(5), 454–459.

- Copeland-Halperin, L. R. (2026). Artificial intelligence in plastic surgery: Current status, limitations, and future directions. *Aesthetic Surgery Journal*, 46(2), 113–121.
- Dumitrascu, D. I., Popa, S. L., Incze, V., Amarie, D. S., Gaspari, L., Aluas, P., Ismaiel, A., Leucuta, D. C., David, L., Mihaileanu, F. V., & Gherman, C. D. (2026). Artificial intelligence and esthetics: Redefining precision and beauty in plastic surgery. *Medicina*, 62(4), Article 633.
- Duong, T. V., Vy, V. P., & Hung, T. N. (2024). Artificial intelligence in plastic surgery: Advancements, applications, and future. *Cosmetics*, 11(4), Article 109.
- Elliott, Z. T., Bheemreddy, A., Fiorella, M., Martin, A. M., Christopher, V., Krein, H., et al. (2023). Artificial intelligence for objectively measuring years regained after facial rejuvenation surgery. *American Journal of Otolaryngology*, 44(2), Article 103775. <https://doi.org/10.1016/j.amjoto.2022.103775>
- Er Bilir, A., & Eroğul, Ö. (2026). Artificial intelligence-assisted quantitative evaluation of upper eyelid blepharoplasty outcomes in dermatochalasis: Methodological validation and clinical correlation. *BMC Ophthalmology*, 26(1), Article 162. <https://doi.org/10.1186/s12886-026-04693-0>
- Flynn, J. P., Cha, E. L., Flynn, T. J., et al. (2021). Machine learning generated synthetic faces for use in facial aesthetic research. *Facial Plastic Surgery & Aesthetic Medicine*, 23(5), 339–343.
- Fuzi, J., Meller, C., Ch'ng, S., et al. (2023). The emerging role of artificial intelligence tools for outcome measurement in facial reanimation surgery: A review. *Facial Plastic Surgery & Aesthetic Medicine*, 25(6), 556–561.
- Goodyear, K., Saffari, P. S., Esfandiari, M., Baugh, S., Rootman, D. B., & Karlin, J. N. (2023). Estimating apparent age using artificial intelligence: Quantifying the effect of blepharoplasty. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 85, 336–343. <https://doi.org/10.1016/j.bjps.2023.07.017>
- Han, S. S., Kim, M. S., Lim, W., Park, G. H., Park, I., & Chang, S. E. (2018). Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. *Journal of Investigative Dermatology*, 138(7), 1529–1538. <https://doi.org/10.1016/j.jid.2018.01.028>
- Hebel, N. S. D., Boonipat, T., Lin, J., Shapiro, D., & Bite, U. (2023). Artificial intelligence in surgical evaluation: A study of facial rejuvenation techniques. *Aesthetic Surgery Journal Open Forum*, 5, Article ojad032. <https://doi.org/10.1093/asjof/ojad032>
- Kambham, A., Nguyen, K., Moghal, M., Masi, B., Matsui, C., De la Cruz Ku, G., et al. (2026). Artificial intelligence-enhanced evaluation of blepharoplasty content on social media: A cross-sectional study. *Journal of Cutaneous and Aesthetic Surgery*. Advance online publication. https://doi.org/10.25259/JCAS_257_2025
- Knoedler, S., Alfertshofer, M., Simon, S., Panayi, A. C., Saadoun, R., Palackic, A., et al. (2024). Turn your vision into reality—AI-powered pre-operative outcome simulation in rhinoplasty surgery. *Aesthetic Plastic Surgery*, 48(23), 4833–4838. <https://doi.org/10.1007/s00266-024-04043-9>
- Kreh, C. C., Roider, L., Firouzbakht, P. K., Nathan, C., Prada, C. A., Lund, H. G., Jr., et al. (2025). Artificial intelligence analysis of periorbital rejuvenation. *Aesthetic Surgery Journal*, 45(2), 215–220. <https://doi.org/10.1093/asj/sjae201>
- Landau, M., Kassir, M., & Goldust, M. (2025). Transforming aesthetic medicine with generative artificial intelligence. *Journal of Cosmetic Dermatology*, 24(2).
- Lim, B., Seth, I., Kah, S., Sofiadellis, F., Ross, R. J., Rozen, W. M., et al. (2023). Using generative artificial intelligence tools in cosmetic surgery: A study on rhinoplasty, facelifts, and blepharoplasty procedures. *Journal of Clinical Medicine*, 12(20), Article 6524. <https://doi.org/10.3390/jcm12206524>

- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Marusza, M., Krupski, A., Stielow, V., & Zatyka, A. (2026). Digital misinformation in aesthetic dermatology: The role of social media and generative artificial intelligence in shaping patient safety, trust, and evidence-based practice: A narrative review (2022–2026). *Quality in Sport*, 53, Article 70151.
- Nogueira, R., Eguchi, M., Kasmirski, J., de Lima, B. V., Dimatos, D. C., Lima, D. L., Glatter, R., Tran, D. L., & Piccinini, P. S. (2025). Machine learning, deep learning, artificial intelligence and aesthetic plastic surgery: A qualitative systematic review. *Aesthetic Plastic Surgery*, 49(1), 389–399.
- Ozmen, B. B., & Schwarz, G. S. (2025). Critical evaluation of generative text-to-image AI models in plastic surgery: Balancing promise and precision. *Aesthetic Surgery Journal*, 45(9), NP183–NP184.
- Qin, F., & Gu, J. (2023). Artificial intelligence in plastic surgery: Current developments and future perspectives. *Plastic and Aesthetic Research*, 10(1).
- Rokhshad, R., Keyhan, S. O., & Yousefi, P. (2023). Artificial intelligence applications and ethical challenges in oral and maxillo-facial cosmetic surgery: A narrative review. *Maxillofacial Plastic and Reconstructive Surgery*, 45(1), Article 14.
- Segura-Bermudez, J. R., Rezwan, S., Barkhordarzadeh, A. D., Tahir, A. F., Khetpal, S., Zhao, T., et al. (2026). Aesthetically ideal noses created using a single artificial intelligence model: Validating literature and exploring ethnic differences. *JPRAS Open*, 49, 517–528. <https://doi.org/10.1016/j.jpra.2026.03.003>
- Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
- Shen, J., Zhang, C. J. P., Jiang, B., Chen, J., Song, J., Liu, Z., et al. (2019). Artificial intelligence versus clinicians in disease diagnosis: Systematic review. *JMIR Medical Informatics*, 7(3), Article e10010. <https://doi.org/10.2196/10010>
- Shiraishi, M., Tanigawa, K., Tomioka, Y., Miyakuni, A., Moriwaki, Y., Yang, R., Oba, J., & Okazaki, M. (2024). Blepharoptosis consultation with artificial intelligence: Aesthetic surgery advice and counseling from Chat Generative Pre-Trained Transformer (ChatGPT). *Aesthetic Plastic Surgery*, 48(11), 2057–2063.
- Shoham, G., Naveh, S., Confino, I., Zoabi, T., Govrin, O., Fliss, E., & Barnea, Y. (2026). A systematic review of applications, challenges, and future trajectories of artificial intelligence in cosmetic surgery. *Aesthetic Surgery Journal*, 46(2), 150–159.
- Stephanian, B., Karki, S., Debnath, K., Saltychev, M., Rossi-Meyer, M., Kandathil, C. K., & Most, S. P. (2024). Role of artificial intelligence and machine learning in facial aesthetic surgery: A systematic review. *Facial Plastic Surgery & Aesthetic Medicine*, 26(6), 679–705.
- Sultan, M. T., Sanan, M., Wenhui, L., & Li, G. (2026). Accuracy of descriptive and illustrative capacity of generative artificial intelligence in aesthetic surgery: A study of accelerated facial aging in depressed medical residents. *Aesthetic Plastic Surgery*. Advance online publication.
- Tokgöz, E., & Carro, M. A. (2023). Applications of artificial intelligence, machine learning, and deep learning on facial plastic surgeries. In *Cosmetic and reconstructive facial plastic surgery: A review of medical and biomedical engineering and science concepts* (pp. 281–306). Springer Nature Switzerland.
- Trabilsy, M., Genovese, A., Prabha, S., Borna, S., Gomez-Cabello, C. A., Haider, S. A., Tao, C., & Forte, A. J. (2025). Utilizing generative text-to-image artificial intelligence models

- to explore race, gender, and age in plastic and aesthetic surgery. *Aesthetic Surgery Journal*, 45(9), NP172–NP178.
- Xie, Y., Seth, I., Hunter-Smith, D. J., Rozen, W. M., Ross, R., & Lee, M. (2023). Aesthetic surgery advice and counseling from artificial intelligence: A rhinoplasty consultation with ChatGPT. *Aesthetic Plastic Surgery*, 47, 1985–1993. <https://doi.org/10.1007/s00266-023-03338-7>
- Yassa, A., Akhavan, A., Ayad, S., Ayad, O., Colon, A., & Ignatiuk, A. (2025). Facial aesthetics in artificial intelligence: First investigation comparing results in a generative AI study. *Eplasty*, 25, Article e13.