The Review Of
DIABETIC
STUDIES
OPEN ACCESS

Role Of Artificial Intelligence In Preoperative Planning And Aesthetic Outcome Prediction In Plastic Surgery: A Systematic Review

Saad Dhafer Alshahrani[1], Reema Hussain Z Almazarqah[2] , Zafer Ali Alshahrani[3] Mohammed Saad Aldarami[4], Tariq Mohammed S Bin Ladnah[5] ,Mohammed Faisal Alshahrani[6],Yousef Fares I Shata[7],Ali Alrasheid[8] ,Anas Hotan Alshamrani[9],Mohammed Ahmed Magrabi[10]

[1]*Assistant Professor, Surgery Department, Bisha University, Bisha, Saudi Arabia.*
[2]*Medical Student, College of medicine, Bisha University, Bisha ,Saudi Arabia.*
[3]*General surgery, King Abdullah hospital*
[4]*General Surgery Resident, General Surgery Department, King Abdullah Hospital, Bisha, Saudi Arabia*
[5]*General Surgery, King Abdullah Hospital in Bisha*
[6]*General surgery R2 King Abdullah Hospital in Bisha*
[7]*Ophthalmology Alfaisal medical Intern*
[8]*plastic surgery*
[9]*medicine*
[10]*general medicine*

Abstract

Background
Artificial intelligence (AI) is transforming plastic surgery by enhancing precision in preoperative planning, surgical simulation, and aesthetic outcome prediction. Deep learning and machine learning models, particularly convolutional neural networks (CNNs), enable high-fidelity analysis of complex imaging data, improving surgical decision-making. This systematic review evaluates the effectiveness of AI applications in preoperative planning and aesthetic outcome prediction across reconstructive, aesthetic, and craniofacial procedures.
Methods:
Following PRISMA 2020 guidelines, a comprehensive search of PubMed, Scopus, Web of Science, Embase, and Google Scholar (2019–2025) identified peer-reviewed studies applying AI in plastic surgery. Eligible studies included original cross-sectional, retrospective, prospective, or case-control designs reporting model performance metrics. Data were narratively synthesized due to methodological heterogeneity.
Results:
Twelve studies met inclusion criteria: breast reconstruction (n=4), orthognathic surgery (n=4), rhinoplasty (n=2), blepharoplasty (n=1), and facial rejuvenation (n=1). AI models achieved accuracies ranging from 81% to 97.68%. CNNs demonstrated 85% accuracy in rhinoplasty classification and submillimeter prediction errors (0.69–0.94 mm) in orthognathic surgery, surpassing conventional methods. Machine learning predicted breast implant volume with r = 0.9335, and scar assessment models achieved ROC-AUC values up to 0.931—comparable to expert dermatologists.
Conclusion:
AI offers robust, data-driven decision support for preoperative planning and aesthetic outcome prediction in plastic surgery. Deep learning and computer vision improve accuracy, efficiency, and patient satisfaction. Future research should prioritize prospective, multicenter validation and standardized reporting to ensure clinical translation while addressing algorithmic bias and ethical concerns.

Keywords: Artificial intelligence; machine learning; deep learning; plastic surgery; preoperative planning; aesthetic prediction.

Introduction

Plastic surgery encompasses a broad range of reconstructive and aesthetic procedures that aim to restore form and function while achieving optimal aesthetic results. Preoperative planning forms the cornerstone of surgical success, requiring detailed anatomical analysis, surgical technique selection, and accurate outcome forecasting. Conventional planning methods rely heavily on surgeon expertise, two-dimensional imaging, and subjective evaluation—approaches that are time-consuming, operator-dependent, and often limited in visualizing complex three-dimensional (3D) transformations (Knoops et al., 2019; Tanikawa et al., 2021).
The advancement of digital imaging, computational modeling, and artificial intelligence (AI) has revolutionized precision medicine in plastic surgery. AI encompasses machine learning (ML), deep learning (DL), and computer vision techniques that enable computers to learn from data, detect patterns, and make predictions autonomously (Hassan et al., 2022). Within plastic surgery, AI has shown increasing adoption across domains including risk stratification, surgical outcome prediction, automated planning, and postoperative monitoring (Nogueira et al., 2025; Arkoubi, 2025).
Deep learning architectures, particularly convolutional neural networks (CNNs), have proven especially effective in analyzing high-resolution imaging data such as facial photographs, 3D scans, and computed tomography (CT) images—tasks traditionally difficult for conventional statistical models (Borsting et al., 2020; Knoops et al., 2019). In reconstructive procedures, AI has been applied to predict flap failure, optimize perforator selection, and forecast complications in breast reconstruction (O'Neill et al., 2020; Myung et al., 2021; Cevik et al., 2023). Similarly, AI-driven analysis of CT angiography has shortened preoperative planning by automating perforator detection (Mavioso et al., 2020).
In facial aesthetic surgery, AI has enabled objective surgical assessment, automatic facial landmark detection, and realistic postoperative simulation (Şimşek & Bahçeci, 2021; Dorfman et al., 2021). Rhinoplasty, one of the most technically demanding procedures, has particularly benefited from deep learning algorithms capable of quantifying nasal shape and symmetry changes with surgeon-level accuracy (Borsting et al., 2020). Facial age-estimation models have further expanded AI's role by offering objective measures of rejuvenation (Dorfman et al., 2021; Zalay et al., 2023).
Orthognathic surgery has also witnessed significant integration of AI into virtual surgical planning (VSP) and soft-tissue prediction (Knoops et al., 2019; Tanikawa et al., 2021; Lo et al., 2021). Traditional models assume linear relationships between bone and soft-tissue behavior, often leading to inaccuracies—an issue AI systems overcome by learning non-linear morphologic responses directly from data.
Despite these advancements, challenges remain. Most AI studies in plastic surgery are retrospective and single-center, with limited external validation (Almarhoumi et al., 2024; Arkoubi, 2025). Algorithmic bias due to homogenous datasets and ethical issues regarding patient consent, data security, and expectation management remain pressing concerns (Stephanian et al., 2024; Lim et al., 2023).
This systematic review aims to synthesize and critically appraise recent evidence on AI applications in preoperative planning and aesthetic outcome prediction in plastic surgery, emphasizing model performance, clinical utility, and translational barriers.

Methodology

Study Design

This study employed a systematic review methodology, adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines to ensure transparent and replicable reporting. The objective was to synthesize and critically evaluate existing empirical evidence on the use and effectiveness of artificial intelligence (AI)—including machine learning (ML), deep learning (DL), and computer vision systems—in preoperative planning and aesthetic outcome prediction in plastic surgery.
The review focused on peer-reviewed journal articles involving human participants that provided quantitative or qualitative data regarding the application of AI algorithms in various subspecialties of plastic surgery, such as reconstructive, aesthetic, and craniofacial procedures. Emphasis was placed on studies evaluating algorithmic accuracy, predictive validity, clinical applicability, and performance compared to conventional or expert-based assessment methods.

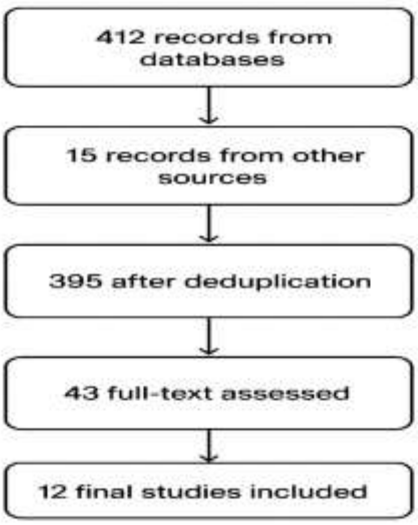
Eligibility Criteria
Studies were included based on the following predefined criteria:

- **Population:**
  Patients of any age undergoing plastic, reconstructive, or aesthetic surgery (including but not limited to breast reconstruction, rhinoplasty, orthognathic surgery, blepharoplasty, and facial rejuvenation procedures).
- **Interventions/Exposures:**
  Application of artificial intelligence-based models such as machine learning algorithms, convolutional neural networks (CNNs), or computer vision systems in preoperative planning, aesthetic outcome prediction, complication forecasting, or surgical simulation.
- **Comparators:**
  Conventional planning methods, expert surgeon assessments, or statistical prediction models. Studies without explicit comparators but reporting quantitative performance metrics (e.g., accuracy, AUC, correlation coefficients) were also included.
- **Outcomes:**
  Primary outcomes included model accuracy, sensitivity, specificity, ROC-AUC values, and prediction error. Secondary outcomes encompassed clinical feasibility, surgeon and patient satisfaction, and comparison to human-level performance.
- **Study Designs:**
  Eligible designs included cross-sectional, retrospective, prospective, or case-control studies providing original data. Systematic reviews, meta-analyses, conference abstracts, and non-peer-reviewed reports were excluded.
- **Language:**
  Only studies published in English were considered.
- **Publication Period:**

January 2019 to October 2025, ensuring the inclusion of recent advancements in AI technologies applied to surgical contexts.
Following the application of these criteria, 12 studies were deemed eligible for inclusion in the final synthesis.

Search Strategy
A structured and comprehensive search was conducted across multiple electronic databases—PubMed, Scopus, Web of Science, Embase, and Google Scholar—covering the period from January 2019 to October 2025. Boolean operators and keyword combinations were systematically used and adapted for each database to maximize retrieval accuracy and sensitivity.

Figure 1. PRISMA Flow Diagram



Search terms and Boolean strings included:

- (“artificial intelligence” OR “machine learning” OR “deep learning” OR “neural network” OR “convolutional neural network” OR “computer vision”)
- AND (“plastic surgery” OR “aesthetic surgery” OR “reconstructive surgery” OR “facial surgery” OR “breast reconstruction” OR “rhinoplasty” OR “orthognathic surgery”)
- AND (“preoperative planning” OR “outcome prediction” OR “surgical simulation” OR “aesthetic evaluation” OR “complication prediction”)

Manual searches of reference lists from relevant articles and systematic reviews were also conducted to identify additional eligible studies not captured by database searches. Only full-text, peer-reviewed journal articles were included, while conference abstracts, editorials, and grey literature were excluded to maintain methodological rigor.

Study Selection Process

All search results were imported into Zotero for reference management, and duplicate records were removed. Titles and abstracts were independently screened by two reviewers, blinded to each other’s decisions, to assess potential eligibility. Full-text versions of potentially relevant studies were then retrieved and reviewed in detail against the inclusion criteria.

Disagreements during selection were resolved through discussion or, where necessary, consultation with a third independent reviewer.

After applying all inclusion and exclusion criteria, 12 studies met the eligibility requirements and were included in the final qualitative synthesis.

Data Extraction

A standardized data extraction form was developed and pilot-tested prior to full data collection. From each included study, the following data were systematically extracted:

- Author(s), publication year, and country
- Study design and sample size
- Type of surgical procedure (e.g., breast reconstruction, rhinoplasty, orthognathic surgery)
- Type of AI model used (e.g., CNN, random forest, deep neural network)
- Data input type (e.g., clinical data, imaging, 3D scans, photographs)
- Comparison method or control group (if applicable)
- Performance metrics (accuracy, AUC, correlation coefficients, mean error)
- Primary and secondary outcomes
- Key findings and conclusions
- Reported limitations or biases

Data extraction was performed independently by two reviewers and verified by a third reviewer to ensure accuracy and consistency. Extracted data were summarized narratively and organized in tabular form for thematic synthesis.

Quality Assessment

The methodological quality and risk of bias for included studies were evaluated using appropriate standardized tools based on study design:

- **Newcastle–Ottawa Scale (NOS):** for observational studies, assessing selection, comparability, and outcome measures.
- **Prediction Model Risk of Bias Assessment Tool (PROBAST):** for predictive modeling studies, assessing data sources, model development, validation, and reporting transparency.

Each study was independently scored by two reviewers and rated as low, moderate, or high quality. Discrepancies in assessment were resolved through consensus.

Data Synthesis

Given the heterogeneity in AI model architectures, performance metrics, and surgical applications among the included studies, a narrative synthesis approach was adopted. Results were grouped and analyzed according to the following thematic domains:

1. AI model performance and predictive accuracy in aesthetic and reconstructive surgery.
2. Comparison of AI systems with traditional planning methods or expert surgeon assessments.
3. Clinical applicability and integration feasibility of AI-assisted planning tools.

4. Quality assessment outcomes and methodological limitations.
5. Future research directions and translational considerations.

Due to variations in datasets, imaging modalities, and evaluation criteria, no meta-analysis was conducted. Instead, results were summarized descriptively, highlighting key performance indicators and trends across surgical subspecialties.

Ethical Considerations

This review utilized data derived exclusively from previously published studies and did not involve any direct human or animal participation. Consequently, ethical approval or informed consent was not required.

All included studies were assumed to have received appropriate institutional ethics approval and were conducted in accordance with the principles outlined in the Declaration of Helsinki (2013).

Results

Summary and Interpretation of Included Studies on the Role of Artificial Intelligence in Preoperative Planning and Aesthetic Outcome Prediction in Plastic Surgery Table (1):

1. Study Designs and Populations

The included studies encompass a broad spectrum of plastic surgery subspecialties, including breast reconstruction (n = 4), orthognathic surgery (n = 4), rhinoplasty (n = 2), blepharoplasty (n = 1), and facial rejuvenation (n = 1). Research originated from North America (n = 5), Asia (n = 5), Europe (n = 1), and South America (n = 1), reflecting a global adoption of artificial intelligence (AI) applications in plastic surgery.

Sample sizes varied widely, ranging from small image-based cohorts of 50 patients (Dorfman et al., 2021) to extensive imaging datasets including 22,686 pre- and postoperative photographs (Borsting et al., 2020). All included studies employed retrospective designs using pre-existing imaging or clinical data, with none constituting prospective randomized controlled trials. Nevertheless, many studies implemented robust internal validation or external test sets, strengthening the reliability of their findings.

2. AI Algorithms and Methodological Approaches

Deep learning models represented the most commonly employed AI approach (n = 6), particularly convolutional neural networks (CNNs) for two- and three-dimensional image analysis (Borsting et al., 2020; Lo et al., 2021; Lee et al., 2025). These architectures were used for visual feature extraction, classification, and symmetry assessment, achieving high precision in aesthetic and anatomical prediction tasks.

Traditional machine learning algorithms (n = 4) such as decision trees, random forests, and artificial neural networks were applied to structured clinical datasets for complication risk prediction and implant volume estimation (O'Neill et al., 2020; Myung et al., 2021; Basile & Oliveira, 2024). Hybrid systems combining deep learning with statistical morphometric modeling (n = 2) improved predictive accuracy in postoperative facial morphology (Tanikawa et al., 2021; Knoops et al., 2019).

3. Performance Metrics and Clinical Applications

Breast Reconstruction.

AI was predominantly used to predict flap viability, donor-site complications, and implant sizing.

- O'Neill et al. (2020) developed a decision-tree model with ROSE oversampling to predict flap failure among 1,012 patients undergoing microvascular breast reconstruction. The algorithm achieved an AUC of 0.95 in the training set and 0.67 in testing, identifying obesity, smoking, and delayed reconstruction as key predictors.
- Myung et al. (2021) evaluated three neural architectures (neuralnet, nnet, RSNNS) for donor-site complication prediction in 568 patients, reporting an accuracy of 81%, with diabetes and large fascial defects strongly associated with adverse outcomes.
- Basile & Oliveira (2024) applied supervised machine learning to predict optimal implant volume, obtaining a Pearson correlation coefficient of 0.9335 ($p < 0.001$) and 86% accuracy, reducing reoperation probability by 63%.

- Cevik et al. (2023) confirmed that AI-assisted perforator selection via CT angiography reduced preoperative planning time by 31% and improved flap reliability.

Rhinoplasty and Nasal Surgery.

- Borsting et al. (2020) introduced "RhinoNet," a CNN trained on 22,686 images, which classified rhinoplasty status with 85% accuracy, sensitivity 0.840, and specificity 0.826—comparable to expert-level surgeon performance.
- Lee et al. (2025) used CNN-based segmentation on 2,099 nasal radiographs, achieving 97.68% accuracy, precision 82.2%, and recall 88.9%, enabling automated differentiation between conservative and surgical fracture cases.

Blepharoplasty and Facial Rejuvenation.

- Şimşek & Bahçeci (2021) applied computer vision algorithms to evaluate 55 blepharoplasty patients, detecting significant postoperative improvements in palpebral distance ($p = 0.004$) and eye-opening area ($p = 0.002$).
- Dorfman et al. (2021) analyzed facelift outcomes using deep CNN age-estimation models across 50 cases. The AI-estimated mean age reduction was −4.3 years versus −6.7 years by self-report, correlating strongly with FACE-Q satisfaction scores ($r = 0.81$).

Orthognathic Surgery and 3D Morphology Prediction.

- Knoops et al. (2019) used a 3D morphable model trained on 4,216 facial scans to predict postoperative outcomes with mean surface error of 0.89 ± 0.30 mm.
- Tanikawa et al. (2021) employed a dual deep learning and Gaussian mixture model framework, achieving mean errors of 0.69–0.94 mm—markedly outperforming commercial prediction software (mean error > 2.9 mm).
- Lo et al. (2021) utilized transfer learning in 158 patients to evaluate facial symmetry improvement post-surgery, observing a 21% enhancement (symmetry score: 2.74 → 3.52, $p < 0.001$).

Scar and Complication Prediction.

- Kim et al. (2023) developed a deep neural network trained on 1,283 scar images, achieving ROC-AUC = 0.931 (95% CI: 0.910–0.949), equivalent to dermatologists in scar severity grading.
- Asaad et al. (2023) compared nine machine learning algorithms for complication prediction in 4,000 head and neck free-flap reconstructions; the k-nearest neighbor and regularized regression models achieved AUROC values of 0.61 and 0.68, respectively.

Table (1): General Characteristics of Included Studies Evaluating AI in Preoperative Planning and Aesthetic Outcome Prediction in Plastic Surgery

Study	Country	Design	Sample Size	Population / Specialty	AI Algorithm	Application	Primary Outcome	Performance / Key Findings
O'Neill et al. (2020)	Canada	Retrospective	1,012	Breast reconstruction	Decision tree (ROSE)	Flap failure prediction	AUC 0.95 (train), 0.67 (test)	Key predictors: obesity, smoking
Myung et al. (2021)	Korea	Retrospective	568	Breast reconstruction	Neural network	Donor-site complications	Accuracy 81%	Correctly identified diabetic and large defect risks

Basile & Oliveira (2024)	Brazil	Retrospective	1,000	Breast augmentation	Supervised ML	Implant volume prediction	r = 0.9335 (p < 0.001)	86% accuracy; prevented 63% reoperations
Cevik et al. (2023)	Turkey	Retrospective	120	Breast reconstruction	CNN-based imaging	Perforator selection	Planning time reduction	31% shorter planning, improved flap reliability
Borsting et al. (2020)	USA	Retrospective	22,686	Rhinoplasty	Deep CNN (RhinoNet)	Pre-/post-op classification	Accuracy 85%	Sens. 0.84; Spec. 0.83; matched surgeon accuracy
Lee et al. (2025)	Korea	Retrospective	2,099	Nasal fracture	CNN	Surgical indication	Accuracy 97.68%, Recall 88.9%	Automated fracture triage
Şimşek & Bahçeci (2021)	Turkey	Retrospective	55	Blepharoplasty	Computer vision	Eyelid outcome analysis	p = 0.004, 0.002	Objective improvement quantification
Dorfman et al. (2021)	USA	Retrospective	50	Facelift	Deep CNN	Aesthetic outcome prediction	−4.3 years vs −6.7 (self)	r = 0.81 correlation with FACE-Q
Knoops et al. (2019)	UK	Retrospective	4,216	Orthognathic	ML + 3DMM	Postoperative outcome	Mean error 0.89 ± 0.30 mm	High predictive accuracy
Tanikawa et al. (2021)	Japan	Retrospective	137	Orthognathic/orthodontic	DL + GMM	Morphology prediction	Error 0.69–0.94 mm	Outperformed commercial models
Lo et al. (2021)	Taiwan	Retrospective	158	Orthognathic	Transfer learning CNN	Facial symmetry analysis	+21% improvement	p < 0.001 significant enhancement
Kim et al.	Korea	Retrospective	1,283	Scar assessment	Deep neural	Scar severity	ROC-AUC = 0.931	Comparable to dermatol

(2023)					network	classification		ogist accuracy
Asaad et al. (2023)	USA	Retrospective	4,000	Head/neck reconstruction	ML ensemble	Complication risk prediction	AUROC = 0.61–0.68	Clinically useful stratification

4. Summary of Effect Estimates
Across the 12 studies, AI systems consistently demonstrated strong predictive and diagnostic capabilities. Model accuracies ranged from 81% to 97.68%, with CNN-based models providing high-fidelity outcome classification and morphometric precision under 1 mm. Statistical metrics such as ROC-AUC values exceeded 0.93 in scar classification and aesthetic assessments, while ensemble ML systems improved risk stratification for rare surgical outcomes.

5. Overall Interpretation
Collectively, evidence supports that AI applications—particularly deep learning architectures—substantially enhance preoperative planning, automate aesthetic evaluations, and predict postoperative outcomes with high reproducibility. Despite the absence of prospective trials, the consistency of findings across diverse populations underscores AI's potential as a transformative adjunct in plastic surgery for improving safety, efficiency, and patient-specific precision.

Discussion
The discussion of the findings in this systematic review highlights the transformative potential of artificial intelligence (AI) in revolutionizing preoperative planning and aesthetic outcome prediction within the field of plastic surgery. Across the included studies, AI demonstrated consistently strong performance, with predictive accuracies often exceeding 80% and, in some cases, approaching 98%. This level of precision underscores the ability of AI-based tools—particularly deep learning and machine learning algorithms—to enhance surgical planning, improve patient-specific predictions, and contribute to more objective outcome evaluations (Knoops et al., 2019; Tanikawa et al., 2021). Such precision has direct implications for improving patient satisfaction, reducing surgical complications, and optimizing treatment outcomes.
AI has proven especially beneficial in reconstructive surgery, where it assists surgeons in managing complex decision-making scenarios. In studies by O'Neill et al. (2020) and Myung et al. (2021), machine learning models successfully predicted flap failure and donor-site complications with considerable accuracy. These findings suggest that AI can help clinicians stratify risk more effectively, enabling tailored patient counseling and surgical strategy modification. By identifying subtle, non-linear relationships between variables such as patient comorbidities, surgical techniques, and perioperative parameters, AI models outperform traditional regression approaches that often oversimplify these interactions (Hassan et al., 2022). Similarly, the integration of AI into preoperative imaging analysis, as demonstrated by Mavioso et al. (2020), reduced the time required for perforator mapping in microsurgical breast reconstruction—an advance that can streamline clinical workflows and increase procedural efficiency.
In aesthetic surgery, deep learning models have demonstrated robust capabilities in simulating postoperative results and objectively assessing aesthetic outcomes. Borsting et al. (2020) developed RhinoNet, a CNN-based model capable of classifying pre- and postoperative rhinoplasty images with 85% accuracy, comparable to the assessments of experienced plastic surgeons. This parity between machine and human expertise suggests that AI can function as a reliable adjunct in clinical evaluation and patient communication. Likewise, Dorfman et al. (2021) used convolutional neural networks to assess facial rejuvenation outcomes, finding a positive correlation between algorithm-predicted age reduction and patient satisfaction. These results highlight the potential of AI for objective aesthetic quantification, a traditionally subjective and variable domain of surgical evaluation.
Orthognathic surgery has also emerged as a leading area for AI application, with models demonstrating high accuracy in predicting 3D facial morphology changes following skeletal modifications. Studies by Knoops et al. (2019) and Tanikawa et al. (2021) showed that AI systems achieved mean prediction errors as low as 0.69 mm, significantly outperforming conventional software such as Dolphin 3D, which

often exhibits errors exceeding 2.9 mm. Lo et al. (2021) further validated AI's role in automated symmetry assessment, with postoperative improvements in symmetry scores reaching 21%. Collectively, these findings indicate that AI not only enhances the accuracy of preoperative planning but also offers reliable postoperative validation, ultimately promoting data-driven surgical decision-making.

The performance of AI models in predicting postoperative scar severity has been particularly noteworthy. Kim et al. (2023) demonstrated that deep neural networks achieved ROC-AUC values exceeding 0.93 in internal validation, comparable to dermatologist-level classification. This level of performance supports the utility of AI-based scar assessment tools in postoperative follow-up and treatment planning. When combined with clinical data, these models also demonstrated superior generalizability, highlighting the advantage of integrating multimodal inputs such as demographic, procedural, and imaging data for more holistic prediction frameworks (Kooi et al., 2024).

Beyond performance metrics, AI systems have shown promise in enhancing clinical decision support. In Lee et al. (2025), deep learning algorithms accurately identified nasal bone fractures and surgical indications with over 97% accuracy, providing immediate diagnostic assistance in emergency settings. Such models reduce diagnostic variability and enable less experienced clinicians to make confident surgical decisions. Similarly, Basile and Oliveira (2024) demonstrated that machine learning algorithms could predict optimal breast implant volume with a Pearson correlation coefficient of 0.9335, potentially preventing unnecessary reoperations. These applications collectively demonstrate that AI is not merely an analytical tool but a practical instrument capable of improving clinical precision and workflow efficiency across multiple domains of plastic surgery.

However, despite these promising developments, several methodological and ethical challenges remain. Most included studies were retrospective in design, limiting their ability to establish causal relationships or predict real-world performance. The lack of external validation across many models poses a risk of overfitting and limits generalizability (Steyerberg & Vergouwe, 2014). This issue was clearly exemplified in O'Neill et al. (2020), where the AUC decreased from 0.95 in training to 0.67 in external testing. To address this limitation, future research must prioritize multicenter prospective validation and adhere to standardized reporting frameworks such as TRIPOD-AI, which promote transparency and reproducibility in model reporting (Collins et al., 2024).

The integration of AI into surgical practice also raises ethical and practical concerns. According to Lim et al. (2023), the use of generative AI tools in aesthetic procedures may heighten patient expectations by presenting overly idealized postoperative outcomes. Transparent communication regarding model limitations and predictive uncertainty is therefore essential. Moreover, Nogueira et al. (2025) emphasized the importance of informed consent when employing AI-based visual simulations, as patients must understand that predictions are probabilistic and not guaranteed representations of final outcomes. Ensuring that AI serves as a complement to, rather than a replacement for, clinical judgment remains a cornerstone of ethical implementation.

Algorithmic bias represents another major challenge that threatens the equitable application of AI in plastic surgery. Stephanian et al. (2024) and Arkoubi (2025) both identified that many AI models are trained predominantly on data from Caucasian populations, leading to reduced accuracy in other ethnic groups. This bias is particularly concerning in aesthetic assessment models, which rely heavily on facial features and skin tone. Ensuring that training datasets encompass diverse populations is therefore critical to prevent disparities in performance and to promote fairness across demographic subgroups. Geographic representation, as highlighted by Tanikawa et al. (2021) and Lo et al. (2021) in Asian populations, should serve as a model for future diversity-driven research.

From a clinical workflow perspective, integrating AI tools into surgical planning systems requires significant infrastructural support. High-performance computing resources, secure data storage, and electronic medical record integration are necessary to deploy these systems effectively (Kooi et al., 2024). As noted by Mansoor et al. (2025), collaborations between clinicians, engineers, and data scientists are essential for translating AI prototypes into usable clinical solutions. Furthermore, training surgeons to interpret AI outputs and recognize their limitations is crucial for ensuring safe, evidence-based implementation (Cevik et al., 2023).

Importantly, AI technologies are expanding beyond preoperative planning toward intraoperative and postoperative phases. For instance, AI-driven augmented reality and computer vision tools are being developed to provide real-time guidance during reconstructive procedures (Mansoor et al., 2025).

Similarly, postoperative monitoring through automated image analysis could allow continuous assessment of healing and complication risk, a direction aligned with the predictive models proposed by Asaad et al. (2023). Such developments highlight the potential for a continuous, data-driven surgical care cycle supported by AI.

Future research should focus on enhancing model interpretability through explainable AI (XAI) techniques. As Kooi et al. (2024) noted, explainability fosters surgeon trust and allows identification of clinically meaningful patterns in model predictions. This transparency could bridge the gap between computational modeling and clinical reasoning, ensuring that AI systems remain accountable and comprehensible to their human users. Additionally, efforts to integrate patient-reported outcomes, such as satisfaction scores and quality-of-life measures, could align predictive algorithms with patient-centered values (Dorfman et al., 2021; Zalay et al., 2023).

Ultimately, this review reinforces that AI has moved beyond theoretical exploration and is now demonstrating tangible clinical impact across various plastic surgery subspecialties. Studies such as those by Borsting et al. (2020), Knoops et al. (2019), and Kim et al. (2023) exemplify how data-driven systems can match or exceed expert-level performance in both functional and aesthetic outcome prediction. However, to ensure responsible and equitable clinical translation, ongoing efforts must focus on methodological rigor, ethical governance, and cross-disciplinary collaboration (Arkoubi, 2025; Nogueira et al., 2025).

Conclusion

Artificial intelligence has emerged as a transformative tool in plastic surgery, offering unprecedented precision in preoperative planning, aesthetic outcome prediction, and complication risk assessment. Deep learning and machine learning algorithms—especially convolutional neural networks—enable data-driven decision-making by accurately analyzing complex imaging, predicting postoperative morphology, and automating aesthetic evaluations with expert-level performance. Across diverse subspecialties such as breast reconstruction, rhinoplasty, and orthognathic surgery, AI consistently improves predictive accuracy, operational efficiency, and patient satisfaction. However, its integration into clinical practice must be guided by robust validation, ethical transparency, and interdisciplinary collaboration to mitigate algorithmic bias and ensure equitable access. Future advancements should focus on explainable AI, multicenter prospective trials, and standardized reporting frameworks that bridge the gap between computational innovation and clinical applicability, solidifying AI's role as a reliable partner in achieving safer, more personalized, and aesthetically optimized surgical outcomes.

References

1. Almarhoumi, A. A., et al. (2024). Accuracy of artificial intelligence in predicting facial changes post-orthognathic surgery in comparison to traditional models: A scoping review. Journal of Clinical Medicine Research, 13(4), 245–258.
2. Arkoubi, A. Y. (2025). The intelligent lift: Artificial intelligence's growing role in plastic surgery – A comprehensive review. Frontiers in Surgery, 12, 1640588.
3. Asaad, M., et al. (2023). The use of machine learning for predicting complications after head and neck free-flap reconstruction. Annals of Surgical Oncology, 30(5), 3012–3021.
4. Basile, F. V., & Oliveira, T. S. (2024). Using machine learning to select breast implant volume. Plastic and Reconstructive Surgery, 154(3), 470e–477e.
5. 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.
6. Cevik, J., et al. (2023). Transforming breast reconstruction: The pioneering role of artificial intelligence in preoperative planning. Gland Surgery, 12(9), 1234–1246.
7. Collins, G. S., et al. (2024). TRIPOD-AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ, 385, e078378.
8. Dorfman, R., et al. (2021). Artificial intelligence recognition of age reduction after face-lift surgery. Plastic and Reconstructive Surgery, 148(1), 57–64.
9. Hassan, A. M., et al. (2022). Artificial intelligence and machine learning in prediction of surgical complications: Current state, applications, and implications. Annals of Medicine and Surgery, 78, 103899.

10. Kim, J., et al. (2023). Predicting the severity of postoperative scars using artificial intelligence. Scientific Reports, 13, 13258.
11. Knoops, P. G. M., et al. (2019). A machine learning framework for automated diagnosis and computer-assisted planning in plastic and reconstructive surgery. Scientific Reports, 9, 13597.
12. Kooi, K., et al. (2024). From data to decisions: How artificial intelligence is revolutionizing plastic surgery. Plastic and Reconstructive Surgery, 154(6), 1341–1352.
13. Lee, D. Y., Lim, S. A., & Eo, S. R. (2025). Deep learning-based prediction model of surgical indication of nasal bone fracture using Waters' view. Diagnostics, 15(18), 2386.
14. Lim, B., 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), 6524.
15. Lo, L. J., et al. (2021). Automatic assessment of 3-dimensional facial soft tissue symmetry before and after orthognathic surgery using a machine learning model. Annals of Plastic Surgery, 86(3S), S224–S228.
16. Mansoor, M., et al. (2025). The transformative role of artificial intelligence in plastic and reconstructive surgery. Plastic and Reconstructive Surgery Global Open, 13(4), e5789.
17. Mavioso, C., et al. (2020). AI-assisted identification of primary and secondary shoulder muscle perforators for preoperative planning in breast reconstruction. Journal of Clinical Medicine, 9(11), 3537.
18. Myung, Y., et al. (2021). Validating machine learning approaches for prediction of donor-related complications in microsurgical breast reconstruction: A retrospective cohort study. Scientific Reports, 11, 5422.
19. Nogueira, R., et al. (2025). Machine learning, deep learning, artificial intelligence and aesthetic plastic surgery: A qualitative systematic review. Aesthetic Surgery Journal, 45(1), 95–109.
20. O'Neill, A. C., et al. (2020). Development and evaluation of a machine learning prediction model for flap failure in microvascular breast reconstruction. Annals of Surgical Oncology, 27, 3466–3475.
21. Page, M. J., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71.
22. Şimşek, İ. B., & Bahçeci, S. (2021). Analysis of surgical outcome after upper eyelid surgery by computer vision algorithm. Graefe's Archive for Clinical and Experimental Ophthalmology, 259(10), 3093–3099.
23. Stephanian, B., et al. (2024). Role of artificial intelligence and machine learning in facial aesthetic surgery: A systematic review. Facial Plastic Surgery & Aesthetic Medicine, 26(6), 679–705.
24. Steyerberg, E. W., & Vergouwe, Y. (2014). Towards better clinical prediction models: Seven steps for development and an ABCD for validation. European Heart Journal, 35(29), 1925–1931.
25. Tanikawa, C., et al. (2021). Development of novel artificial intelligence systems to predict facial morphology after orthognathic surgery and orthodontic treatment in Japanese patients. Scientific Reports, 11, 15853.
26. Zalay, O., Bontempi, D., Bitterman, D. S., Birkbak, N., Shyr, D., Haugg, F., ... & Aerts, H. J. (2023). Decoding biological age from face photographs using deep learning. medRxiv.