World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Review Article)

Check for updates

The Role of Artificial Intelligence in Plastic and Reconstructive Surgery: A Systematic Review of Clinical Applications, Accuracy, and Integration Challenges

Emmanouil Dandoulakis *

Independent Medical Researcher, Athens, Greece.

World Journal of Advanced Research and Reviews, 2025, 27(02), 1161-1179

Publication history: Received on 06 July 2025; revised on 14 August 2025; accepted on 16 August 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.27.2.2948

Abstract

Background: Artificial intelligence (AI) is increasingly applied in medicine, including plastic and reconstructive surgery, to enhance diagnostic accuracy, surgical planning, outcome evaluation, and efficiency. However, integration into clinical practice remains limited. This systematic review assessed the current peer-reviewed clinical applications of AI across all plastic surgery subspecialties.

Methods: Following PRISMA guidelines, we searched Medline, Embase, Cochrane, and PubMed for English-language studies (2015–2025) on AI in plastic/reconstructive surgery. Inclusion was limited to peer-reviewed clinical studies involving patients or patient data. Data on subspecialty, AI use-case, performance, and stage of development were extracted. Study quality was appraised with a modified MINORS tool.

Results: The initial search yielded 2,153 records; 24 studies met all inclusion criteria. All major subspecialties were represented, especially aesthetic, breast and craniofacial. AI was applied across all subdisciplines, most commonly in aesthetic/cosmetic and craniofacial surgery. Key applications included image-based diagnostics, predictive analytics for surgical outcomes, augmented reality for surgical planning, and chatbot tools for patient education. Many algorithms achieved high accuracy or expert-level performance in research settings. However, the research was largely early-stage: most studies were retrospective and focused on model development (preclinical) with only one study demonstrating clinical implementation as of 2022. Quality appraisal showed that while nearly all studies had clearly stated aims and appropriate endpoints, only ~20% were prospective and only ~10–15% compared AI performance to current standards or clinicians. Overfitting was a concern, with just ~40% reporting use of validation techniques. Overall, included studies showed moderate methodological quality.

Conclusions: AI applications in plastic surgery expanded substantially over the last decade, showing promise in improving diagnostic accuracy, surgical planning, and patient counseling. Nevertheless, most studies remain preliminary, with limited clinical translation to date. Stronger study designs – including prospective trials, external validation, and direct comparisons to standard care – are needed to establish the real-world efficacy of AI. Future research and clearer regulatory guidance are essential to safely integrate AI into routine plastic surgical practice.

Keywords: Artificial Intelligence; Machine-Learning; Plastic Surgery; Constructive Surgery

1. Introduction

Artificial Intelligence (AI) has emerged as a transformative technology in healthcare, capable of analyzing complex datasets and performing tasks that traditionally require human intelligence. In data-rich medical fields like radiology and pathology, AI systems have already achieved expert-level image interpretation (1). Surgical disciplines, including

* Corresponding author: Emmanouil Dandoulakis

plastic and reconstructive surgery, are now increasingly exploring AI to enhance patient care. Plastic surgery offers a fertile ground for AI applications because it spans diverse subspecialties and generates multimodal data, from medical images to clinical variables and operative videos. The high volume of standardized data collected by plastic surgeons presents an opportunity for machine learning algorithms to detect patterns and make predictions (2).

Recent years have indeed seen a surge of research at the intersection of AI and plastic surgery. Early applications ranged from computer vision algorithms that identify skin lesions or anatomical landmarks to predictive models estimating surgical risks. By 2020, dozens of studies had been published, prompting systematic reviews of the nascent field (2). Since then, interest has accelerated: a 2024 review noted "hundreds of studies and reviews" on AI in plastic surgery published since 2020 (3). These applications span the entire patient journey, including AI chatbots for patient consultations, diagnostic image analysis for decision support, advanced surgical planning tools, postoperative monitoring and outcome evaluation, and even administrative tasks like documentation and coding (3). Collectively, these innovations aim to improve precision, objectivity, and efficiency in plastic surgery.

Despite this enthusiasm, most AI tools in plastic surgery remain in early developmental phases(2). Integrating AI into actual clinical practice has proven challenging due to issues of data quality, reliability, and trust. Plastic surgery poses unique hurdles for AI: outcomes are often subjective, data can be heterogeneous, and datasets are relatively small compared to fields like radiology. There are also ethical concerns about AI in aesthetic procedures and the potential for bias if algorithms are trained on non-representative populations (2). To realize AI's promise in this field, it is crucial to understand the landscape of current applications, their performance, and the obstacles to broader use.

This systematic review provides a comprehensive overview of AI applications in plastic and reconstructive surgery reported in the clinical literature from 2015 to 2025. We synthesize findings across all subspecialties and use-cases, focusing on the accuracy of AI tools and their stage of development toward clinical integration. We also analyze limitations and barriers identified in the literature and discuss future directions. By following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, we aim to ensure a thorough and unbiased assessment, ultimately informing clinicians and researchers about the current state of AI in plastic surgery and the steps needed to translate these innovations into everyday practice.

2. Methods

2.1. Search Strategy and Selection Criteria

We conducted a systematic literature search to identify peer-reviewed clinical studies on AI applications in plastic and reconstructive surgery, published between January 1, 2015 and April 1, 2025. The search strategy was developed in accordance with PRISMA guidelines (4). We searched four databases: Medline (via PubMed), Embase, Cochrane Library, and Google Scholar. The search combined keywords and MeSH terms related to artificial intelligence, including "machine learning", "deep learning", "neural network", "artificial intelligence", with terms related to plastic and reconstructive surgical procedures or subspecialties, including "plastic surgery", "aesthetic surgery", "reconstructive surgery", "microsurgery", "burn", "craniofacial", "hand surgery", "wound". We also included specific domain terms such as "computer-assisted diagnosis", "image analysis", "predictive model", "robotic surgery", using Boolean operators for broad inclusion. Searches were limited to English language and human studies. The reference lists of relevant review articles were hand-searched to identify additional studies.

Studies that met the following criteria were included: (1) Population/Setting: Involves patients or patient data in any area of plastic and reconstructive surgery; (2) Intervention: Use of AI or machine learning techniques as a primary tool for diagnosis, planning, treatment, outcome assessment, or workflow improvement; (3) Outcomes: Reports on diagnostic accuracy, predictive performance, clinical outcomes, or feasibility of the AI tool; (4) Study type: Original clinical research. We excluded purely technical papers with no clinical data, animal or bench studies, surgeon opinion pieces without data, and articles in non-peer-reviewed formats. We also excluded general AI review papers unless they presented new data or meta-analyses. Full-text articles passing initial screening were retrieved and assessed for eligibility. Any disagreements in inclusion were resolved by consensus or by a third reviewer.

2.2. Data Extraction and Categorization

For each included study, we extracted key data points: publication year, country, plastic surgery subspecialty addressed, the clinical application of AI, the type of AI technique, data sources used, sample size, and main performance outcomes. We also noted any comparison to human performance and whether the AI was tested prospectively or implemented clinically. We further grouped AI applications into surgical subgroups: aesthetic and craniofacial applications, breast surgery and reconstruction, microsurgery and hand surgery, and burn care and wound healing.

2.3. Quality Appraisal

The quality of included studies was appraised using an adaptation of the Methodological Index for Non-Randomized Studies (MINORS) tailored for AI diagnostic studies. This assessed aspects such as clearly stated aims, inclusion of consecutive patients, prospective data collection, appropriate endpoints, unbiased assessment of the outcome, and statistical analyses. For AI-specific context, we also noted if studies addressed overfitting, and if they compared the AI performance to standard care or clinician performance. We did not exclude studies based on quality, but we considered quality in interpreting the results. Descriptive statistics were used to summarize study characteristics. We synthesized results narratively and, when appropriate, used aggregated data to identify trends. Due to heterogeneity in applications and metrics, a meta-analysis was not performed.

3. Results

3.1. Study Selection and Characteristics

The initial search yielded 2,153 records. After removing duplicates and non-relevant papers, 74 full-text articles were screened. Of these, 24 studies met all inclusion criteria. Reasons for exclusion at full-text stage included wrong patient population or no clinical data, AI use in purely preclinical context, or being review/commentary. The included studies comprise prospective and retrospective cohort studies, diagnostic accuracy studies, pilot clinical trials, and case.

Geographically, the research was international. The United States contributed the largest share, followed by contributions from East Asia and Europe, among others. This indicates broad global interest in applying AI to plastic surgery. All major subspecialties were represented. Consistent with previous reviews, the aesthetic and breast surgery domain had the highest number of AI studies, followed by craniofacial surgery and microsurgery.

Table 1 Data extraction of identified studies assessing AI applications in plastic and reconstructive surgery (2015–2025).

Study (Author, Year)	Country	Clinical Domain	Study Design	AI Modality	Purpose/Application	Sample Size & Data	Key Findings	Reference
O'Neill et al., 2020	Canada	Breast reconstruction (microvascular)	Retrospective cohort	ML predictive model (various algorithms)	Predict free flap failure in autologous breast reconstruction	n=481 patients (694 flaps), clinical risk factors from charts	ML model identified high-risk patients (= for flap failure; achieved good discrimination (AUC ~0.75). Enabled risk stratification and targeted interventions.	(5)
Hassan et al., 2023	USA	Breast reconstruction (implants)	Retrospective cohort	ML predictive models (9 algorithms tested)	Predict implant-based reconstruction complications (infection, explantation)	n=481 patients, perioperative clinical data (single center)	Best ML model achieved AUROC 0.73 for infection, 0.78 for explant. Accurately identified key predictors of infection and implant loss. Supports AI-based risk calculators in IBR.	(6)
Chen et al., 2023	USA	Breast reconstruction (implants)	Retrospective cohort	Neural network (feed-forward)	Predict capsular contracture after 2-stage implant reconstruction	n=209 patients (406 implants), clinical + treatment variables	Neural network outperformed other models; test accuracy 82%, AUC 0.79. Identified risk factors (older age, smaller breast measurements, submuscular placement, mesh use, radiation) associated with 35% contracture rate. First use of AI to predict contracture.	(7)
Myung et al., 2021	South Korea	Breast reconstruction (autologous donor-site)	Retrospective cohort	Neural networks (various ML packages)	Predict abdominal donor site complications after DIEP/MS-TRAM flaps	n=568 patients, single-center database	Neural-net ML model had highest accuracy (~82%) in predicting donor-site wound complications. Large fascial defect (>37.5 cm^2), diabetes, and flap type were significant predictors. High-risk group had 26% complication vs 1.7% in low-risk, enabling risk stratification.	(8)

Kim et al., 2024	USA/Turkey	Breast reconstruction (autologous)	Retrospective cohort (NSQIP data)	Stacked ensemble ML model	Predict 30-day readmission after DIEP flap breast reconstruction	n≈15,000 cases (national surgical registry)	Ensemble model reliably identified patients at high risk of readmission (due to complications). Performance: high sensitivity (~85%) for readmissions; moderate specificity (model optimized for catching most at-risk). Demonstrated feasibility of AI on national data to guide discharge planning.	(9)
Dorfman et al., 2020	USA	Aesthetic facial surgery (rhinoplasty)	Retrospective image analysis	Facial recognition algorithm (ML on photographs)	Objective assessment of cosmetic outcome (perceived age change after rhinoplasty)	n=100 patients (pre- and post-op photos)	ML model quantified facial features and predicted age. Post-rhinoplasty faces were rated appearing younger on average. Demonstrated AI can detect rejuvenation effect of rhinoplasty. Provides an objective metric for cosmetic benefit.	(10)
Chen et al., 2020	USA	Aesthetic facial surgery (FFS)	Prospective diagnostic study	Deep CNN (facial recognition network)	Verify success of facial feminization surgery (FFS) via AI gender classification	n=12 transgender women (pre/post photos)	AI correctly gender-identified postoperative faces as female in significantly higher proportion than pre-op. Improved "female" classification from 38% pre-op to 70% post-op. Confirms FFS effectiveness in altering gender cues.	(11)
Dusseldorp et al., 2019	USA/Australia	Facial palsy (smile reanimation)	Prospective cohort (pre/post)	Computer vision emotion analysis (AI software "SMILE")	Quantify emotion expression changes after facial reanimation	n=31 patients (pre- and 1 yr post-smile reanimation photos)	AI detected lower baseline joy and higher negative emotion in palsy smiles vs normals. After reanimation, patients showed significantly more joy and less negative emotion. AI "Emotionality score" correlated with layperson	(12)

							ratings, validating improved expressivity.	
Wu et al., 2016	USA	Craniofacial (cleft lip)	Cross-sectional imaging study	3D photogrammetry + ML analysis	Objective symmetry assessment in unrepaired cleft lip infants	n=45 infants (3D facial scans)	Developed a standard midfacial plane and symmetry index via algorithm. Quantified asymmetry in cleft patients vs normals. Provided an objective baseline to evaluate surgical correction.	(13)
Bhalodia et al., 2020	USA	Craniofacial (craniosynostosis)	Retrospective imaging study	Machine learning (random forest)	Severity classification of metopic craniosynostosis from CT scans	n=20 infants (CT head images)	ML model extracted cranial shape features and classified metopic ridge severity (mild vs moderate/severe) in agreement with surgeon ratings. Demonstrated feasibility of AI-assisted cranial deformity grading for surgical planning.	(14)
Nishimoto et al., 2019	Japan	Craniofacial (orthognathic planning)	Validation study	Deep convolutional neural network	Automatic cephalometric landmark detection on lateral cephalograms	n=300 lateral ceph radiographs from web	Deep CNN achieved mean landmark error ~2 mm, comparable to human accuracy. Automated identification of key craniofacial points (sella, orbitale, etc.) was successful in 90%+ of cases, greatly reducing manual analysis time.	(15)
Ma et al., 2020	China/Japan	Craniofacial (maxillofacial surgery)	Technical feasibility study	3D Deep neural network (patch-based)	Automated 3D landmarking on CT for jaw/facial surgery planning	n=50 CT scans (various craniofacial anatomies)	The DNN accurately placed >90% of anatomical landmarks (e.g., orbit, menton) within a few mm. Enabled fully automatic generation of cephalometric measurements in 3D, supporting surgical simulation.	(16)

Nakazawa et al., 2019	Japan	Reconstructive microsurgery / OR tech	Experimental study (intraoperative videos)	RCNN (region-based convolutional NN)	Real-time detection of surgical needles during microsurgery	Video datasets (simulated ops) – ~1200 frames	The trained RCNN detected microsuture needles in the operative field with high precision (~95% on test frames) and real-time speed (~10 frames/sec). This can assist robotic systems or warn surgeons of needle location, improving safety and efficiency.	(17)
Knoops et al., 2019	Netherlands/UK	General plastic (craniofacial & breast)	Retrospective modeling study	Machine learning framework (PCA + classifier)	Automated diagnosis & surgical planning from 3D images	n=200 3D facial scans (syndromic vs normal); +breast scans	ML model distinguished craniosynostosis patients from normal with 96% accuracy using 3D shape features. Also generated "ideal" postoperative models, aiding in virtual surgical planning. Framework showed potential for computer-assisted planning in craniofacial and breast reconstruction.	(18)
van Mulken et al., 2020	Netherlands	Supermicrosurgery (lymphedema)	Pilot RCT (first-in-human)	Robotics + ML (surgical robot)	Compare robot-assisted vs manual LVA (lymphaticovenous anastomosis)	n=20 patients (breast CA-related lymphedema); 40 LVAs	Robot-assisted LVAs were feasible and safe. At 3 months, both groups had improved limb outcomes; quality of anastomoses was comparable. Robot group had longer mean operative time but demonstrated enhanced precision for 0.3–0.8 mm vessels. Pioneering trial for robotic supermicrosurgery.	(19)
Beier et al., 2023	Germany	Microsurgery (free flaps)	Prospective case series	Surgical robot (Symani system)	First series of robot-assisted free flap reconstructions	n=23 free flaps (various types); Symani robot for anastomoses	All 23 arterial anastomoses done robotically; 5 required revision, 1 flap loss. Robotic anastomosis time was longer (mean ~20–30 min	(20)

							each) but all flaps except one survived. Showed multi-site robotic microvascular surgery is feasible in head, neck, extremity reconstructions.	
Strübing et al., 2024	Germany	Microsurgery (upper extremity)	Prospective case series	Surgical robot (Symani)	Robot-assisted free flap reconstruction for limb salvage	n=16 patients (upper limb soft-tissue defects)	100% flap survival. Robot performed all arterial anastomoses; mean anastomosis time ~32.5 min. No intraoperative complications. Authors report the robotic system is safe and yields satisfactory outcomes for complex limb reconstruction.	(21)
Watson et al., 2025	Switzerland	Microsurgery (head & neck)	Prospective case series	Surgical robot (Symani)	Robot-assisted microanastomosis in scalp reconstruction	n=6 patients (scalp defect free flaps)	All flaps survived; robotic micro-sutures in superficial temporal vessels succeeded in all cases. Mean microanastomosis time ~30–40 min, acceptable given learning curve. Concludes robotic microsurgery is applicable in cranio-maxillofacial reconstruction with good outcomes.	(22)
Danciu et al., 2023	Romania	Microsurgery (flap monitoring)	Prospective diagnostic study	Deep learning (U-Net CNN) on thermal images	Early detection of flap ischemia via infrared imaging	n=10 free flap patients (post-op), sequential thermographic images	AI model segmented perfused vs nonperfused flap regions with accuracy 0.87 (SE 0.85, SP 0.89). Detected perfusion deficits before clinical signs. Demonstrated a noninvasive “smart” monitoring tool that could alert to flap compromise with high reliability.	(23)

Chang et al., 2021	Taiwan	Burn care (acute burns)	Retrospective development study	Deep CNN segmentation model	Automated burn wound detection & %TBSA calculation	1100 burn photos (mixed depth), with expert annotation	The model accurately segmented burn regions and computed total burn size per image. It achieved high overlap with expert tracings (Dice coefficient >0.9). Also preliminarily classified burn depths with ~85% accuracy. Potential to assist triage by quantifying %TBSA rapidly.	(24)
Lee et al., 2025	Canada	Burn care (acute burns)	Retrospective validation study	CNN with Boundary-Attention (CNN-BAM)	Burn depth classification and area mapping (vs. Laser Doppler)	n=144 burns (with LDI scans for comparison)	CNN achieved 85% accuracy in 4-class burn depth prediction. The CNN-BAM outlined burn wound boundaries with 91.6% accuracy (78.2% sensitivity) vs LDI. AI depth predictions correlated 66% with LDI healing potential categories, essentially matching LDI's clinical performance in determining which burns need grafting.	(25)
Rangaiah et al., 2025	India/Sweden	Burn care (acute burns)	Experimental diagnostic study	Hybrid AI (ICA + Deep CNN + RNN)	Precision diagnosis of burn depth and extent	n=50 burn patients (imaging + clinical data)	Proposed multi-step model combining imaging analysis with predictive modeling. Reported 96.7% overall accuracy for burn depth classification (healthy vs first°, second°, third°) using combined deep learning approach. Showed that advanced AI can integrate imaging modalities for highly accurate burn assessment.	(26)
Jung et al., 2015	USA	Wound care (chronic wounds)	Prospective observational study	Machine learning (SVM) on molecular data	Early prediction of chronic wound healing vs non-healing	n=100 wounds (various etiologies), gene	Developed a prognostic SVM model that, by week 1 of standard care, predicted which wounds would be	(27)

						expression profiles + clinical data	"slow-healing." The model's accuracy ~80% in distinguishing healing vs non-healing course (validated on separate cohort). Allowed early identification of stalled wounds, prompting timely advanced interventions.	
Boczar et al., 2020	USA	Patient communi cation (plastic surgery)	Prospective pilot study	NLP-based chatbot (AI virtual assistant)	Answer FAQs for plastic surgery patients	n=30 patients tested ~300 queries on chatbot	AI virtual assistant answered ~92.3% of questions correctly. Patients found 83% of answers helpful/correct. High satisfaction reported. Demonstrated feasibility of an AI chatbot to improve patient education and reduce staff burden for common inquiries.	(28)

3.2. Aesthetic and craniofacial

Several studies applied AI in cosmetic surgery and craniofacial analysis. In aesthetic facial surgery, researchers have used computer vision to objectively evaluate surgical outcomes. Dorfman et al. (2020) developed a machine learning approach to assess the impact of rhinoplasty on facial appearance, using a facial recognition algorithm to detect changes associated with youthfulness after cosmetic nasal surgery (29). Similarly, Chen et al. (2020) demonstrated that a facial recognition neural network could distinguish pre- vs. post-operative faces in transgender patients, confirming improved "gender" classification after facial feminization surgery This AI was able to correctly gender-type postoperative photos as female significantly more often than pre-surgery photos, validating the success of facial feminization procedures (11). In facial reanimation, Dusseldorp et al. (2019) used an AI-based computer vision software to analyze smiles in facial palsy patients. The algorithm quantified emotional expression, finding that before surgery these patients' smiles showed lower joy and higher negative emotion probability compared to controls, while after smile reanimation surgery their expressions showed significantly more joy and less negative emotion. The computed "Emotionality Quotient" correlated well with layperson assessments (30), indicating AI can objectively track improvements in facial expressiveness.

In craniofacial reconstruction, machine learning has been leveraged for imaging diagnostics and surgical planning. Spoer et al. (2022) introduced a 3D analysis for infants with cleft lip, using computer vision to define a midfacial symmetry plane as a standard measure (31). ML algorithms have also been trained to quantify craniofacial deformities: Bhalodia et al. (2020) developed a pilot machine learning model to classify the severity of metopic craniosynostosis from CT scans (32). The model's severity predictions aligned with clinical assessments, suggesting utility in standardizing craniosynostosis evaluation. Deep learning has improved cephalometric planning as well. Nishimoto et al. (2019) achieved automated cephalometric landmark detection on lateral skull radiographs using a deep convolutional neural network, with accuracy comparable to human examiners in locating cranial landmarks, streamlining orthodontic and orthognathic surgical planning (33). Extending this to 3D, Ma et al. (2020) created an automatic 3D landmarking model for craniofacial CT images using a patch-based deep neural network, significantly reducing manual effort in identifying anatomical points (34). These advances indicate that computer vision can assist plastic and craniofacial surgeons in diagnosis and treatment planning by providing objective, reproducible measurements and predictions.

3.3. Breast surgery and reconstruction

Artificial intelligence has been applied extensively in breast reconstruction for both outcome prediction and aesthetic assessment. Multiple retrospective studies trained machine learning models on clinical datasets to predict complications after breast surgery. For instance, Hassan et al. (2023) developed several ML algorithms to predict implant-based reconstruction complications. Using data from 481 patients, their best models achieved AUROC of 0.73 for predicting postoperative implant infection and 0.78 for implant loss. The ML model identified key risk factors and provided patient-specific risk estimates (6). This suggests AI can stratify patients preoperatively by infection risk, aiding in counseling and potentially guiding preventive strategiespubmed.ncbi.nlm.nih.gov. Likewise, Chen et al. (2023) addressed capsular contracture, a common problem in two-stage implant reconstruction. In a cohort of 209 patients, a neural network model was most accurate, correctly predicting capsular contracture with 82% accuracy (AUC ~0.79) (7). The model flagged risk factors such as older age, smaller pre-op breast measurements, submuscular implant placement, use of surgical mesh, and history of radiation, allowing surgeons to identify high-risk cases for closer follow-up or alternative techniques (7).

For autologous breast reconstruction, AI-based predictive analytics have similarly been explored. O'Neill et al. (2020) built a machine learning model to predict free flap failure in microvascular breast reconstruction using pre- and intraoperative variables (35). Their model successfully identified high-risk patients who were more likely to suffer flap thrombosis or failure. This enables targeted preventive measures or enhanced postoperative monitoring in those patients (35). In an even larger series of 568 patients, Myung et al. (2021) validated ML approaches for predicting donor-site complications after abdominal flap (DIEP) breast reconstructions. Their neural network model outperformed other statistical methods, yielding an overall predictive accuracy ~82% for abdominal wound complications The model highlighted that a fascial defect size >37.5 cm^2, patient diabetes, and certain flap techniques significantly increased donor site risk (8). Patients above the risk threshold had a 26% donor complication rate vs only 1.7% in low-risk patientspubmed.ncbi.nlm.nih.gov, demonstrating how AI risk calculators can discriminate those who might benefit from prophylactic mesh or modified closure techniques. Additionally, a multi-institutional group led by Ozmen et al. (2025) harnessed AI for 30-day readmission prediction after DIEP flap surgery. Using NSQIP national data, they constructed an ensemble ML model that reliably identified patients likely to be readmitted, enabling targeted perioperative interventions (36).

3.4. Microsurgery and hand surgery

In reconstructive microsurgery, AI techniques have aimed to improve intraoperative precision and postoperative monitoring. One notable area is robotic-assisted microsurgery. In a landmark randomized trial, van Mulken et al. (2020) reported the first-in-human use of a dedicated microsurgical robot for supermicrosurgery in lymphedema patients (19). In this pilot RCT, 20 breast cancer survivors requiring lymphaticovenular anastomosis (LVA) were randomized to robot-assisted vs manual LVA. The robot group achieved successful anastomoses with comparable 3-month outcomes and improved suturing precision, albeit with longer operative times. This study demonstrated the feasibility of robotic supermicrosurgery in patients and showed a trend toward reduced fatigue and tremor-related errors (19). Following this, several case series have implemented the new Symani Surgical System for microsurgical free flaps. Beier et al. (2023) documented 23 free flap transfers using the Symani robot for microvascular anastomoses. All 23 arterial anastomoses and a few venous anastomoses were completed robotically; while the robotic suturing took longer than manual norms, the success rate was high (only 1 flap loss) (20). Similarly, Strübing et al. (2024) reported a series of 16 patients undergoing robot-assisted free flap reconstruction of the upper extremity. They found the technique feasible and safe, noting that all flaps survived and the robotic approach was especially useful in deep or narrow fields where traditional hand suturing is challenging (21). A smaller study by Watson et al. (2025) in Zurich used the Symani robot for 6 scalp reconstruction cases, concluding that robotic microanastomosis in head and neck reconstruction is safe and yields satisfactory outcomes, with no flap failures and reasonable operative times (22).

Artificial intelligence is also enhancing intraoperative vision and postoperative monitoring. Nakazawa et al. (2020) developed a real-time computer vision system using a region-based convolutional neural network to automatically detect surgical needles in the operative field (37). Such technology can be integrated with robotic platforms to guide suture placement or avoid needle loss, improving safety. In free flap monitoring, traditionally reliant on clinical exam and hand-held Dopplers, AI-based tools are emerging to detect perfusion problems earlier. Danciu et al. (2023) introduced a deep learning system analyzing thermal imaging of flaps to detect ischemia. In a pilot involving postoperative flap patients, their model segmented perfused vs. non-perfused areas on infrared images with 87% accuracy (85% sensitivity, 89% specificity) (23), outperforming prior techniques. This noninvasive monitoring tool could alert staff to compromised flaps earlier than clinical observation. Additionally, predictive models have been developed for microsurgery outcomes: for example, an AI algorithm by Shi et al. (2022) used machine learning to predict which patients might require return to the OR for microvascular revision, allowing proactive management (35). While no studies on hand surgery-specific AI met the inclusion criteria, the advances in microsurgery and nerve repair imply future applications in hand reanimation and transplant surgery.

3.5. Burn care and wound healing

Burn surgery has seen significant AI-driven developments, particularly in burn depth assessment and wound management – areas where accurate early diagnosis is critical. Traditional burn depth estimation by visual exam is error-prone, with up to 25–39% (38). AI algorithms using imaging have shown promise in distinguishing burns requiring grafting from those that will heal spontaneously (38). For instance, Chang et al. (2021) developed a deep learning model for automated burn wound diagnosis. Using a large dataset of burn photographs, their model could segment burn regions and calculate total body surface area (%TBSA) involvement on a pixel-wise basis. This tool achieved high agreement with clinicians for burn size and provided an objective TBSA computation, which is useful for fluid resuscitation planning (24). In terms of depth, Rangaiah et al. (2025) proposed a combined imaging + AI framework for "precision diagnosis" of acute burns. They utilized advanced imaging modalities and an ensemble of deep learning models to estimate burn depth and need for surgery. Notably, their approach, which included an adaptive independent component analysis and a recurrent neural network (RNN), achieved 96.7% accuracy in classifying burn depth as superficial vs. deep (26). Such high accuracy, if confirmed, exceeds most prior methods and hints that AI can outperform even costly devices like laser Doppler imaging (LDI). Lee et al. (2025) conducted a multi-center validation comparing an AI-driven burn assessment tool against standard LDI. Their convolutional neural network with a boundary-attention mechanism (CNN-BAM) correctly classified burn wound depth with ~85% accuracy (4-level classification) and could automatically delineate wound boundaries with 91.6% accuracy compared to expert LDI maps (25). The AI's burn-depth predictions showed substantial correlation with healing outcomes and LDI findings (66% agreement), effectively matching LDI's clinical accuracy. This suggests an accessible AI algorithm on a mobile device could provide point-of-care burn triage comparable to expensive imaging systems (25). Other groups have similarly reported deep CNN models achieving 80–95% accuracy in classifying burn depths (38), often using photographic data augmented by thermal images or clinical data.

In chronic wound care, AI has been applied to predict and identify non-healing wounds. Jung et al. (2016) developed a prognostic model for chronic wound healing based on wound tissue gene expression and clinical feature. Using machine learning, they could predict by week 1 which wounds were likely to be "slow-healing" versus those on track to heal, with the model discriminating healing status with significant accuracy (27). This early identification allows proactive interventions for wounds predicted to stall. Another included study leveraged an AI system to detect risk of surgical site infection in wounds, demonstrating how NLP and predictive analytics can flag at-risk surgical wounds by mining operative reports for warning phrases (35).

3.6. Quality Assessment

The methodological quality of included studies was variable, with several common limitations. All studies clearly stated their aims and the vast majority used appropriate endpoints and statistical analyses. However, only 25% (6/24) enrolled consecutive patients and just 29% (7/24) were conducted prospectively, indicating a predominance of retrospective study designs. While 58% (14/24) of studies were judged to have unbiased outcome assessments, 10 relied on subjective or potentially biased measures. Only 42% (10/24) explicitly reported strategies to mitigate overfitting—such as cross-validation or external validation—which raises concerns about model generalisability. Notably, just 25% (6/24) of studies compared AI performance against standard care or clinician assessment, limiting the ability to contextualise their clinical relevance (Table 2)

Table 2 MINORS critical appraisal of identified literature

Study (Author, Year)	Clearly stated aim	Consecutive patients	Prospective design	Appropriate endpoints	Unbiased outcome assessment	Appropriate statistical analysis	Overfitting mitigation reported	Comparison to standard/clinician
O'Neill et al., 2020	✓	✗	✗	✓	✓	✓	✓	✗
Hassan et al., 2023	✓	✗	✗	✓	✓	✓	✓	✗
Chen et al., 2023	✓	✗	✗	✓	✓	✓	✓	✗
Myung et al., 2021	✓	✗	✗	✓	✓	✓	✓	✗
Ozmen et al., 2025	✓	✓	✗	✓	✓	✓	✓	✗
Dorfman et al., 2020	✓	✗	✗	✓	✓	✓	✓	✗
Chen et al., 2020	✓	✗	✗	✓	✓	✓	✓	✗
Dusseldorp et al., 2019	✓	✗	✗	✓	✓	✓	✓	✗
Wu et al., 2016	✓	✗	✗	✓	✓	✓	✓	✗
Bhalodia et al., 2020	✓	✗	✗	✓	✓	✓	✓	✗
Nishimoto et al., 2019	✓	✗	✗	✓	✓	✓	✓	✗
Ma et al., 2020	✓	✗	✗	✓	✓	✓	✓	✗
Nakazawa et al., 2020	✓	✗	✗	✓	✓	✓	✓	✗

Knoops et al., 2019	✓	✗	✗	✓	✓	✓	✓	✗
van Mulken et al., 2020	✓	✗	✗	✓	✓	✓	✓	✗
Beier et al., 2023	✓	✗	✗	✓	✓	✓	✗	✗
Strübing et al., 2024	✓	✗	✗	✓	✓	✓	✗	✗
Watson et al., 2025	✓	✗	✓	✓	✓	✓	✗	✗
Danciu et al., 2023	✓	✗	✓	✓	✓	✓	✗	✗
Chang et al., 2021	✓	✗	✓	✓	✗	✓	✗	✗
Lee et al., 2025	✓	✓	✓	✓	✗	✓	✗	✓
Rangaiah et al., 2025	✓	✓	✓	✓	✗	✓	✗	✓
Jung et al., 2016	✓	✓	✓	✓	✗	✓	✗	✓
Boczar et al., 2020	✓	✓	✓	✓	✗	✓	✗	✓

4. Discussion

Our systematic review found that AI is being employed across a wide range of plastic and reconstructive surgery domains. From 2015 to 2025, the literature on AI in plastic surgery grew markedly, yielding dozens of studies spanning aesthetic surgery, breast reconstruction, craniofacial surgery, microsurgery, and wound care (2). The most common applications involved computer vision and machine learning for diagnosis or outcome prediction. For example, multiple studies applied deep learning to analyze wound or burn images for automated assessment (39). Others used machine learning models to predict postoperative complications or patient-reported outcomes with encouraging accuracy (40, 41). Emerging AI tools such as augmented reality (AR) have also been explored to assist surgical planning, notably in perforator flap surgery, where AR visualization can enhance preoperative mapping of blood vessels. Likewise, natural language AI (chatbots and large language models) are being tested for patient education and surgical training support (42).

Our findings align with earlier reviews that noted the breadth of AI's potential in plastic surgery alongside its nascent stage. Jarvis et al. (2020) similarly identified numerous emerging AI applications – including machine learning for outcome prediction and facial image analysis – but emphasized that these were early explorations requiring further development (43). Spoer et al. (2022) performed a systematic review up to early 2021 and included 44 studies, reporting that most research was in phase 0–1 (discovery or technical feasibility) with very few reaching clinical efficacy testing (2). Our updated review confirms that even with an influx of studies by 2025 (approximately 70 included), the majority remain at preclinical phases. Notably, Spoer et al. observed only one study with translation to practice, and we found little additional progress beyond that in subsequent years. This underscores a persistent gap between algorithm development and clinical implementation.

Subspecialty-focused reviews mirror our conclusions. For example, a recent review of AI in facial plastic surgery noted that AI could aid in diagnosis and surgical planning but that evidence was limited and fragmented across case studies (42). Similarly, a narrative review by Liang et al. (2021) highlighted various AI tools and even demonstrated a Markov model for keloid treatment, but ultimately pointed out the challenges to applying these models in practice (44). Our results also expand on prior literature by incorporating newer AI modalities. Earlier reviews mostly discussed machine learning and computer vision; our review includes the rise of AR and chatbots in plastic surgery. The 2025 systematic review by Herzog et al. noted AR as an especially promising tool for improving surgical visualization and patient consultation, a finding echoed in our analysis of recent studies. Additionally, our quality appraisal offers a contrast with prior assessments: whereas Spoer et al. (2022) and Nogueira et al. (2025) used standardized risk-of-bias tools and found many studies at moderate to high risk of bias (31, 45), our adapted criteria specifically highlight deficits like lack of prospective validation and limited reporting of overfitting countermeasures in the current literature.

4.1. Strengths and limitations

This systematic review provides a comprehensive and up-to-date synthesis of AI applications in plastic and reconstructive surgery through 2025. A key strength is the broad inclusion of diverse AI modalities, which allowed us to capture the full landscape of AI use in this specialty. We also implemented a rigorous quality appraisal using an adapted MINORS framework, tailored to AI diagnostic studies, which to our knowledge is among the first attempts to quantitatively assess the methodological quality of this body of literature. By not excluding studies based on quality, we were able to identify common limitations across the field. Our analysis revealed that while nearly all studies clearly stated their objectives and used appropriate outcome measures, many had significant methodological shortcomings (see Table 1). For instance, only about a quarter of studies explicitly reported enrolling consecutive patients, and only ~20% were designed prospectively. This indicates potential selection biases and a predominance of retrospective analyses. Furthermore, outcome assessment was not always blinded or independent, with roughly half the studies risked biased assessment by using non-independent ground truth or unblinded evaluators. Another notable limitation was the scant attention to overfitting: only ~40% of studies described measures such as cross-validation or external testing to ensure their AI models would generalize to new data. Additionally, only a small minority of studies directly compared the AI tool to standard care or clinician performance, underscoring that most research has not yet benchmarked AI against the current gold standard. These quality issues limit the confidence and generalizability of reported findings.

Our review itself has limitations. The heterogeneity of included studies, spanning different AI techniques, clinical aims, and outcome metrics, precluded any quantitative meta-analysis, and we relied on narrative synthesis (42). There is also an inherent risk of publication bias; studies reporting positive AI results may be overrepresented in the literature. We attempted to mitigate bias by including all relevant languages and by critically appraising study design rather than only reported accuracies. Nonetheless, the rapid evolution of AI means that conclusions could become outdated as new studies emerge; our search covered up to 2025, and subsequent breakthroughs or validations might not be captured. Finally, while our adapted quality criteria were suited to evaluating AI studies, the scoring was somewhat subjective (e.g. what constitutes "appropriate" statistics or "unbiased" assessment), and other reviews have used formal risk-of-bias tools that could yield different evaluations (45). Despite these limitations, our work provides a necessary assessment of both the promise and current evidence gaps of AI in plastic surgery.

4.2. Future directions

To facilitate the transition from research to clinical use, several explicit strategies must be addressed. First, regulatory approval remains a major hurdle. Most AI systems must demonstrate clinical safety and efficacy through prospective clinical trials or equivalent regulatory pathways, which few plastic surgery AI tools have achieved. Second, ethical considerations are critical, especially in aesthetic contexts, where algorithmic bias or overreach into patient decision-making must be avoided. Third, ensuring robust data privacy, particularly with identifiable data like facial images, requires adherence to strict de-identification protocols and institutional governance frameworks. Fourth, clinician acceptance hinges on transparency: AI must be explainable and seamlessly integrate into clinical workflows to foster trust and utility. Finally, real-world integration depends on practical considerations such as software interoperability with EMRs and PACS, speed of inference, and minimal workflow disruption.

Recent research reinforces these priorities, highlighting the importance of prospective validation and real-world testing to ensure that AI systems developed in research settings remain reliable when deployed in clinical environments (46). Explainable AI techniques are gaining traction as essential tools to identify and manage data drift, thus improving model transparency and clinician trust. Stakeholder engagement throughout the development process, particularly involving clinicians early and consistently, has also been recognised as critical to successful adoption. Studies further stress the value of integrating AI into existing clinical workflows through iterative design, user-centred interfaces, and seamless interoperability with electronic medical records. Frameworks such as the FUTURE-AI guideline offer structured, end-to-end recommendations covering everything from development and validation to deployment and monitoring, helping to facilitate the safe, effective, and ethical implementation of AI tools in healthcare (47).

5. Conclusion

AI is rapidly emerging as a valuable tool in plastic and reconstructive surgery, with applications spanning diagnostics, surgical planning, outcome measurement, and workflow optimization. This review of clinical studies from 2015–2025 highlights that AI models often achieve high accuracy, sometimes matching expert performance, and have been explored across all subspecialties. However, most AI solutions remain in early development or validation, with limited adoption in routine clinical practice due to challenges such as insufficient data, lack of robust validation, and cautious clinical uptake. Realizing AI's full potential will require collaboration among surgeons, data scientists, and industry to improve data quality, algorithm transparency, and generalizability, as well as to establish ethical guidelines. If these hurdles are addressed, AI could soon become a routine part of clinical care, enhancing decision-making, surgical precision, and patient outcomes, and marking a transformative shift in the field's future.

References

[1] Choi E, Leonard KW, Jassal JS, Levin AM, Ramachandra V, Jones LR. Artificial Intelligence in Facial Plastic Surgery: A Review of Current Applications, Future Applications, and Ethical Considerations. Facial Plast Surg. 2023;39(5):454–9.

[2] Spoer DL, Kiene JM, Dekker PK, Huffman SS, Kim KG, Abadeer AI, et al. A Systematic Review of Artificial Intelligence Applications in Plastic Surgery: Looking to the Future. Plast Reconstr Surg Glob Open. 2022;10(12):e4608.

[3] Hijazi Y, Kiwan O, Rafie A, Al-Kalbani M, Myatt A. Artificial intelligence in plastic surgery, where do we stand? JPRAS Open. 2024;42.

[4] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Bmj. 2021;372:n71.

[5] O'Neill AC, Yang D, Roy M, Sebastiampillai S, Hofer SOP, Xu W. Correction to: Development and Evaluation of a Machine Learning Prediction Model for Flap Failure in Microvascular Breast Reconstruction. Annals of Surgical Oncology. 2022;29(6):3867–.

[6] Hassan AM, Biaggi-Ondina A, Asaad M, Morris N, Liu J, Selber JC, et al. Artificial Intelligence Modeling to Predict Periprosthetic Infection and Explantation following Implant-Based Reconstruction. Plast Reconstr Surg. 2023;152(5):929–38.

[7] Chen Y, Wang ML, Black GG, Qin N, Zhou G, Bernstein JL, et al. Machine-Learning Prediction of Capsular Contraction after Two-Stage Breast Reconstruction. JPRAS Open. 2023;38:1–13.

[8] Myung Y, Jeon S, Heo C, Kim EK, Kang E, Shin HC, et al. Validating machine learning approaches for prediction of donor related complication in microsurgical breast reconstruction: a retrospective cohort study. Sci Rep. 2021;11(1):5615.

[9] Kim DK, Corpuz GS, Ta CN, Weng C, Rohde CH. Applying unsupervised machine learning approaches to characterize autologous breast reconstruction patient subgroups: an NSQIP analysis of 14,274 patients. J Plast Reconstr Aesthet Surg. 2024;88:330–9.

[10] Dorfman R, Chang I, Saadat S, Roostaeian J. Making the Subjective Objective: Machine Learning and Rhinoplasty. Aesthet Surg J. 2020;40(5):493–8.

[11] Chen K, Lu SM, Cheng R, Fisher M, Zhang BH, Di Maggio M, et al. Facial Recognition Neural Networks Confirm Success of Facial Feminization Surgery. Plast Reconstr Surg. 2020;145(1):203–9.

[12] Dusseldorp JR, Guarin DL, van Veen MM, Miller M, Jowett N, Hadlock TA. Automated Spontaneity Assessment after Smile Reanimation: A Machine Learning Approach. Plast Reconstr Surg. 2022;149(6):1393–402.

[13] Wu J, Heike C, Birgfeld C, Evans K, Maga M, Morrison C, et al. Measuring Symmetry in Children With Unrepaired Cleft Lip: Defining a Standard for the Three-Dimensional Midfacial Reference Plane. Cleft Palate Craniofac J. 2016;53(6):695–704.

[14] Bhalodia R, Dvoracek LA, Ayyash AM, Kavan L, Whitaker R, Goldstein JA. Quantifying the Severity of Metopic Craniosynostosis: A Pilot Study Application of Machine Learning in Craniofacial Surgery. J Craniofac Surg. 2020;31(3):697–701.

[15] Nishimoto S, Saito T, Ishise H, Fujiwara T, Kawai K, Kakibuchi M. Three-Dimensional Craniofacial Landmark Detection in Series of CT Slices Using Multi-Phased Regression Networks. Diagnostics [Internet]. 2023; 13(11).

[16] Ma Q, Kobayashi E, Fan B, Nakagawa K, Sakuma I, Masamune K, et al. Automatic 3D landmarking model using patch-based deep neural networks for CT image of oral and maxillofacial surgery. The International Journal of Medical Robotics and Computer Assisted Surgery. 2020;16(3):e2093.

[17] Nakazawa A, Harada K, Mitsuishi M, Jannin P. Real-time surgical needle detection using region-based convolutional neural networks. International Journal of Computer Assisted Radiology and Surgery. 2019;15.

[18] Knoops PGM, Papaioannou A, Borghi A, Breakey RWF, Wilson AT, Jeelani O, et al. A machine learning framework for automated diagnosis and computer-assisted planning in plastic and reconstructive surgery. Scientific Reports. 2019;9(1):13597.

[19] van Mulken TJM, Schols RM, Scharmga AMJ, Winkens B, Cau R, Schoenmakers FBF, et al. First-in-human robotic supermicrosurgery using a dedicated microsurgical robot for treating breast cancer-related lymphedema: a randomized pilot trial. Nature Communications. 2020;11(1):757.

[20] Beier JP, Hackenberg S, Boos AM, Modabber A, Duong Dinh TA, Hölzle F. First Series of Free Flap Reconstruction Using a Dedicated Robotic System in a Multidisciplinary Microsurgical Center. Plast Reconstr Surg Glob Open. 2023;11(9):e5240.

[21] Struebing F, Bigdeli AK, Boecker A, Weigel J, Kneser U, Gazyakan E. Hands-On Robotic Microsurgery: Robotic-Assisted Free Flap Reconstruction of the Upper Extremity. J Clin Med. 2024;13(23).

[22] Watson JA, Könneker S, Esposito G, Hofmann L, Kim B-S, Giovanoli P, et al. Robotic assisted free flap reconstruction of the scalp using the Symani® surgical system. Acta Neurochirurgica. 2025;167(1):93.

[23] Danciu R, Danciu BA, Vasiu L-S, Avino A, Filip CI, Hariga C-S, et al. Deep Learning-Based Flap Detection System Using Thermographic Images in Plastic Surgery. Applied System Innovation [Internet]. 2024; 7(6).

[24] Chang CW, Lai F, Christian M, Chen YC, Hsu C, Chen YS, et al. Deep Learning-Assisted Burn Wound Diagnosis: Diagnostic Model Development Study. JMIR Med Inform. 2021;9(12):e22798.

[25] Lee JJ, Abdolahnejad M, Morzycki A, Freeman T, Chan H, Hong C, et al. Comparing Artificial Intelligence Guided Image Assessment to Current Methods of Burn Assessment. J Burn Care Res. 2025;46(1):6–13.

[26] Rangaiah PKB, kumar BPP, Huss F, Augustine R. Precision diagnosis of burn injuries using imaging and predictive modeling for clinical applications. Scientific Reports. 2025;15(1):7604.

[27] Jung K, Covington S, Sen CK, Januszyk M, Kirsner RS, Gurtner GC, et al. Rapid identification of slow healing wounds. Wound Repair Regen. 2016;24(1):181–8.

[28] Boczar D, Sisti A, Oliver JD, Helmi H, Restrepo DJ, Huayllani MT, et al. Artificial Intelligent Virtual Assistant for Plastic Surgery Patient's Frequently Asked Questions: A Pilot Study. Ann Plast Surg. 2020;84(4):e16–e21.

[29] Yalçın MZ, Toplu Y, Kurt O. Can artificial ıntelligence detect the anti-aging effect of rhinoplasty? J Plast Surg Hand Surg. 2025;60:84–90.

[30] Dusseldorp JR, Guarin DL, van Veen MM, Jowett N, Hadlock TA. In the Eye of the Beholder: Changes in Perceived Emotion Expression after Smile Reanimation. Plast Reconstr Surg. 2019;144(2):457–71.

[31] Spoer DL, Kiene JM, Dekker PK, Huffman SS, Kim KG, Abadeer AI, et al. A Systematic Review of Artificial Intelligence Applications in Plastic Surgery: Looking to the Future. Plastic and Reconstructive Surgery – Global Open. 2022;10(12).

[32] Bhalodia R, Dvoracek LA, Ayyash AM, Kavan L, Whitaker R, Goldstein JA. Quantifying the severity of metopic craniosynostosis: a pilot study application of machine learning in craniofacial surgery. Journal of Craniofacial Surgery. 2020;31(3):697–701.

[33] Nishimoto S, Sotsuka Y, Kawai K, Ishise H, Kakibuchi M. Personal computer-based cephalometric landmark detection with deep learning, using cephalograms on the internet. Journal of Craniofacial Surgery. 2019;30(1):91–5.

[34] Ma Q, Kobayashi E, Fan B, Nakagawa K, Sakuma I, Masamune K, et al. Automatic 3D landmarking model using patch-based deep neural networks for CT image of oral and maxillofacial surgery. The International Journal of Medical Robotics and Computer Assisted Surgery. 2020;16(3):e2093.

[35] Yizhuo S, Andrew JM, Renee G, Justin MB. Artificial intelligence in breast reconstruction. Artificial Intelligence Surgery. 2025;5(1):150–9.

[36] Ozmen BB, Phuyal D, Schwarz GS. SP02. Predicting Reoperation in DIEP Flap Surgery With Machine Learning: A Random Forest Approach. Plast Reconstr Surg Glob Open. 2025;13(Suppl 1).

[37] Nakazawa A, Harada K, Mitsuishi M, Jannin P. Real-time surgical needle detection using region-based convolutional neural networks. International Journal of Computer Assisted Radiology and Surgery. 2020;15:41–7.

[38] Boissin C, Laflamme L, Fransén J, Lundin M, Huss F, Wallis L, et al. Development and evaluation of deep learning algorithms for assessment of acute burns and the need for surgery. Scientific Reports. 2023;13(1):1794.

[39] Chang CW, Ho CY, Lai F, Christian M, Huang SC, Chang DH, et al. Application of multiple deep learning models for automatic burn wound assessment. Burns. 2023;49(5):1039–51.

[40] Chen J, Gabay A, Boe LA, Shammas RL, Stern C, Pusic A, et al. Machine Learning Accurately Predicts Patient-Reported Outcomes 1 Year After Breast Reconstruction. Annals of Surgery. 9900.

[41] Mantelakis A, Khajuria A. The applications of machine learning in plastic and reconstructive surgery: protocol of a systematic review. Systematic Reviews. 2020;9(1):44.

[42] Herzog C, Jardim G, Decol A, Palma V, Alves G, Negri R, et al. Limitations and Challenges in the Incorporation of Artificial Intelligence in Plastic Surgery, A. Systematic Review. Brazilian Journal of Plastic Surgery. 2025;40.

[43] Jarvis T, Thornburg D, Rebecca AM, Teven CM. Artificial Intelligence in Plastic Surgery: Current Applications, Future Directions, and Ethical Implications. Plast Reconstr Surg Glob Open. 2020;8(10):e3200.

[44] Liang X, Yang X, Yin S, Malay S, Chung KC, Ma J, et al. Artificial Intelligence in Plastic Surgery: Applications and Challenges. Aesthetic Plast Surg. 2021;45(2):784–90.

[45] Nogueira R, Eguchi M, Kasmirski J, de Lima BV, Dimatos DC, Lima DL, et al. Machine Learning, Deep Learning, Artificial Intelligence and Aesthetic Plastic Surgery: A Qualitative Systematic Review. Aesthetic Plast Surg. 2025;49(1):389–99.

[46] Choi Y, Yu W, Nagarajan MB, Teng P, Goldin JG, Raman SS, et al. Translating AI to Clinical Practice: Overcoming Data Shift with Explainability. RadioGraphics. 2023;43(5):e220105.

[47] Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. Bmj.;388:e081554.