E-ISSN 2250-0944
ISSN 2250-1150
doi: 10.48047/ijprt/16.02.165
Research Article

Artificial Intelligence in Surgical Decision-Making: A Comprehensive Review of Applications Across the Preoperative, Intraoperative, and Postoperative Continuum

Dr. N. Junior Sundresh MS, FRCS, FACS[1], Dr. Jayaraman MS[2], Dr. Shukla Bikku[3]

[1]Professor, Department of General Surgery, Government Medical College and Hospital, Cuddalore, Tamil Nadu, India
[2]Assistant Professor, Department of General Surgery, Government Medical College and Hospital, Cuddalore, Tamil Nadu, India
[3]Post-graduate, Department of General Surgery, Government Medical College and Hospital, Cuddalore, Tamil Nadu, India
Corresponding Author: Dr. Shukla Bikku
Post-graduate, Department of General Surgery, Government Medical College and Hospital, Cuddalore, Tamil Nadu, India.
Received: 06.05.2026, Revised: 04.06.2026, Accepted: 01.07.2026

ABSTRACT

Background: Artificial intelligence (AI) and machine learning (ML) are increasingly embedded across the surgical care pathway, moving beyond early risk-scoring tools toward real-time intraoperative guidance, automated complication prediction, and clinical decision support for multidisciplinary oncologic planning. Whether this expanding footprint translates into genuinely improved surgical decision-making-as opposed to impressive but clinically unvalidated model performance-remains an open and actively debated question.

Objective: To synthesize current evidence on AI applications supporting surgical decision-making across the preoperative, intraoperative, and postoperative phases of care, and to critically appraise the regulatory, ethical, and trust-related barriers to clinical adoption.

Methods: A structured narrative review of the literature indexed in PubMed, Embase, and Scopus was performed, prioritizing systematic reviews, prognostic model development and validation studies, and regulatory analyses published between 2019 and 2026. Domains reviewed included preoperative risk prediction, intraoperative computer vision and guidance, postoperative complication prediction (with anastomotic leak as an exemplar), AI-based clinical decision support in surgical oncology multidisciplinary teams (MDTs), the U.S. FDA regulatory pathway for AI/ML-enabled devices, and the ethical and explainability challenges specific to surgical AI.

Results: Preoperative ML models for mortality and major adverse cardiac and cerebrovascular events now achieve area-under-the-curve (AUC) values of 0.90–0.97 using only routinely available electronic health record data, consistently outperforming legacy risk calculators such as the ACS-NSQIP Surgical Risk Calculator and the ASA physical status classification. Intraoperative computer-vision systems have demonstrated feasibility for real-time surgical phase recognition, instrument tracking, and critical-structure identification, with precision frequently exceeding 90% in experimental settings, though a 2026 scoping review found only 12% of eligible studies had evaluated real-time intraoperative integration. Deep learning models applied to intraoperative video and preoperative clinical data have shown promise for predicting anastomotic leak, a leading source of morbidity in colorectal surgery, though most studies remain small, single-technique proofs of concept. AI-based clinical decision support tools are being piloted within surgical oncology MDTs but show highly variable performance across tumor types and institutions. Regulatory scrutiny has intensified: as of recent analyses, roughly 96–97% of FDA-cleared AI/ML medical devices reached market via the lower-evidentiary-bar 510(k) pathway, and fewer than 10% of orthopaedic AI devices have undergone formal prospective clinical trial validation. Persistent barriers to adoption include algorithmic "black box" opacity, documented tension between model interpretability and predictive performance, inconsistent evidence that explainability actually improves clinician trust or decision quality, and unresolved questions of legal liability when AI-guided decisions contribute to patient harm.

Conclusion: AI-based tools already outperform traditional risk-scoring systems for preoperative risk stratification and show credible early promise for intraoperative guidance and postoperative complication prediction, but the evidence base remains dominated by retrospective, single-center, internally validated studies with limited prospective or multicenter confirmation. Wider clinical adoption should proceed in parallel with more rigorous prospective validation, standardized performance and bias reporting, clearer regulatory requirements for adaptive AI/ML devices, and explicit frameworks preserving surgeon authority and accountability over AI-assisted decisions.

Keywords: Artificial intelligence, machine learning, surgical decision-making, computer vision, clinical decision support, risk prediction, explainable AI, algorithmic bias, regulatory science.

1. INTRODUCTION

Surgical decision-making has always depended on the synthesis of large volumes of heterogeneous information-patient comorbidity, imaging, laboratory data, intraoperative findings, and accumulated surgeon experience-into a judgment made under time pressure and uncertainty. Artificial intelligence (AI) and machine learning (ML) are increasingly positioned as tools to augment this process, promising to convert the ever-growing volume of structured and unstructured perioperative data into actionable, individualized guidance. Contemporary reviews describe AI's expanding role as extending well beyond early risk-prediction calculators toward real-time intraoperative decision support, automated documentation, and even components of robotic task automation.[1] As these systems progress from experimental prototypes toward integration into routine practice, understanding both their demonstrated capabilities and their well-documented limitations-including overfitting, dataset bias, and the interpretive opacity of complex models-is essential for their safe and ethically sound adoption.[1,2]

This review synthesizes the current evidence on AI-supported surgical decision-making across three phases of surgical care: preoperative risk stratification and planning, intraoperative guidance, and postoperative complication prediction, using anastomotic leak prediction as a detailed exemplar of the latter. It further considers AI's emerging role in surgical oncology multidisciplinary team (MDT) decision-making, the regulatory pathway through which these tools reach clinical practice, and the ethical, legal, and trust-related barriers-chief among them algorithmic bias and the "black box" problem-that continue to constrain confident clinical adoption.[2,4,5,6]

2. METHODS

This is a structured narrative review rather than a formal systematic review with meta-analysis. A literature search was conducted across PubMed/MEDLINE, Embase, and Scopus for studies published between January 2019 and February 2026, using combinations of the terms "artificial intelligence", "machine learning", "deep learning", "computer vision", "surgery", "surgical decision-making", "clinical decision support", "risk prediction" and procedure-or complication-specific terms (e.g., "anastomotic leak", "surgical oncology", "multidisciplinary team"). Priority was given to systematic and scoping reviews, prognostic model development and external validation studies reporting discrimination metrics (area under the receiver operating characteristic curve [AUC/AUROC]), and regulatory landscape analyses of FDA-cleared AI/ML-enabled devices. Case reports, non-peer-reviewed preprints without empirical validation, and studies without a clearly defined comparator or outcome were deprioritized, though several recent, methodologically transparent preprints were retained where they represented the most current evidence available in fast-moving subdomains (e.g., real-time intraoperative computer vision).

3. RESULTS

3.1 PREOPERATIVE RISK STRATIFICATION AND SURGICAL PLANNING

Preoperative risk prediction represents the most mature domain of AI application in surgery, with a substantial body of evidence now directly comparing ML models against established clinical risk calculators. In a large prognostic study of 1,477,561 surgical patients, an automated ML model achieved an AUROC of 0.972 for predicting postoperative mortality and 0.923 for major adverse cardiac and cerebrovascular events or mortality, outperforming the American College of

Surgeons NSQIP Surgical Risk Calculator by 0.048 AUROC points.[7] A separate model trained on 53,097 surgical patients and using only structured, automatically extractable preoperative electronic health record (EHR) data achieved an AUC of 0.936 for in-hospital mortality, outperforming the Preoperative Score to Predict Postoperative Mortality (POSPOM), the Charlson comorbidity score, and the ASA physical status classification-the latter of which requires manual clinician chart review and is therefore not fully automatable.[8] Notably, the non-linear ML models in this study outperformed logistic regression on both discrimination and calibration, suggesting genuine methodological advantage rather than mere data-volume effects.[8]

These findings have been replicated across increasingly specific clinical targets. A CatBoost-based model developed using data from 36,502 patients and externally validated on 404 patients achieved an AUC of 0.867 for perioperative stroke prediction, substantially outperforming the revised cardiac risk index (AUC 0.528).[9] Similar model-development efforts have targeted postoperative acute kidney injury (deep neural network and LightGBM models, AUC 0.83-0.84, trained on 239,267 surgeries),[12] perioperative venous thromboembolism in lung cancer surgery (XGBoost model incorporating six routinely available predictors),[11] and postoperative malnutrition risk in esophageal cancer patients, where a random forest model (AUC 0.82) performed comparably to a simplified, clinically interpretable nomogram (AUC 0.80)-illustrating that in some contexts, transparent statistical tools can approach ML performance while preserving interpretability.[13] A broader review of AI and ML in surgical complication prediction found that ensemble methods such as random forest and extreme gradient boosting consistently outperformed single-model approaches for outcomes ranging from cardiopulmonary complications after lung resection (AUC 0.75) to severe complications after cardiac surgery (AUC 0.72), generally exceeding the discrimination of traditional logistic regression-based scores.[14]

3.2 INTRAOPERATIVE DECISION SUPPORT: COMPUTER VISION AND REAL-TIME GUIDANCE

Intraoperative AI applications center predominantly on computer vision (CV) systems trained to interpret live or recorded surgical video. A 2026 scoping review following PRISMA-ScR and Joanna Briggs Institute methodology screened 490 articles and included 113 studies of CV-AI systems in general surgery, finding that anatomical delineation and surgical phase recognition represented the most mature application areas, but that only 13 studies (12%) had evaluated real-time intraoperative integration, with the remainder confined to retrospective feasibility analyses; the review concluded that it remains unclear whether CV-AI meaningfully improves surgeon performance, intraoperative judgment, or operative outcomes.[16] A related systematic review of AI-based intraoperative guidance specifically within robotic-assisted ventral cavity surgery screened 2,601 records and included 95 studies, categorizing guidance methods into phase recognition, semantic segmentation, object/action detection, and three-dimensional model overlay or registration, but identified limited clinical validation, inconsistent methodology, and evaluation metrics of unclear clinical relevance as key shortcomings limiting translation to practice.[20]

Specialty-specific proof-of-concept systems illustrate both the promise and the current ceiling of this technology. In ophthalmic surgery, a region-based convolutional neural network was able to track pupil position and identify surgical phase during phacoemulsification cataract surgery with a mean AUROC greater than 95%, triggering real-time computer-vision-based guidance tools integrated into the operating microscope.[17] A separate systematic review of AI in vitreoretinal surgery found intraoperative CV systems achieving precision above 90% for real-time instrument detection and tracking in experimental and early clinical settings, alongside deep learning models exceeding 90% accuracy for predicting anatomical outcomes such as macular hole closure.[15] In hepato-pancreato-biliary surgery, computer-vision systems have been used to recognize critical intraoperative landmarks such as the critical view of safety during cholecystectomy and to detect hazards including bleeding in near real time, though a systematic review of this literature found most studies to be retrospective, single-center feasibility analyses lacking external validation.[21] Complementary work in orthopedic surgery describes AI-driven imaging analysis, anatomical segmentation, and implant selection extending from preoperative planning into intraoperative image processing and adaptive feedback integrated with robotic platforms, with applications spanning joint replacement, spine, and trauma surgery.[18]

Urologic commentary on this emerging field notes that AI-powered computer vision may additionally reduce operating-room workload through automated endoscope tracking and stabilization, though the authors characterize the field as still in its infancy.[19]

3.3 PREDICTING AND PREVENTING POSTOPERATIVE COMPLICATIONS: ANASTOMOTIC LEAK AS AN EXEMPLAR

Anastomotic leak (AL) after colorectal and rectal cancer surgery is among the most extensively studied targets for AI-based postoperative complication prediction, both because of its severe clinical consequences and because no reliable real-time intraoperative assessment tool has existed in conventional practice. A 2025 proof-of-concept study developed a convolutional neural network trained on 5,356 annotated video frames from 26 patients across three international high-volume centers, testing four CNN architectures (EfficientNetB0, EfficientNetB7, ResNet50, and MobileNetV2) to classify intraoperative laparoscopic images of the anastomosis as at-risk or normal; the authors concluded that such models could enable real-time, intraoperative risk stratification, representing a potential paradigm shift from the current reliance on postoperative clinical deterioration to trigger AL diagnosis.[22,23]

A parallel narrative review of ML and deep learning approaches to AL prevention after rectal cancer surgery concluded that AI-based models have consistently demonstrated superior predictive power compared with traditional statistical methods, particularly when incorporating preoperative risk factors such as malnutrition, body composition (notably visceral fat), and radiological features, though the review emphasized that rigorous validation through multicenter randomized trials remains necessary before such tools can be considered reliable and generalizable.[24]

Complementary approaches have explored preoperative-only models: a Swiss pilot study developed an ML-based analytic model using preoperative clinical variables to predict AL risk in colorectal surgery,[25] while more recent exploratory work has examined preoperative CT-based mapping of bowel blood supply combined with deep embedding and topological data analysis methods to generate calibrated, interpretable leak-risk scores integrated with a case-based retrieval decision-support tool.[26]

Even emerging quantum machine learning approaches have been benchmarked against classical models (logistic regression, multilayer perceptrons, boosting algorithms) for AL prediction using a 200-patient clinical dataset, reflecting the breadth of methodological experimentation in this space, though such approaches remain far from clinical translation.[27] Across this literature, a consistent limitation is sample size: most AL-prediction studies, including the most methodologically ambitious multicenter video-based model, are trained on cohorts of well under 100 patients, constraining generalizability pending larger, prospectively validated datasets.[22,25]

3.4 AI-BASED CLINICAL DECISION SUPPORT IN SURGICAL ONCOLOGY MULTIDISCIPLINARY TEAMS

Beyond individual-patient risk prediction, AI-based clinical decision support systems (CDSSs) are increasingly piloted to support surgical oncology multidisciplinary team (MDT) meetings, where they aim to recommend treatment strategies, prioritize complex cases, and improve adherence to evidence-based guidelines amid rising caseload and complexity.[4]

A 2026 systematic review following PRISMA guidelines, searching Cochrane, Ovid MEDLINE, and Embase, mapped the current landscape of AI-based CDSSs in surgical oncology MDTs and found that performance varied substantially across tumor types and healthcare systems; in one retrospective analysis of 138 soft tissue sarcoma cases discussed at MDT meetings, a "Watson for Oncology"-type system demonstrated highly variable concordance with MDT decisions, illustrating the difficulty of generalizing oncology decision-support tools across heterogeneous, rare, or complex tumor presentations.[4] A related systematic review focused specifically on the ethical, legal, and informed-consent dimensions of AI in oncologic therapeutic decision-making, registered prospectively on PROSPERO and following Cochrane Handbook methodology, examined the literature from January 2015 to May 2025 and highlighted persistent gaps in transparency and explainability as key barriers to ethical, patient-centered implementation of such tools.[6]

3.5 REGULATORY LANDSCAPE AND CLINICAL VALIDATION

The regulatory pathway through which AI/ML-enabled surgical and surgery-adjacent devices reach clinical practice has direct implications for the strength of evidence underpinning their use. Across multiple device categories, analyses consistently find that the substantial majority of FDA-cleared AI/ML-enabled medical devices reach market via the 510(k) pathway, which requires demonstration of substantial equivalence to a predicate device rather than de novo clinical trial evidence: one analysis of AI/ML-enabled cardiovascular devices found 96.0% cleared via 510(k), with only 9.0% of device summaries including prospective study results and just 5.7% reporting clinical trial data.[28] An analysis of FDA-cleared AI algorithms for medical imaging similarly found 98.0% cleared via 510(k) rather than the more rigorous de novo or premarket approval pathways.[29] Within surgical subspecialties specifically, a retrospective analysis of 70 FDA-cleared AI/ML-based orthopedic surgery devices found that while the pace of clearance accelerated sharply (from 3.0 devices/year in 2017-2019 to 16.6 devices/year in 2022-2024), 22.8% of cleared devices lacked any clinical testing altogether, 68.6% were validated only against retrospective datasets, and just 8.6% had undergone a formal prospective clinical trial.[30] In recognition of the unique challenges posed by adaptive, continuously updating AI/ML software, the FDA has introduced the Predetermined Change Control Plan (PCCP) framework, allowing manufacturers to pre-specify the scope of permissible future model modifications without requiring a new marketing submission for each update-a regulatory innovation whose real-world implementation is itself the subject of ongoing analysis.[31,32] Collectively, this regulatory picture suggests that formal market clearance of a surgical AI tool should not, by itself, be interpreted as equivalent to robust prospective clinical validation.

3.6 EXPLAINABILITY, ALGORITHMIC BIAS, AND CLINICIAN TRUST

A persistent barrier to confident clinical adoption of AI in surgical decision-making is the tension between model performance and interpretability: complex deep learning models frequently achieve superior discrimination but function as "black boxes" whose internal reasoning is not transparent to the clinicians who must act on their output.[1,33] A systematic review of explainable AI (XAI) in healthcare found that the relationship between explainability and clinician trust is not straightforward-providing explanations for AI outputs did not uniformly increase appropriate trust, and in some studies risked either fostering unwarranted distrust of accurate models or, conversely, encouraging automation bias in which clinicians deferred excessively to AI recommendations even when those recommendations were incorrect.[34] Work specific to surgical applications distinguishes model interpretability (the inner logic mapping inputs to outputs) from explainability (post hoc techniques such as SHAP or LIME that surface which patient-specific variables most influenced a given prediction), and notes that a model can be fully explainable while still being poorly validated, biased, or clinically untrustworthy-explainability and trustworthiness are related but distinct properties.[33] A separate ethical and legal analysis argued that clinician trust in AI ought to be grounded primarily in transparency about model training and validation quality, rather than in algorithmic explainability per se, since black-box and explainable models are equally susceptible to bias, unfairness, and data-driven inequity if the underlying training data or development process is flawed.[35,36]

Algorithmic bias represents a related and clinically consequential concern: because most AI models in surgery are trained on retrospective, often single-institution datasets, they risk underperforming for patient populations underrepresented in training data-a concern raised explicitly in reviews of AI ethics in thoracic and general surgery, which caution that algorithmic bias may systematically underestimate the needs of marginalized populations, and that unresolved questions of legal liability arise when AI-guided decisions contribute to patient harm.[2,4]

Table No. 1: Representative AI Applications Across the Surgical Care Continuum

Phase of Care	Representative AI Application	Typical Model Type	Reported Performance	Maturity
Preoperative	30-day mortality/MACCE	Automated	AUROC 0.90-0.97	Validated, multi-

	prediction from EHR data	ML/random forest		site
Preoperative	Perioperative stroke risk prediction	Gradient boosting (CatBoost)	AUC 0.87 (external validation)	Externally validated
Preoperative	Postoperative AKI risk prediction	Deep neural network/LightGBM	AUC 0.83-0.84	Single-center, web-deployed
Intraoperative	Surgical phase/instrument recognition	CNN-based computer vision	Precision >90% (experimental)	Early clinical, limited real-time use
Intraoperative	Real-time anastomotic leak risk from video	CNN (EfficientNet, ResNet)	Proof-of-concept, multicenter	Early feasibility
Intraoperative	Critical-structure/hazard detection (e.g., biliary anatomy)	Deep learning segmentation	Feasibility demonstrated	Pilot/feasibility
Postoperative	Anastomotic leak prediction (preoperative/clinical data)	ML ensembles, neural networks	Superior to logistic regression in most studies	Multiple pilot validations
Oncologic MDT	Treatment-recommendation clinical decision support	Guideline-based/ML hybrid	Highly variable by tumor type	Heterogeneous, early adoption

Note: Performance figures are drawn from the individual cited studies [7-14,16,17,22-25] and reflect discrimination in the original development or validation cohort; cross-study comparison should be made cautiously given differing populations, outcome definitions, and validation methodology.

4. DISCUSSION

The evidence reviewed here supports a consistent and clinically important pattern: AI-based models now reliably outperform legacy, clinician-scored risk tools for preoperative risk stratification, with the magnitude of improvement (often 0.05–0.3 AUC points over comparators such as ASA classification, POSPOM, or the revised cardiac risk index) large enough to be plausibly clinically meaningful rather than a statistical artefact.[7-9] This preoperative domain also benefits from the largest and most methodologically rigorous evidence base, likely because structured EHR data are comparatively abundant, standardized, and amenable to large-cohort, multi-site model development.

By contrast, intraoperative and postoperative applications remain considerably earlier in their translational trajectory. The finding that only 12% of eligible computer-vision studies in a recent scoping review had evaluated real-time intraoperative integration,[16] and that anastomotic leak prediction models-arguably the most clinically anticipated postoperative AI application in general surgery-remain confined to small, single-technique proof-of-concept cohorts,[22,25] suggests that the gap between technical feasibility and clinical readiness is substantially wider in these domains than the preoperative literature might suggest. This pattern is consistent with a broader pattern seen throughout medical AI: development and internal validation proceed rapidly, but prospective, multicenter, outcome-linked clinical trials-the evidentiary standard against which surgical interventions are conventionally judged-lag well behind.[30]

The regulatory findings reinforce this caution. That the overwhelming majority of FDA-cleared AI/ML devices reach market through a pathway requiring only substantial equivalence to a predicate device, with a minority reporting prospective or clinical trial data,[28-30] means that regulatory clearance is at present a weak proxy for genuine clinical validation. This has direct implications for surgeons and institutions evaluating whether to adopt a given AI tool: FDA clearance confirms a device meets a baseline safety and performance bar relative to existing technology, but it does not, by itself, establish that the tool improves patient-centered surgical outcomes, and clinicians should seek out the underlying validation literature rather than treating clearance as sufficient reassurance.

The explainability literature further complicates a simple narrative in which more transparent AI is unambiguously better. Evidence that explanations do not consistently improve appropriate clinician trust-and may in some contexts either increase unwarranted skepticism or, more concerningly, encourage automation bias-suggests that the surgical AI community's substantial investment in explainability techniques (e.g., SHAP, LIME, GradCAM) should be paired with equal investment in studying how surgeons actually use and respond to these explanations in practice, rather than assuming that explainability is intrinsically protective.[33,34]

The alternative position-that clinician trust should rest on rigorous, transparent validation of model development and performance rather than on post hoc explainability of individual predictions-offers a pragmatic path forward that does not require solving the (arguably unsolved) technical problem of making deep neural networks fully interpretable.[35]

Finally, questions of accountability remain largely unresolved in the literature reviewed. As AI-based tools move from passive risk-scoring toward active intraoperative guidance and treatment recommendation, the locus of responsibility when an AI-informed decision contributes to patient harm-the treating surgeon, the institution, or the device manufacturer-is not settled either legally or ethically, and this ambiguity is repeatedly flagged as a barrier to confident adoption even by authors otherwise enthusiastic about AI's clinical potential.[2,6]

5. LIMITATIONS

This review has several limitations. First, as a narrative rather than systematic review, it does not apply formal risk-of-bias tools to each primary study, and inherits any bias or overstatement present in the underlying literature, particularly given that AI performance studies are known to be susceptible to overfitting and reporting bias favoring positive results. Second, the field is evolving extremely rapidly-several sources cited here are 2025-2026 publications or preprints that have not yet undergone extended post-publication scrutiny, and specific performance figures (AUCs, precision values) should be regarded as illustrative of current capability rather than fixed benchmarks. Third, most primary studies reviewed were retrospective, single-center, and internally validated, meaning the true real-world, cross-institutional performance of most tools discussed remains unknown; this limitation is explicitly acknowledged across nearly every domain reviewed, from intraoperative computer vision [16,20,21] to anastomotic leak prediction.[22,24,25] Fourth, this review focused predominantly on general, colorectal, oncologic, and orthopedic surgery, with more limited coverage of AI applications in cardiac, neurosurgical, and pediatric surgical decision-making, which may have distinct evidence bases and adoption trajectories. Finally, the regulatory analysis is US-FDA-centric; regulatory frameworks in the European Union (EU MDR, AI Act), United Kingdom, and other jurisdictions differ meaningfully and were not analyzed in comparable depth.

6. CONCLUSION

Artificial intelligence is reshaping surgical decision-making at every phase of care, with the strongest and most mature evidence supporting its use in preoperative risk stratification, where ML models now consistently outperform traditional clinical risk scores. Intraoperative computer-vision guidance and postoperative complication-prediction tools, including anastomotic leak prediction, show credible technical promise but remain earlier in their translational pathway, constrained by small, single-center datasets and limited real-time clinical integration. The regulatory pathway through which most AI/ML surgical devices reach market provides a comparatively low evidentiary bar, meaning that clinical adoption decisions should rest on independent scrutiny of the underlying validation literature rather than regulatory clearance alone. Unresolved challenges around algorithmic bias, the uncertain relationship between explainability and appropriate clinical trust, and unsettled questions of liability for AI-assisted decisions represent the principal barriers to safe, ethically sound, and clinically meaningful integration of AI into surgical decision-making. Priority areas for future research include prospective multicenter validation trials with patient-centered outcomes, standardized and transparent bias and performance reporting, and empirical study of how explainability techniques actually influence surgeon behavior and decision quality in practice.

REFERENCES

1. Limon D, Satish V, Raghavan N, Nguyen P, Rajesh A. Artificial Intelligence in Surgery Revisited: A 2025 Guide to Understanding and Applying AI Models in Clinical Practice. Am Surg. 2026;92(3):687-697. Doi: 10.1177/00031348251403592.
2. Yu E, Rosenberg GM, Udelsman BV, Harano T, Atay SM, Kim AW, Shakhsheer BA, Wightman SC. Narrative review of the ethics of artificial intelligence: are we ready for artificial intelligence in surgery? J Thorac

Dis. 2026;18(2):173. Doi: 10.21037/jtd-2025-1814.
3. Ruiz *et al*. Accuracy and Reliability of Artificial Intelligence in Surgical Decision-Making: A Literature Review. 2025.
4. Frampton AE, *et al*. Leveraging artificial intelligence to support surgical oncology multidisciplinary team decision-making: a systematic review. Front Artif Intell. 2026;9:1804004. Doi: 10.3389/frai.2026.1804004.
5. The ethical considerations of integrating artificial intelligence into surgery: a review. 2025.
6. Froicu EM, Creangă-Murariu I, Afrăsânie VA, Gafton B, Alexa-Stratulat T, Miron L, Pușcașu DM, Poroch V, Bacoanu G, Radu I, Marinca MV. Artificial Intelligence and Decision-Making in Oncology: A Review of Ethical, Legal and Informed Consent Challenges. Curr Oncol Rep. 2025. Doi: 10.1007/s11912-025-01698-8.
7. Development and Validation of a Machine Learning Model to Identify Patients Before Surgery at High Risk for Postoperative Adverse Events. JAMA Netw Open. 2023.
8. An automated machine learning-based model predicts postoperative mortality using readily-extractable preoperative electronic health record data. Br J Anaesth. 2019.
9. Compact machine learning model for perioperative stroke prediction prior to surgery: a retrospective cohort study. Sci Rep. 2025.
10. Machine learning and preoperative risk prediction: the machines are coming. J Thorac Cardiovasc Surg. 2024.
11. Machine learning-based preoperative prediction of perioperative venous thromboembolism in Chinese lung cancer patients: a retrospective cohort study. 2025.
12. A Risk Prediction Model (CMC-AKIX) for Postoperative Acute Kidney Injury Using Machine Learning: Algorithm Development and Validation. JMIR Med Inform. 2025.
13. Machine learning and the nomogram as the accurate tools for predicting postoperative malnutrition risk in esophageal cancer patients. 2025.
14. Artificial Intelligence and Machine Learning in Prediction of Surgical Complications: Current State, Applications, and Implications. Am Surg. 2022.
15. Artificial Intelligence in Vitreoretinal Surgery: A Systematic Review of Current Applications and Future Directions. Ophthalmol Ther. 2026. Doi: 10.1007/s40123-026-01347-8.
16. Surgical computer vision for intraoperative decision-support: a scoping review on performance metrics and readiness for real-time deployment. Artif Intell Surg. 2026.
17. Evaluation of Artificial Intelligence-Based Intraoperative Guidance Tools for Phacoemulsification Cataract Surgery. JAMA Ophthalmol. 2022.
18. AI-Enhanced Surgical Decision-Making in Orthopedics: From Preoperative Planning to Intraoperative Guidance and Real-Time Adaptation. 2025.
19. De Backer P. Intra-operative visual guidance through AI. Uroweb. 2024.
20. Artificial intelligence for intraoperative surgical guidance in robotic-assisted ventral cavity surgery: a systematic review on the current state of validation methods. npj Digit Med/npj Digit Surg. 2026. Doi: 10.1038/s44484-026-00013-7.
21. Artificial Intelligence and Digital Tools Across the Hepato-Pancreato-Biliary Surgical Pathway: A Systematic Review. 2025.
22. A Novel Deep Learning Model for Predicting Colorectal Anastomotic Leakage: A Pioneer Multicenter Transatlantic Study. J Clin Med. 2025;14(15):5462. Doi: 10.3390/jcm14155462.
23. A Novel Deep Learning Model for Predicting Colorectal Anastomotic Leakage: A Pioneer Multicenter Transatlantic Study [supplementary/PubMed record]. 2025. PMID 40807083.
24. Machine learning and deep learning to improve prevention of anastomotic leak after rectal cancer surgery. World J Gastrointest Surg. 2025. PMID 39872776.
25. Machine learning-based preoperative analytics for the prediction of anastomotic leakage in colorectal surgery: a Swiss pilot study. 2024.
26. Predicting the risk of colorectal anastomotic leak based on preoperative mapping of the blood supply of the bowel. arXiv. 2026.
27. Quantum Machine Learning for Predicting Anastomotic Leak: A Clinical Study. arXiv. 2025.
28. Analysis of FDA-Approved Artificial Intelligence and Machine Learning-Enabled

Cardiovascular Devices. JACC Adv. 2025. Doi: 10.1016/j.jacadv.2025.102174.
29. Trends in Clinical Validation and Usage of Food and Drug Administration (FDA)-Cleared Artificial Intelligence (AI) Algorithms for Medical Imaging. medRxiv. 2022.
30. FDA-Cleared Artificial Intelligence Medical Devices in Orthopaedic Surgery. J Am Acad Orthop Surg. 2025.
31. U.S. Food and Drug Administration. Artificial Intelligence in Software as a Medical Device. FDA.gov.
32. Regulating Flexibility for Artificial Intelligence: FDA Experience with Predetermined Change Control Plans. medRxiv. 2025.
33. Unveiling the Algorithm: The Role of Explainable Artificial Intelligence in Modern Surgery. Healthcare (Basel). 2025;13(24):3208. Doi: 10.3390/healthcare13243208.
34. How Explainable Artificial Intelligence Can Increase or Decrease Clinicians' Trust in AI Applications in Health Care: Systematic Review. J Med Internet Res. 2024.
35. On the practical, ethical, and legal necessity of clinical Artificial Intelligence explainability: an examination of key arguments. BMC Med Inform Decis Mak. 2025.
36. Explainable artificial intelligence in orthopedic surgery. J Am Acad Orthop Surg Glob Res Rev/PMC. 2024.