Epidemiology & Health Data Insights
(eISSN 3080-8111)

Review Article

Artificial Intelligence in Surgery: Current Applications and Future Prospects

Peter Aduvie Josiah[1], Chinedu Nwosu-Ijiomah[2], Ndidi Atasie Eboh[3]

[1]Department of Medicine and Surgery, College of Health Sciences, Niger Delta University, Bayelsa State, Nigeria
[2]Department of Health Informatics, School of Computing, University of West London, UK
[3]Department of Nursing Science, Stark State College, North Canton, OH, USA

Received: Apr 05, 2026
Accepted: May 28, 2026

Corresponding author's email:
peterjosi51@gmail.com

Citation: Josiah PA, Nwosu-Ijiomah C, Eboh NA. Artificial Intelligence in Surgery: Current Applications and Future Prospects. Epidemiology and Health Data Insights. 2026;2(4):ehdi043.
https://doi.org/10.63946/ehdi/18696

ABSTRACT

Artificial intelligence (AI) is transforming the practice of surgery through the support of clinical decision-making, improved accuracy and efficiency, and better patient outcomes across the perioperative period. This review evaluates and critically analyzes current, technological and evidence-based applications of AI in surgery. Machine learning, deep learning, computer vision and natural language processing are some of the core AI technologies that are enabling developments in all phases of surgery. Some of these applications include, improved preoperative diagnosis, intraoperative real-time imaging, workflow analysis, robotic surgery assistance and intraoperative clinical decision support. Postoperatively, the potential of AI is also vast. Through the use of predictive models for surveillance and early complications detection, as well as remote management of patients, AI is optimizing postoperative care.

Although the findings suggest promise for the future, the clinical adoption of AI in surgery is limited by a number of issues, including data quality and heterogeneity, lack of validation in large prospective studies, potential for bias, ethical concerns and high costs of implementation. Other factors including clinician acceptance, data privacy, regulatory approval and medico-legal implications need to be addressed before any new technology is widely adopted.

Future directions of AI in surgery include progression to semi-autonomous systems that augment the efforts of the surgeon, integration with emerging technologies such as genomics and digital twins, and increased use in low resource settings to help address existing global inequities in surgical care. For successful AI adoption in surgery, AI systems will need to undergo robust validation, require interdisciplinary approaches, and all systems must be developed in a manner that is not only transparent but also clinically appropriate for surgeons, so that it can maximize the benefits of human talent in the operating room while maintaining patient safety.

Keywords: Artificial Intelligence; Robotic Surgery; Machine Learning; Perioperative Care; Clinical Decision Support

Introduction

Background and Rationale

Surgical care has become more complex in recent times due to various patient, institutional and societal factors. Patient populations continue to grow older with more co-morbidities, there are expanded therapeutic options, and higher societal expectations of medical intervention [1]. Healthcare systems worldwide face mounting pressure to reduce complications, hospital stay, and optimize healthcare resources [2,3]. This places significant pressure on modern surgeons who already work in high stakes, time critical environments and must integrate patient clinical history, imaging and laboratory data in performance of their duties. All these heighten the demand for innovative tools to support surgical decision making [4].

Surgical practice traditionally has relied heavily on human expertise, pattern recognition, and manual dexterity. Although surgical training and experience forms the foundations of surgical practice, human performance has fundamental limitations due to cognitive overload, fatigue, possibility of bias and the variability of practitioners [5,6]. These factors can affect technical execution, intraoperative decision making, and postoperative management, ultimately leading to different patient outcomes. There is therefore a drive to introduce technologies and systems that augments human capabilities, improves situational awareness and facilitates evidence based decision making throughout the surgical pathway [7].

Definition and Scope of Artificial Intelligence

Artificial intelligence (AI) refers to computational systems that have been designed to mimic and replicate actions that usually require human intelligence such as pattern recognition, learning, prediction and reasoning. Machine learning, a concept within the broader field of AI involves the programming of algorithms to learn from provided data, predict, classify and execute multiple tasks without being explicitly programmed to perform each task. A subset of machine learning known as deep learning utilizes multiple layered neural networks to effectively process highly complex, high-dimensional data, such as images and videos [8].

The current applications of AI in healthcare is somewhat narrow and can be significantly broadened. Current applications do not approach human level cognition across domains, rather it is tailored to perform specific, well defined tasks. Computer vision, predictive analytics, natural language processing, and reinforcement learning are currently the most relevant AI modalities in surgery [8,9].

Evolution of AI in Surgical Practice

There already have been significant improvements in the applications of AI in surgery—from the early computer assisted systems that were primarily used for image guidance and systems navigation to currently more advanced, data-driven approaches. The development of powerful computational systems with large data-set and advanced algorithms have enabled the availability of models that performs real time analysis and prediction. The potential applications of AI in surgery during the perioperative, intraoperative and postoperative periods have been expanded by the integration of AI in robotic platforms, advanced imaging modalities and electronic health records [10,11].

Aim and Structure of the Review

This review aims to critically evaluate the current applications of AI in surgery, assess the strength of existing evidence, as well as explore future and potential applications while recognizing and addressing persisting challenges. It begins with an overview of core AI concepts, followed by an assessment of perioperative applications, ethical and regulatory concerns, and finally, future directions for research and clinical implementation.

Methodology

This study was conducted as a narrative review of the literature on artificial intelligence applications in surgery. A comprehensive search was performed across multiple databases including PubMed, Google Scholar, and the Cochrane Library. The primary search terms used included “AI in surgery,” “machine learning in surgery,” “deep learning in surgery,” “computer vision surgery,” “robotic surgery AI,” and related combinations. No strict time limits were applied, although priority was given to studies published within the last 10 years to capture recent advancements. The reference lists of retrieved articles were also manually screened for additional relevant studies (snowballing technique).

Articles were selected based on relevance to the core themes of this review (preoperative, intraoperative, and postoperative applications of AI in surgery, surgical training, as well as ethical and regulatory considerations). Both original research studies and high-quality review articles were included. As this is a narrative (rather than systematic) review, formal PRISMA reporting, predefined inclusion/exclusion criteria, and systematic quality appraisal of all studies were not performed. However, emphasis was placed on landmark studies, high-impact publications, and those with robust methodologies to ensure a balanced and critical synthesis of the current evidence.

Fundamentals of Artificial Intelligence in Surgery

Core AI Technologies

Artificial intelligence applications in surgery are primarily powered by machine learning techniques that utilizes complex datasets to detect patterns and make predictions [12]. Tasks such as stratifying patient risks, predicting postoperative outcomes, and classifying medical images, are carried out by supervised learning algorithms that have been trained on labelled data [13,14]. By contrast, unsupervised learning algorithms have proven very valuable in grouping similar patients and analyzing surgical workflows by being able to uncover hidden structures within unlabelled data [15]. Reinforcement learning, in which agents learn optimal behavior through trial and error interactions with an environment, is enabling adaptive intraoperative decision making and attracting increasing interest in control of surgical robots [16].

Deep learning is a very important subset of machine learning. It captures intricate, non-linear relationships in high dimensional data utilizing multilayered artificial neural networks. Most image and video based surgical applications are formed particularly by convolutional neural networks which are prolific in processing visual information. Deep learning powers computer vision, enabling the automatic identification of anatomical structures, surgical equipment, and procedural phases from intraoperative videos or still images. Another key modality is natural language processing which is able to extract vital medical information from unstructured text, pathology narratives, operation notes and electronic medical record notes [17-19].

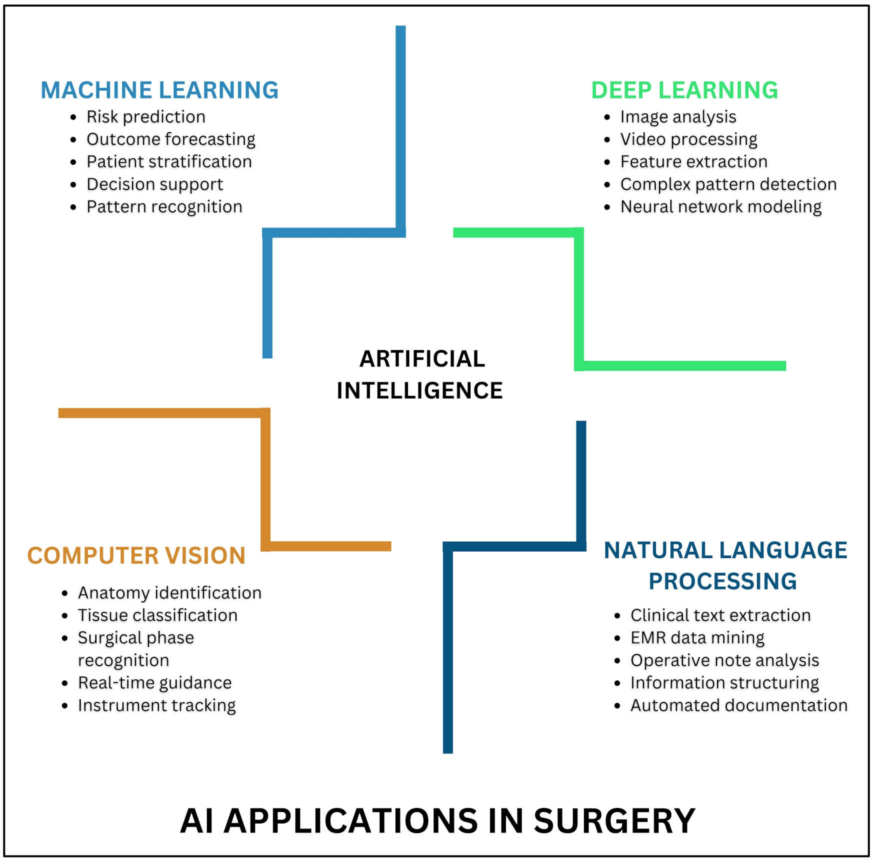
Data Sources for Surgical AI

The quality, volume and diversity of the data used in training and evaluation is critical to the effectiveness of surgical AI systems. The primary data input is provided by stills and videos from imaging modalities such as computed tomography, ultrasonography, MRI and endoscopy [20,21]. These data are used for diagnostic, planning and guidance applications. Intraoperative video streams from robotic and minimally invasive procedures provided valuable temporal data that enhances real time navigation, technical skill assessment and workflow recognition [22]. Electronic medical records provide structured data such as demographics, vital signs, co-morbidities and laboratory results, and unstructured narrative context, along with longitudinal outcome measures. Intraoperative monitoring systems and wearable sensors are emerging sources that provide continuous streams of physiological and performance related data, further boosting the dataset landscape [23,24].

Model Development and Validation

An optimal model development requires unbiased performance evaluation through the partitioning of data into separate training, validation and independent testing sets [25]. An enduring challenge of AI in Surgery is overfitting: This is when AI models perform optimally on the training datasets but poorly in institutional, new patients and clinical contexts [26]. This reinforces the need for external validation and training on large multi-center, and geographically diverse datasets. There is also growing interest in explainable AI methods. Techniques with more transparent model reasoning helps build physician confidence, promotes regulatory approval and facilitates safer, and more responsible integration of AI in clinical decision making [27,28]. The various AI applications useful in surgery are described in Figure 1.

Figure 1: Overview of Artificial Intelligence Applications in Surgery



Preoperative Applications of Artificial Intelligence

Diagnostic Support and Risk Stratification

Preoperative assessment is a key component of the surgical process. Accurate diagnosis and risk assessment is required in order to determine the most appropriate surgical technique to be employed and to predict the patient's postoperative course. Artificial intelligence holds much promise for improving diagnostic accuracy from medical images. Machine learning and deep learning algorithms have been shown to be capable of automatically recognising anatomical structures, identifying diseases and staging them to an accuracy that is often comparable to, or even greater than, that of a human pathologist. In oncology, artificial intelligence has been employed to improve tumour detection, margin evaluation and lymph node assessment, all of which can aid in preoperative planning of the surgical intervention [29-33].

The vast majority of works related to Artificial Intelligence (AI) in surgery have been focused on the diagnostic aspect. However, there is a growing interest for using AI to predict surgical risk and postoperative complications. Predictive models using patients' demographics, comorbidities, lab values and imaging characteristics can be used to estimate the probability of certain postoperative complications such as Surgical Site Infections (SSIs), Adverse Lung Injury (ALI) or hospital mortality. AI-based risk scores are expected to outperform traditional risk scores that are most of the time based on simple and often outdated scoring systems. Indeed, AI-based risk scores will enable a more refined stratification of surgical risk and will allow for a more personalized management of high risk surgical patients through the implementation of prehabilitation programs to improve the overall physical condition of high risk patients before surgery. This can be achieved through the implementation of less traumatic surgical techniques such as robotic-assisted surgery to minimize tissue injury and finally through the implementation of close postoperative surveillance to allow for early detection of postoperative complications that can then be managed more appropriately [34-36].

Patient Selection and Surgical Decision-Making

Optimization of patient selection is still crucial to ensure a good surgical outcome and safety. There are currently steps being taken to facilitate the use of artificial intelligence (AI) for designing a clinical decision support system (CDS) that will help a surgeon decide on the best operation for a patient given the individual patient's topography, medical history etc. and will also weigh out the benefits and harms of the proposed surgery for the patient [37]. The proposed system will utilize big datasets to form algorithms that will offer suggestions to the surgeon but will not be restrictive in cases where personal discretion is required. Using the AI in complex high risk surgical procedures will identify more variables that can impact the patient's risk to undergo such a surgery, and it will direct the patient and the surgeon to an alternative treatment that has a strong evidence base to support the choice. Ultimately a CDS will assist in shared decision making between the physician and the patient [35,37].

Using AI for preoperative planning does not necessarily mean that human judgement will be less needed in the actual operation. A whole new set of questions arise regarding explainability, interpretability and continued clinician involvement. Models have to be explainable and aligned with medical knowledge and practice in order for clinicians to be able to properly understand, agree with and trust the information provided [28].

Preoperative Planning and Simulation

Artificial intelligence (AI) is revolutionising preoperative planning through the use of automated anatomical segmentation and patient specific 3D reconstruction [38]. With accurate segmentation of key anatomy, surgeons are able to create detailed operative maps and more bespoke surgical approaches. Using the power of AI, VR and AR, surgeons can rehearse and optimize complex surgical interventions in a risk free environment, using the patient's own scan data [39-42].

Artificial Intelligence (AI) has been used to train and develop predictive models to estimate the operative difficulty of the surgery, estimated time for surgery and the postoperative recovery trajectory. These predictive models can be used to prepare the patient, surgical team and the healthcare institution in advance of the procedure to ensure optimal readiness. Ultimately, AI can be used to reduce the preoperative and intraoperative uncertainty, and to prepare the surgical team and healthcare institution well in advance for what is to come in the operating room [34,35,36,43].

Intraoperative Applications of Artificial Intelligence

Computer Vision and Real-Time Surgical Guidance

One of the most prevalent uses of AI in the OR is based on computer vision, and while these uses are currently more demonstrated than proven, they are also the most immediate ones that can be implemented in the OR [44,45]. Many surgeries are now performed using minimally invasive techniques or robotic systems and, therefore, there is a high definition video feed coming from the cameras [11,23]. All the while, deep learning algorithms can be trained and applied to the real time video feed to automatically and in real time identify anatomical structures such as organs, vessels, nerves and pathologies [39,46]. This therefore enhances the surgeon's knowledge of the environment in which they are operating and can prevent injuries, particularly during dissections, where these injuries are more likely to occur, and in cases where the field is distorted by prior surgeries, inflammation or tumor growth.

Artificial Intelligence (AI) also goes beyond simple recognition to being very instrumental in pathology object detection, tissue classification and boundary detection. These applications in pathology are much more than just being able to differentiate between normal and abnormal/pathological and malignant/benign tissues but can also be used to identify and describe the tumor margins [29-32]. This is highly critical in oncology because, on one side it is important to ensure that tumours are resected with sufficiently wide margins, on the other side, it is important to avoid injury to surrounding critical structures. Augmented Reality (AR) brings additional benefits on top of the already mentioned uses of the AI in Pathology by providing AI generated annotations of the preoperative anatomy which are then superimposed on to the live view of the surgeon and the intraoperative real time video stream of the working field [38,42]. In this way, AI assist in enabling a more accurate and controlled surgery, especially in the key and critical stages of the surgery when the required surgical performance is at the highest level.

AI in Minimally Invasive and Robotic Surgery

The increasing prevalence of artificial intelligence (AI) in the field of minimally invasive and robotic surgery has opened up numerous avenues for intraoperative support. With the use of machine learning (ML) techniques, the analysis of instrument motion, force, and vision allow for more dexterous, stable and precise surgery. Motion-tracking technology allows the physiological tremor inherent to the human hand to be filtered out, resulting in more accurate and smooth movements. Tasks that demand extreme precision, such as micro-suturing, vessel anastomosis or dissection, benefit from the improvement in accuracy achieved by this technology [11,12,23,47].

Everyday, the new, more advanced minimally invasive and robotic surgical systems are doing more of the work with the surgeon having to make only the most important decisions. The system's AI does the smaller tasks. For instance, the system can control the camera, do the suturing through the arms of the robot, or grip and pick up tissue. The AI is far from being able to do all the tasks on its own, and there are many years of research before the AI is ready to replace the human surgeon, but in the meantime, the systems are helping to ease the work load of the surgeon allowing them to be more efficient and deliver better results for longer and more complex surgeries [4,7,11,23].

Surgical Workflow Analysis

AI also supports deep analysis of surgical workflows. The main goal is to describe in a highly detailed, automated manner, the individual surgical actions performed during the surgery and to divide the surgery workflow into major stages to be able to track the surgical progress, predict the upcoming tasks for the surgeon and give real-time clinical hints and alerts to assist him. Workflow analysis provides a quantitative evaluation of the surgical performance and gives the opportunity to identify the bottlenecks, unnecessary steps and inter-patient variability in terms of the surgeon's behaviour. Based on these insights, the efficiency improving interventions and tools can be then tailored for each individual surgeon and hospital [43,44].

Artificial intelligence (AI) can also enhance the surgical workflow monitoring and thereby error detection and prevention. AI can real time monitor the steps taken during surgery and the use of surgical instruments and alert for any deviation from a pre-defined normative surgical workflow. In case any aberration is detected in real time, the AI system flags the surgical team in a timely manner, thereby preventing small errors from potentially being translated into significant harm to the patient. Thus, AI acts as a real time quality assurance mechanism in the operating room and can be considered as an extension to existing quality and patient safety assurance measures in place in the OR [43,44,48].

Intraoperative Decision Support

Intraoperative decision support is one of the most exciting and challenging applications of AI in surgery. Ideally, all available information from all sources should be combined and utilized in real time to detect any potential intraoperative complications such as uncontrolled bleeding, drop in blood pressure or surgical errors. Alerts should be generated in real time to allow the surgeon to correct an emerging complication [34,37]. In addition, the AI system can provide the probability of occurrence of any intraoperative or postoperative adverse event which can be very useful to guide the timing and nature of any corrective intervention [35,36,49].

Alerts are just a small part of what a surgical AI system can do. The system can also help a surgeon to modify or change the steps that he is taking based on the real time data that is made available to him. The AI system always remains on the sidelines of a procedure and ultimate control and decision making lies with the surgeon based on his assessment of the patient's needs. What the system does is enable the surgeon to make those decisions more effectively by helping them make it more quickly, with more certainty, and with less risk of error. That is important because surgery is a very dynamic environment and there is a lot of uncertainty involved in any surgical procedure [34,36,37].

Postoperative Applications of Artificial Intelligence

Outcome Prediction and Complication Detection

The postoperative period is a dynamic state of clinical changes and it is the period when the risk of complications are substantial which can influence the recovery and long term survival of the surgical patient. Artificial intelligence has been shown to be effective in the prediction and early detection of postoperative complications by taking into account perioperative data, including preoperative patient characteristics (demographics, comorbidities) as well as intraoperative information and serial trends in laboratory values. Examples of postoperative complications that could be detected early in the postoperative period include surgical site infection, sepsis, venous thromboembolic event and organ failure [35,36,50,51].

Modern artificial intelligence (AI) monitoring systems evolve beyond simple one-time risk scores and instead provide a dynamic, real-time continuous assessment of the patient status. By utilising machine learning, the AI system can rapidly detect minute and often asymptomatic changes to the patients vital signs, respiratory patterns or laboratory values, often hours, sometimes days before these changes result in a recognisable clinical deterioration. Early warning systems can enable timely and preventable clinical interventions and thus help reduce avoidable "failure to rescue" adverse events, ultimately enhancing patient safety, especially in the postoperative period [35,36,50,52,53].

Postoperative Care Optimization

Using Artificial Intelligence (AI) to enhance postoperative care and optimize resource utilization has become a growing trend in the field of healthcare. Length of stay (LOS) predictive models in the hospital setting allows healthcare providers to assess and better anticipate the postoperative needs of their patients. Identifying patients who may require additional interventions, such as early mobilization, increased physiotherapy, and nutrition or social work services before discharge is invaluable. Similarly, using readmission risk models that incorporate traditional clinical risk factors along with socio-economic risk factors can alert providers to patients at high risk for readmission and allow for additional interventions prior to discharge [35,50,54].

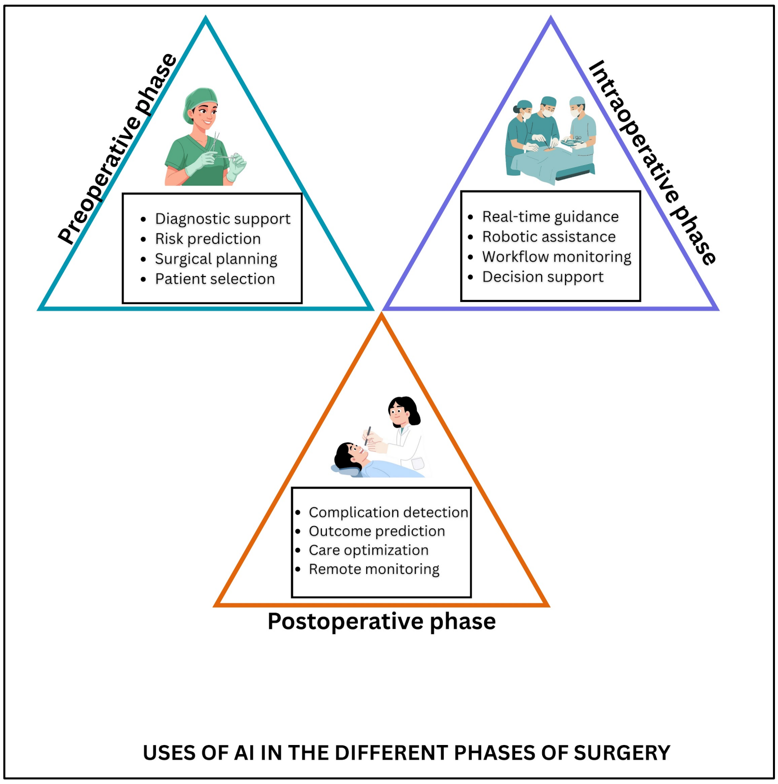
Predictive models also aid in the effective implementation of enhanced recovery after surgery (ERAS) protocols. The AI is thought to improve on a one size fits all approach of current care pathways for surgery, by providing a tailored postoperative plan, which optimises the individual postoperative outcomes. It does this by identifying the patient's risks profile and tailoring analgesia regimens, feeding schedules and activity goals. The concept therefore has the potential for real world improvements in patient outcomes, with reduced interventions while saving costs for the health care systems [55].

Remote Monitoring and Follow-Up

The application of artificial intelligence to wearable sensors and remote monitoring has revolutionised the way post-surgical care is delivered. Many types of wearable sensors have become available to track heart rate and rhythm, respiratory rate, activity, sleep and a myriad of other parameters. The wearable sensor data can be transmitted to an AI platform, where analysis of large amounts of data derived from an individual patient can identify patterns of change that are indicative of a developing complication, such as wound infection, arrhythmia or thromboembolic events before any symptoms develop [24,56].

The AI-based remote monitoring solution that integrates with telemedicine platforms ensures that the communication between the patient and the clinical team is timely and relevant, less frequent and hence more convenient for patients, while maintaining a high level of clinical engagement. Overall, AI provides continuity of care, supports safer and earlier hospital discharge, improves patient engagement and overall recovery experience [55,56]. Figure 2 illustrates the multiple roles of AI across the entire surgical workflow, highlighting how these applications interconnect to support continuity of care from preoperative planning through postoperative recovery.

Figure 2: Roles of Artificial Intelligence Across the Surgical Workflow



Artificial Intelligence in Surgical Training and Education

Objective Assessment of Surgical Skills

Current approaches for evaluating surgical competence primarily rely on clinical assessment provided by supervising physicians, which is inherently subjective and may result in high inter-observer variability, as well as being restricted by time and number of observations that can be feasibly performed. AI can provide a more objective, and scalable, approach for the evaluation of surgical performance [57].

By processing video data from intraoperative footage or simulation exercises, AI can analyze and measure the technical aspects of surgical skills such as motion characteristics of surgical instruments, task performance, handling of tissues and overall surgical flow. AI can provide quantitative and precise measures that can serve as a tool to augment human evaluation, thereby enabling inter and intra-individual benchmarking between surgeons and trainees across institutions [44,57,58].

This AI-based skill assessment feature enables longitudinal tracking of learner skill development, identifies areas for improvement and maps current skill levels to industry-wide benchmarks for comparison. Objective assessments like this help bring about fairness in hiring and promotion processes, more clarity in career advancement and an official stamp of approval in formal certifications [57,58].

Simulation-Based Training

Simulation is a key component of modern surgical education. Advances in artificial intelligence have transformed simulation-based learning into a highly-realistic and effective educational tool. An AI-based Virtual Reality (VR) simulator can provide trainees with a highly-realistic and anatomically-accurate simulation of the complex surgical procedures they need to learn. The trainee can practice repeatedly within a risk-free environment to an unlimited number of times to enhance their understanding and skills. The VBS (VR-based-Simulator) will continuously assess the abilities and the learning pace of the trainee. As the trainee demonstrates an understanding of the skills that need to be applied, the difficulty level and the challenges will increase [59,60].

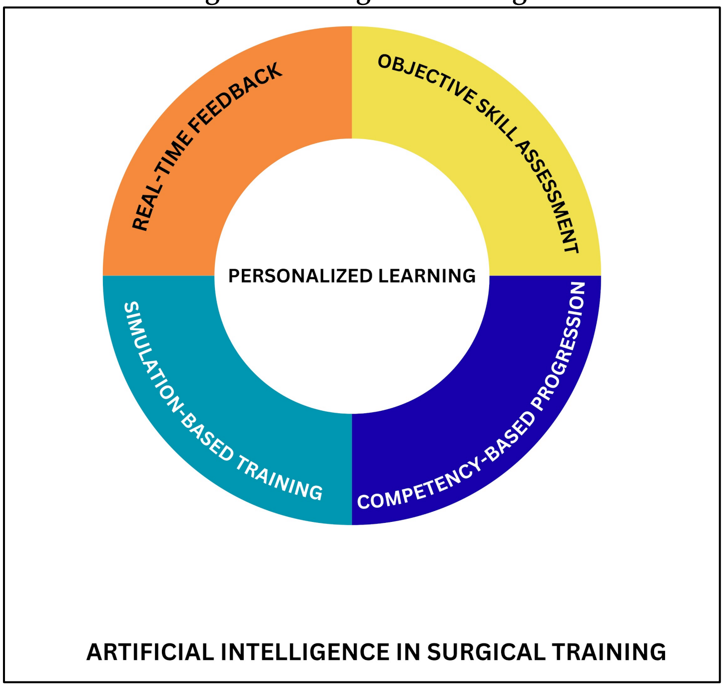
One of the key benefits of using an AI-enhanced simulation for training is that the learner receives instant and highly relevant feedback. As opposed to receiving instructor comments sometime in the future, the simulation provides a stream of quantitative feedback on an infinite number of trials immediately in real time. This quantitative, instant feedback enables the learner to track and optimize their metrics such as effectiveness, accuracy and frequency of their mistakes. Instant feedback to the learner promotes faster learning, increases effectiveness of structured practice and is highly beneficial for learning of complex procedures [58,59,60].

Competency-Based Surgical Education

The move to competency based surgical education has evolved to be more focused on attainment of skills and understanding rather than time based progression through a curriculum. Artificial intelligence allows for the assessment of key competency milestones in a consistent and objective manner despite the diverse settings in which trainees may work. Utilisation of Artificial intelligence in assessment technology reduces human bias and allows for rapid and reproducible assessment of clinical, technical and procedural skills [58,61].

Artificial Intelligence (AI) can be a valuable performance benchmarking and competency standard-setting tool in surgical education. AI can assist educators in setting learning targets for their students and in designing individualized curricula tailored to the needs of the learner. Ultimately, the judicious use of AI in surgical education can help to improve the standard of surgical training, to enhance patient safety and to produce highly competent future surgeons [9,39,58,61,62]. The key advantages of incorporating AI into surgical education are synthesized in Figure 3.

Figure 3: Benefits of Artificial Intelligence in Surgical Training and Education



Evidence Base and Clinical Validation

Current Level of Evidence

The body of literature about the application of artificial intelligence (AI) in surgery is expanding. While the scope and quality of studies vary, the majority of reports are based on single centre, observational, retrospective cohort studies, as well as proof of concept studies validating the performance of an algorithm in a specific clinical or simulated setting. Although these studies suggest that AI can be used for a wide range of applications, from image and workflow analysis to the prediction of intraoperative and postoperative variables, they all are based on predefined, and in many cases highly controlled, scenarios. In these scenarios, AI performed with high precision and reproducibility [34,35,43].

The first pilot trials and feasibility studies are now underway to evaluate the usability, safety and early impact of these systems in real-world clinical practice. While they look promising in terms of time savings, accuracy and helping surgeons with their decision-making, these initial trials are designed to prove the feasibility of the technology and will not provide the final word on its clinical impact [63].

Evidence Appraisal: Current Demonstrated Benefit vs Potential Applications

While artificial intelligence shows considerable promise across the surgical pathway, the strength of evidence varies significantly by application. This section distinguishes between applications with demonstrated clinical benefit in prospective or randomized studies, those undergoing active testing, and those that remain largely proposed or proof-of-concept.

Applications with Demonstrated Clinical Benefit

- Intraoperative hypotension prediction: The HYPE randomized clinical trial demonstrated that a machine learning-derived early warning system significantly reduced the depth and duration of intraoperative hypotension compared with standard care in patients undergoing elective noncardiac surgery [46].

- Preoperative risk stratification: The POTTER AI calculator has been prospectively validated in a bi-institutional study of emergency general surgery patients undergoing laparotomy, showing high accuracy in predicting mortality and complications, superior to traditional tools in some aspects, and rated as user-friendly for preoperative counseling [33].

- Postoperative complication risk prediction: MySurgeryRisk, a machine learning algorithm using preoperative clinical data, has undergone large-scale retrospective development and external validation across tens of thousands of patients, demonstrating strong predictive performance for major complications and mortality [35].

Applications Undergoing Testing / Early Clinical Evaluation

Several computer vision and workflow analysis systems for phase recognition, tool tracking, and critical view assessment are being evaluated in real-world operating rooms. Similarly, AI-assisted robotic applications (e.g., improved pedicle screw placement accuracy) have shown promising results in controlled clinical studies, though larger prospective trials are still needed. AI-based objective surgical skill assessment tools are also being tested for integration into training curricula [39,46].

Proposed or Emerging Applications

Most other applications—including advanced real-time intraoperative decision support, fully contextual gesture analysis for functional outcomes, remote postoperative monitoring with wearables, digital twin simulations, and semi-autonomous robotic systems—currently rely on retrospective data, single-center studies, or simulation-based validation. While these demonstrate technical feasibility and encouraging performance metrics, robust prospective evidence linking them to improved patient-oriented outcomes (e.g., reduced complications, shorter hospital stays, or better long-term functional results) remains limited [47,51,57].

This evidence appraisal highlights that while certain narrow AI applications have already transitioned into clinical use with measurable benefits, the majority of AI innovations in surgery are still in earlier stages of development and validation. Future high-quality prospective multicenter trials will be essential to confirm broader clinical impact.

Challenges in AI Model Validation and Real-World Implementation

Although many AI models in surgery show strong internal performance (e.g., high accuracy or AUC), critical aspects of validation remain inadequately addressed.

External and Temporal Validation: Most studies are limited to internal or single-center validation. External validation across diverse institutions, populations, and regions, as well as temporal validation on newer data, is still uncommon, limiting generalizability.

Transportability Across Centres: Variations in surgical practices, patient demographics, equipment, and data systems often cause significant performance decay when models are moved from high-volume academic centers to community or low-resource settings.

Model Drift and Post-Deployment Monitoring: AI systems are susceptible to performance degradation over time due to concept drift and data drift. Few studies implement ongoing monitoring protocols to detect and address such issues after deployment.

Robust multi-center external validation, temporal testing, and continuous post-deployment monitoring are essential for safe clinical use. Future efforts should align with evolving regulatory expectations (e.g., FDA) and adopt standardized reporting to build clinician trust and facilitate responsible adoption.

Limitations of Existing Evidence

Although many AI applications in surgery show promising results, the current body of evidence has important limitations. Most studies are retrospective, single-centre, and based on relatively small sample sizes, which reduces statistical power and introduces potential centre-specific biases that limit generalizability to real-world practice. Comparisons across studies are further hindered by heterogeneity in data quality, annotation protocols, and outcome measures [64].

A major shortcoming is the scarcity of randomized controlled trials evaluating patient-oriented outcomes. Without robust comparators, it remains unclear whether the observed improvements translate into meaningful clinical benefits. Additional common sources of bias in the surgical AI literature include selection bias, data leakage, missing data, unrepresentative training datasets, and biased outcome labelling. Consequently, many reported performance metrics may be overly optimistic and may not hold in broader clinical settings [64-66].

Future research should prioritize large-scale, multicenter prospective studies and systematic reviews that incorporate standardized quality appraisal tools to provide more reliable evidence on the true clinical value of AI in surgery.

Regulatory Approval and Clinical Adoption

Regulatory clearance is a major step in the translation of surgical AI into the clinic. In the US, the FDA has recently established new regulatory classifications for software-based medical devices. Other regulatory bodies have begun to evolve their own policies. Although the software clearance addresses many of the technological hurdles to adoption of surgical AI, there are many other issues that still need to be addressed in order to fully implement and deploy a surgical AI system in a hospital. Examples of some of these challenges include integration of the AI system into the hospital healthcare infrastructure, acceptance of the technology by human surgeons and clinicians, ethical considerations and legal liability that may arise, and the on-going necessity of validating, tracking and updating an AI based surgical system [67-71].

Table 1: Summary of Key Empirical Studies on AI Applications in Surgery

Study	Key AI Application	Study Design	Sample Size	Validation Approach	Comparator	Performance Metrics	Key Clinical Outcomes
Panossian et al (2025) [33]	Preoperative risk prediction (POTTER AI calculator – emergency general surgery)	Prospective, bi-institutional cohort study	361 patients (EGS laparotomy)	Prospective real-world external validation	Traditional surgeon judgment / existing risk tools	High accuracy (c-statistic ~0.90 for mortality; 0.80–0.89 for complications)	Superior prediction of mortality & morbidity; useful bedside tool for preoperative counseling
Bihorac et al (2019) [35]	Preoperative/postoperative risk prediction for major complications & death	Retrospective development & validation (multi-cohort)	51,457 patients (major inpatient surgery)	Internal + external validation on independent cohorts	Traditional risk scores	High AUCs outperforming conventional models	Accurate prediction of complications & mortality; supports perioperative decision-making
Wijnberge et al	Intraoperative hypotension early warning	Randomized controlled trial	~60–70 per arm (elective	Randomized, real-time	Standard care (no AI alert)	Significant reduction in depth &	Reduced intraoperative

Study	Key AI Application	Study Design	Sample Size	Validation Approach	Comparator	Performance Metrics	Key Clinical Outcomes
(2020) [46]	system (HYPE Trial)		noncardiac surgery)	intraoperative use		duration of hypotension	hypotension exposure
Golany et al (2022) [47]	Computer vision – surgical phase recognition (laparoscopic cholecystectomy)	Retrospective video analysis + model development	371 videos (multi-hospital, incl. complex cases)	80:20 train-test split stratified by complexity/ hospital/adverse events	Human surgeon annotations	Mean accuracy 89% (comparable to inter-surgeon agreement)	Robust phase recognition even in complex cases; supports real-time workflow, alerts & education
Xiao et al (2025) [49]	Robotic AI-assisted pedicle screw placement	Retrospective controlled trial	50 patients (20 AI-robotic vs 30 manual)	Retrospective CT-based evaluation	Manual fluoroscopy-guided placement	Higher screw accuracy (Gertzbein-Robbins grading), reduced fluoroscopy	Improved accuracy, lower radiation exposure, shorter hospital stay, better pain scores
Heard et al (2026) [51]	Contextual gesture recognition (anatomy + function) for outcome prediction (RARP)	Retrospective multi-center video analysis	147 patients (26 surgeons)	Internal validation with transformer model	Gesture sequences alone	AUC improved from 0.78 to 0.85 with contextual features	Better prediction of 1-year erectile function recovery; improved explainability
Yanagida et al (2025) [57]	AI-based surgical skill assessment via phase recognition (laparoscopic cholecystectomy)	Retrospective video analysis	Multiple laparoscopic cholecystectomy videos	Comparison against expert surgeons	Expert surgeon assessment	Strong correlation of automated metrics with skill levels	Objective, scalable skill assessment for training & competency evaluation

Abbreviations: AI, artificial intelligence; EGS, emergency general surgery; AUC, area under the curve; CT, computed tomography; RARP, robot-assisted radical prostatectomy; c-statistic, concordance statistic; vs, versus.

Ethical, Legal, and Practical Considerations

Data Privacy and Security

The development and application of artificial intelligence (AI) in surgery require substantial volumes of high value, personal data. This raises a multitude of privacy and security issues. Gaining a valid consent for reuse of such data for the purpose of training new generations of algorithms is another on-going ethical issue. This should be addressed through robust regulation of data governance; including clarification of data custody and use rights, stipulation of acceptable data uses and the allocation of respective duties and liabilities between hospital and clinician and/or the developer of the technology. Increasing digitisation of and connectivity within surgical devices has also created new cyber security threats including the potential for the system to be compromised, or data that is held within the system to be stolen. All of these represent new patient privacy and safety risks that have not previously been addressed in current surgical practice [72-74].

Bias, Fairness, and Equity

Algorithmic bias is arguably the most significant ethical challenge in surgical AI. The failure of an AI system to provide equally accurate or predictable performance across all relevant demographic groups, clinical presentations, or healthcare settings is a source of potential bias and discrimination [75]. If the training dataset lacks representation from various demographics, conditions, or clinical settings, then a broader deployment of the model is likely to yield a suboptimal performance in untested groups or be marred by the introduction of new bias. Populations that have been historically marginalised or neglected such as ethnic minorities, low socioeconomic populations, and those from low resource settings or with complex comorbidities are at risk of having their disparities in healthcare worsened. Reducing bias in an AI system therefore requires the use of a diverse training set, identification of any unanticipated performance gaps when the model is tested in subgroups, and the regular surveillance of how the system performs in these populations to prevent future discrimination [75,76].

Accountability and Liability

The use of AI in surgical decision making raises a host of new and complex issues regarding professional accountability and medico-legal liability. In the event of a surgical complication occurring following the use of an AI-assisted decision-making tool, there are many unclear and contentious issues that remain to be resolved. However, as AI tools are intended to complement rather than replace the human element of surgery, the balance of accountability for the success or failure of surgical interventions in the future is likely to be complex, with significant uncertainty as to the respective roles and burdens of responsibility for surgeons, hospitals and manufacturers of such technology. It is therefore essential that there is a clear and updated medico-legal framework together with professional and collegiate guidance, regarding the appropriate use and documentation of such technologies, as well as the acceptability of clinical reliance on their outputs [73,74,76-78].

Surgeon–AI Relationship

Surgeon-AI interaction will be critical for the success of any clinical deployment. An AI system should be viewed as a tool that supports the clinical activities of surgeons, augmenting their clinical skills, expertise and dexterity. For a system to be acceptable to surgeons, it must be transparent, interpretable and deliver clear clinical benefit. Explainability, good user centred design will be critical to obtaining surgeon trust and ensuring that any decision that is assisted by the system is one that is clinically appropriate and makes sense. There will be a need for multi-professional collaboration between surgeons, engineers, ethicists and policy makers to ensure that AI is successfully delivered as a tool in surgery [10,28,72,73].

Future Prospects of Artificial Intelligence in Surgery

Toward Autonomous Surgery

There are many other future applications of AI in surgery that are of great interest and discussion, including the ability to bring surgery to an increasingly autonomous level. Currently, most of the AI in surgery applications fall back into one of the lower levels of autonomy and are therefore limited to the provision of intelligent decision support, or the performance of small, discrete components of the surgical task under close human supervision. Most current modelling frameworks also propose a more graduated approach to increasing the autonomy of surgical AI systems, starting from the task-assistance level and gradually advancing towards conditional autonomy [79].

Significant progress has been made in robotics, computer vision and reinforcement learning, yet carrying out surgery autonomously in its entirety remains far from being technically possible. The reasons are manifold: from the variability of the human body to the uncertainties inherent to the operating room and the need for human professionals to bring together contextual knowledge and good sense of judgement at any given time. Finally, ethical, legal and medical safety issues impose effective brakes on any hasty drive for full autonomy [16,23,73].

Integration with Emerging Technologies

The potential of AI in surgery will be further enhanced when it is combined with other emerging advancements. When combined with genomics, AI can lead to true surgical precision. Precision surgery using genomics in combination with AI technology will use genomic risk assessment, tumor molecular characteristics and drug response to tailor an individualized surgical plan that maximizes the benefit to the patient while minimizing surgical trauma. These AI driven precision medicine approaches can then further aid in the selection of the most appropriate time for surgical intervention and with optimization of the perioperative period [80,81].

One of the most exciting trends in surgery today is digital twin technology – creating highly detailed, patient specific virtual replicas of the human body that accurately represent the anatomical, physiological and pathological characteristics of a given individual. These digital twins can be combined with artificial intelligence (AI) to allow surgeons to practice procedures, predict complications and refine their approach to surgery in a completely safe environment. Using digital twins with AI has the potential to make surgery safer, more individualized and better in terms of patient outcomes [82].

Global and Low-Resource Settings

Artificial Intelligence (AI) has the potential to reduce disparities in global surgery and to address the global surgical workforce shortage. AI-based decision-support tools, real-time remote surgical assistance, image analysis and automated diagnostics may facilitate access to surgical expertise where it is needed most, such as in low resource settings where specialized surgical skills are in short supply. The use of non-specialized health workers combined with real-time remote expert consultation through technology has the potential to increase surgical capacity in low resource settings and improve patient outcomes worldwide [83].

That potential can only be fully realized if practical concerns are thoughtfully addressed, such as ensuring that the infrastructure exists for the technology, that affordable implementation pathways are developed, and that the technology is properly adapted to local healthcare settings and realities.

Research Priorities

Future research should involve multicentre prospective studies evaluating the real-world clinical impact of AI-assisted surgical techniques, with a focus on patient relevant outcomes. There is a need for standardized reporting guidelines and methodologies for validation of AI algorithms [66].

Realisation of the benefits of the new technologies in safe, efficient and equitable surgical practice will depend on long-term collaborative effort among surgeons, data scientists, engineers, ethicists, policy makers and, most importantly, patients [81].

Barriers to Implementation of Artificial Intelligence in Surgery

Despite the great promise of artificial intelligence in surgery, significant barriers and limitations remain to its widespread adoption. The main barriers to widespread clinical adoption of AI in surgery are summarized in Figure 4.

From a technical perspective, most contemporary AI-based systems are highly specialized to a specific task and show reduced robustness when faced with real clinical variability. Small differences in patient anatomy, slight differences in surgical technique, unanticipated events during the procedure or changes in clinical context can have a significant impact on the performance of these systems. What works well in a constrained and highly controlled validation setting often fails to translate into the uncontrolled, dynamic, and variable environment of the operating room (OR), limiting both the operational flexibility and the dependability of the technology [84,85].

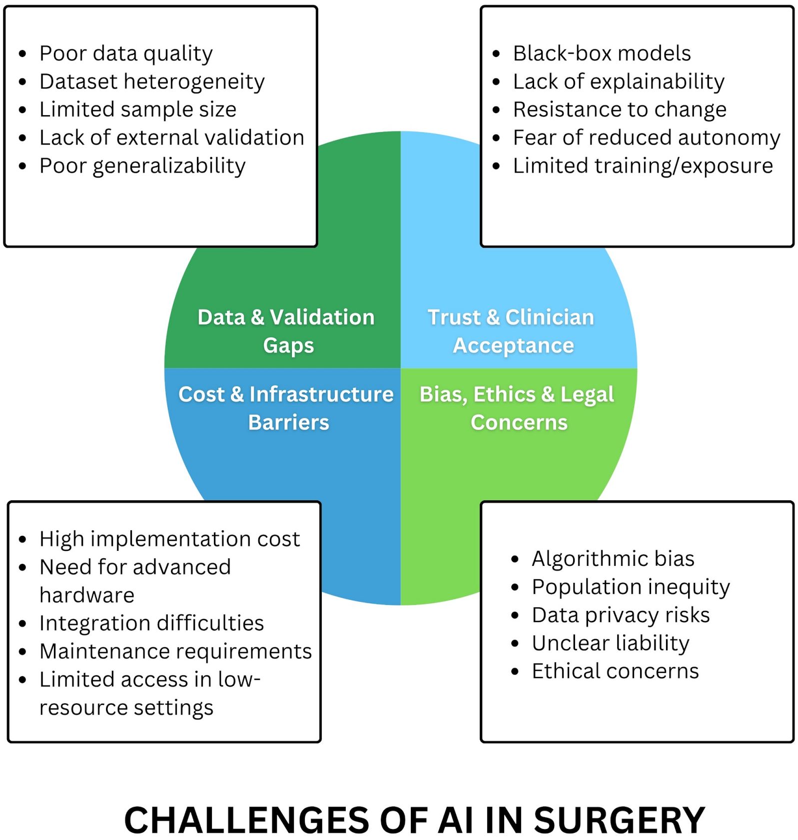
One of the most persistent and significant challenges remains data quality. Achieving the promised benefits of surgical AI requires that a large quantity of high quality and accurately annotated data be available. Operation data are often characterized by significant heterogeneity, lack of complete information, poor documentation and inconsistent data capture. Variability in image acquisition protocols, video format, such as resolutions and frame rates, light exposure and annotation practices all add sources of data noise and potential bias that impact model performance and reliability. The paucity of quality data from a broad spectrum of populations, such as those in low-resource settings and from diverse ethnicities also remains a major concern related to addressing issues of fairness and validity of models in validation sets [64,84,85].

Other limitations in the adoption of AI surgical systems are the cost of the implementation and absence of necessary infrastructure. Most of the AI applications require a lot of computational power and they need to be connected to the hospital information system. They also need to be fully integrated into the workflow of the surgeons, and the model needs to be updated and retrained from time to time, which also requires a highly competent IT department. This will likely be unaffordable for a lot of small hospitals, community clinics and low resource health care settings, potentially creating a new digital divide in surgical health care [7,83,85].

Finally, a hurdle of significant relevance is the issue of clinician acceptance. As already discussed, surgeons have well-reasoned concerns regarding the use of AI in surgical care. Valid concerns exist concerning the impact of automation on workflow, loss of professional autonomy, medicolegal accountability, and having faith in a "black-box" decision support system that is lacking in transparency and justification of the advised action. Mitigating this hurdle would obviously require more than the ability of the technology to accurately assess what actions are needed in a given situation at the time it is required, but also a range of factors, including but not limited to: a clinically significant contribution to surgical care, a clear understanding of the system design with a transparent and justifiable decision-making process, adequate incorporation into established surgical protocols and effective educational and training strategies for the operating room team, as well as strong institutional endorsement [7,86,87].

Improving the performance of current systems in these key areas is therefore a pressing priority if surgical AI is to fulfill its promised potential and safely and fairly transform modern surgical care.

Figure 4: Key Challenges in the Clinical Implementation of Artificial Intelligence in Surgery



Limitations of this Review

This study is a narrative review and as such a systematic literature search with predefined inclusion/exclusion criteria was not conducted, and no formal quality assessment or risk of bias evaluation of individual studies was performed. Consequently, the potential for selection bias exists and the reproducibility of our findings is limited. Additionally, the authors' perspective, primarily from clinical surgery, may have influenced the selection and interpretation of studies. Given the rapid pace of development in AI surgery research, some cited studies may be supplemented or contextualized by newer evidence in the near future. Despite these limitations, we have attempted to provide a balanced and critical synthesis of the current literature. Readers should therefore interpret the findings with appropriate caution.

Conclusion

Artificial intelligence is rapidly transforming perioperative outcomes in surgical practice through the provision of powerful tools that enhance surgical precision, intraoperative decision making, efficiency and postoperative outcomes. AI has demonstrated considerable ability to augment surgical care and education, from preoperative risk stratification and planning to intraoperative guidance and postoperative monitoring. However, several limitations to its application exist. These include issues with clinician acceptance as well as low confidence in its clinical applicability due to methodological constraints in its evidence base. Other persisting challenges are data quality concerns, regulatory hurdles, algorithm bias, cost of implementation and lack of infrastructure especially in low resource settings.

Looking to the future, the development of transparent, reliable systems with robust clinical validation, ethical implementation and interdisciplinary collaboration, represents the trajectory of AI in surgery, rather than the development of systems to replace human expertise. A more personalized, precise and equitable era of surgical care can be realised with careful integration of AI in surgery.

Acknowledgements

Author Contributions: Conceptualization, P. J.; methodology, P. J., C. N. and N. E.; validation, P. J., C. N. and N. E.; investigation, P. J., C. N. and N. E. writing – original draft preparation, P. J., C. N. and N. E.; writing – review and editing, P. J., C. N. and N. E.; Supervision, P. J. All authors have read and agreed to the published version of the manuscript.

Disclosure of AI Use: AI-assisted language editing tools were used solely to improve grammar, spelling, and clarity of the manuscript.

Funding: No funding was received.

References

1. Wu J, Zhang H, Shao J, Chen D, Xue E, Huang S, et al. Healthcare for older adults with multimorbidity: a scoping review of reviews. Clin Interv Aging. 2023;18:1723–1735. doi:10.2147/CIA.S425576
2. Alharbi S, Aljohani B, Elmasry L, Baldovino F, Raviz K, Altowairqi L, et al. Reduction of hospital bed cost for inpatient overstay through optimisation of patient flow. BMJ Open Qual. 2023;12(2):e002142. doi:10.1136/bmjoq-2022-002142
3. Batiha AM. Cost-Reduction Strategies and Patient Care Quality: Insights From Hospital Leaders. J Eval Clin Pract. 2026;32(2):e70351. doi:10.1111/jep.70351
4. Loftus TJ, Tighe PJ, Filiberto AC, Efron PA, Brakenridge SC, Mohr AM, et al. Artificial Intelligence and Surgical Decision-making. JAMA Surg. 2020;155(2):148–158. doi:10.1001/jamasurg.2019.4917
5. Udyavar R, Cornwell EE 3rd, Havens JM, Hashmi ZG, Cooper Z, Askari R, et al. Surgeon-driven variability in emergency general surgery outcomes: Does it matter who is on call? Surgery. 2018;164(5):1109–1116. doi:10.1016/j.surg.2018.07.008
6. Reijmerink IM, van der Laan MJ, Wietasch JKG, Hooft L, Cnossen F. Impact of fatigue in surgeons on performance and patient outcome: systematic review. Br J Surg. 2024;111(1):znad397. doi:10.1093/bjs/znad397
7. Wah JNK. The rise of robotics and AI-assisted surgery in modern healthcare. J Robot Surg. 2025;19:311. doi:10.1007/s11701-025-02485-0
8. Soori M, Arezoo B, Dastres R. Artificial intelligence, machine learning and deep learning in advanced robotics: a review. Cognit Robot. 2023;3:54–70. doi:10.1016/j.cogr.2023.04.001
9. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44–56. doi:10.1038/s41591-018-0300-7
10. Guo C, He Y, Shi Z, Wang L. Artificial intelligence in surgical medicine: a brief review. Ann Med Surg (Lond). 2025;87(4):2180–2186. doi:10.1097/MS9.0000000000003115
11. Iftikhar M, Saqib M, Zareen M, Mumtaz H. Artificial intelligence: revolutionizing robotic surgery: review. Ann Med Surg (Lond). 2024;86(9):5401–5409. doi:10.1097/MS9.0000000000002426
12. Satapathy P, Pradhan KB, Rustagi S, Suresh V, Al-Tameemi ZHA. Application of machine learning in surgery research: current uses and future directions – editorial. Int J Surg. 2023;109(6):1550–1551. doi:10.1097/JS9.0000000000000421
13. Thankappan K. Artificial Intelligence and Machine Learning in Head and Neck Oncology. J Head Neck Physicians Surg. 2022;10(2):117–120. doi:10.4103/jhnps.jhnps_81_22
14. Yang X. Comparison of Supervised Learning and Unsupervised Learning in the Diagnosis of Medical Imaging Diseases. Appl Comput Eng. 2025;145:164–169. doi:10.54254/2755-2721/2025.22239
15. Iqbal T, Elahi A, Wijns W, Shahzad A. Exploring unsupervised machine learning classification methods for physiological stress detection. Front Med Technol. 2022;4:782756. doi:10.3389/fmedt.2022.782756

16. Datta S. Reinforcement learning in surgery. Surgery. 2021;170(1):329–332. doi:10.1016/j.surg.2020.11.040
17. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–444. doi:10.1038/nature14539
18. Mienye ID, Swart TG, Obaido G, Jordan M, Ilono P. Deep Convolutional Neural Networks in Medical Image Analysis: A Review. Information. 2025;16(3):195. doi:10.3390/info16030195
19. Liu L, Blake V, Barman M, Gallego B, Churches T, Kennedy G, et al. Using natural language processing to extract information from clinical text in electronic medical records for populating clinical registries: a systematic review. J Am Med Inform Assoc. 2026;33(2):484–499. doi:10.1093/jamia/ocaf176
20. Carstens M, Vasisht S, Zhang Z, Barbur I, Kolbinger FR. Artificial intelligence for surgical scene understanding: a systematic review and reporting quality meta-analysis. NPJ Digit Med. 2026;9:59. doi:10.1038/s41746-025-02227-4
21. Kasimieh O, Ahmed MM, Othman ZK, Ali I, Hassan MM, Maulion PM, et al. Deep learning in real-time image-guided surgery: a systematic review of applications, methodologies, and clinical relevance. Art Int Surg. 2025;5:557-571. doi:10.20517/ais.2025.92
22. Gharios M, El-Hajj VG, Frisk H, Ohlsson M, Omar A, Edström E, et al. The use of hybrid operating rooms in neurosurgery, advantages, disadvantages, and future perspectives: a systematic review. Acta Neurochir (Wien). 2023;165(9):2343-2358. doi:10.1007/s00701-023-05756-7
23. Knudsen JE, Ghaffar U, Ma R, Hung AJ. Clinical applications of artificial intelligence in robotic surgery. J Robot Surg. 2024;18:102. doi:10.1007/s11701-024-01867-0
24. Ogut E. Artificial Intelligence in Clinical Medicine: Challenges Across Diagnostic Imaging, Clinical Decision Support, Surgery, Pathology, and Drug Discovery. Clin Pract. 2025;15(9):169. doi:10.3390/clinpract15090169
25. van Smeden M, Heinze G, Van Calster B, Asselbergs FW, Vardas PE, Bruining N, et al. Critical appraisal of artificial intelligence-based prediction models for cardiovascular disease. Eur Heart J. 2022;43(31):2921–2930. doi:10.1093/eurheartj/ehac238
26. Charilaou P, Battat R. Machine learning models and over-fitting considerations. World J Gastroenterol. 2022;28(5):605–607. doi:10.3748/wjg.v28.i5.605
27. Pennestrì F, Cabitza F, Picerno N, Banfi G. Sharing reliable information worldwide: healthcare strategies based on artificial intelligence need external validation. Position paper. BMC Med Inform Decis Mak. 2025;25:56. doi:10.1186/s12911-025-02883-2
28. Hulsen T. Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare. AI. 2023;4(3):652–666. doi:10.3390/ai4030034
29. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–118. doi:10.1038/nature21056
30. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60–88. doi:10.1016/j.media.2017.07.005
31. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Werneck Krauss Silva V, Busam KJ, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med. 2019;25(8):1301–1309. doi:10.1038/s41591-019-0508-1
32. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med. 2019;25(6):954–961. doi:10.1038/s41591-019-0447-x
33. Panossian VS, Argandykov D, Arnold SC, Gebran A, Paranjape CN, Hwabejire JO, et al. Validation of Artificial Intelligence-Based POTTER Calculator in Emergency General Surgery Patients Undergoing Laparotomy:

Prospective, Bi-Institutional Study. J Am Coll Surg. 2025;240(3):254-262. doi:10.1097/XCS.0000000000001234
34. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. Ann Surg. 2018;268(1):70–76. doi:10.1097/SLA.0000000000002693
35. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, Motaei A, Madkour M, Pardalos PM, et al. MySurgeryRisk: Development and validation of a machine-learning risk algorithm for major complications and death after surgery. Ann Surg. 2019;269(4):652–662. doi:10.1097/SLA.0000000000002706
36. Lee M, Zakhari V, Shanmugaraj A, Sivasundaram L, Horner NS. Machine learning for predicting outcomes, complications and resource utilisation after hip arthroscopy: a systematic review. Knee Surg Sports Traumatol Arthrosc. 2026; doi:10.1002/ksa.70304
37. Elhaddad M, Hamam S. AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential. Cureus. 2024;16(4):e57728. doi:10.7759/cureus.57728
38. Piazzolla P, Volpi G, Piana A, Checcucci E, Amparore D. Artificial intelligence guidance for 3D augmented reality robotic surgery: When the machine falls the human assistance is still alive. Urol Video J. 2025;25:100307. doi:10.1016/j.urolvj.2024.100307
39. Shen D, Wu G, Suk HI. Deep Learning in Medical Image Analysis. Annu Rev Biomed Eng. 2017;19:221–248. doi:10.1146/annurev-bioeng-071516-044442
40. Kersten-Oertel M, Jannin P, Collins DL. The state of the art of visualization in mixed reality image guided surgery. Comput Med Imaging Graph. 2013;37(2):98–112. doi:10.1016/j.compmedimag.2013.01.009
41. Khor WS, Baker B, Amin K, Chan A, Patel K, Wong J. Augmented and virtual reality in surgery---the digital surgical environment: applications, limitations and legal pitfalls. Ann Transl Med. 2016;4(23):454. doi:10.21037/atm.2016.12.23
42. Longo UG, De Salvatore S, Candela V, Zollo G. Augmented Reality, Virtual Reality and Artificial Intelligence in Orthopedic Surgery: A Systematic Review. Appl Sci (Basel). 2021;11(7):3253. doi:10.3390/app11073253
43. Padoy N. Machine and deep learning for workflow recognition during surgery. Minim Invasive Ther Allied Technol. 2019;28(2):82–90. doi:10.1080/13645706.2019.1584116
44. Twinanda AP, Shehata S, Mutter D, Marescaux J, de Mathelin M, Padoy N. EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos. IEEE Trans Med Imaging. 2017;36(1):86–97. doi:10.1109/TMI.2016.2593957
45. Mascagni P, Alapatt D, Sestini L, Altieri MS, Madani A, Watanabe Y, et al. Computer vision in surgery: from potential to clinical value. NPJ Digit Med. 2022;5(1):163. doi:10.1038/s41746-022-00707-5
46. Wijnberge M, Geerts BF, Hol L, Lemmers N, Mulder MP, Berge P, et al. Effect of a Machine Learning-Derived Early Warning System for Intraoperative Hypotension vs Standard Care on Depth and Duration of Intraoperative Hypotension During Elective Noncardiac Surgery: The HYPE Randomized Clinical Trial. JAMA. 2020;323(11):1052-1060. doi:10.1001/jama.2020.0592
47. Golany T, Aides A, Freedman D, Rabani N, Liu Y, Rivlin E, et al. Artificial intelligence for phase recognition in complex laparoscopic cholecystectomy. Surg Endosc. 2022;36(12):9215-9223. doi:10.1007/s00464-022-09405-5
48. Shin Y. Artificial Intelligence-Powered Quality Assurance: Transforming Diagnostics, Surgery, and Patient Care—Innovations, Limitations, and Future Directions. Life (Basel). 2025;15(4):654. doi:10.3390/life15040654
49. Xiao X, Wang X, Meng B, Pan X, Zhao H. Comparison of robotic AI-assisted and manual pedicle screw fixation for treating thoracolumbar fractures: a retrospective controlled trial. Front Bioeng Biotechnol.

2025;13:1491775. doi:10.3389/fbioe.2025.1491775

50. Golas SB, Shibahara T, Agboola S, Otaki H, Sato J, Nakae T, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. BMC Med Inform Decis Mak. 2018;18(1):44. doi:10.1186/s12911-018-0620-z
51. Heard JR, Deo A, Ghaffar U, et al. AI model predicts patient outcomes from surgical gestures and provides insights into explainability. npj Digit Surg. 2026;1:4. doi:10.1038/s44484-025-00006-y
52. Meyer A, Zverinski D, Pfahringer B, Kempfert J, Kuehne T, et al. Machine learning for real-time prediction of complications in critical care: a retrospective study. Lancet Respir Med. 2018;6(12):905–914. doi:10.1016/S2213-2600(18)30300-X
53. Balian J, Cho NY, Vadlakonda A, Choinski K, Fazl Alizadeh M, Lyou Y, et al. Failure to rescue following emergency general surgery: A national analysis. Surg Open Sci. 2024;20:77–81. doi:10.1016/j.sopen.2024.05.013
54. Kehlet H, Wilmore DW. Evidence-based surgical care and the evolution of fast-track surgery. Ann Surg. 2008;248(2):189–198. doi:10.1097/SLA.0b013e31817f2c1a
55. Zain Z, Almadhoun MKI, Alsadoun L, Bokhari SF. Leveraging Artificial Intelligence and Machine Learning to Optimize Enhanced Recovery After Surgery (ERAS) Protocols. Cureus. 2024;16(3):e56668. doi:10.7759/cureus.56668
56. Khan MM, Shah N. AI-driven wearable sensors for postoperative monitoring in surgical patients: A systematic review. Comput Biol Med. 2025;196(Pt B):110783. doi:10.1016/j.compbiomed.2025.110783
57. Yanagida Y, Takenaka S, Kitaguchi D, Hamano S, Tanaka A, Mitarai H, et al. Surgical skill assessment using an AI-based surgical phase recognition model for laparoscopic cholecystectomy. Surg Endosc. 2025;39(8):5018–5026. doi:10.1007/s00464-025-11903-1
58. Escobar-Castillejos D, Barrera-Animas AY, Noguez J, Cárdenas-Robledo LA, Rosas-Fernández JB. Transforming Surgical Training With AI Techniques for Training, Assessment, and Evaluation: Scoping Review. J Med Internet Res. 2025;27:e58966. doi:10.2196/58966
59. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg. 2002;236(4):458–464; discussion 463–464. doi:10.1097/00000658-200210000-00008
60. Lakshmi AA, Vijayaraj A, Sruthi Shree CK, Kumar SR, Santhosh V, Tharun Kumar S. Artificial intelligence enhanced virtual reality surgery training. In: 2025 2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE); 2025; Chennai, India. p. 1–7. doi:10.1109/RMKMATE64874.2025.11042515
61. Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. Acad Emerg Med. 2008;15(11):988–994. doi:10.1111/j.1553-2712.2008.00227.x
62. Aggarwal R, Mytton OT, Derbrew M, Hananel D, Heydenburg M, Issenberg B, et al. Training and simulation for patient safety. Qual Saf Health Care. 2010;19(Suppl 2):i34–i43. doi:10.1136/qshc.2009.038562
63. Yilmaz R, Alsayegh A, Bakhaidar M, Fazlollahi AM, Abou Hamdan N, Tee T, et al. Combining real-time AI and in-person expert instruction in simulated surgical skills training – Randomized crossover trial. npj Artif Intell. 2025;1:36. doi:10.1038/s44387-025-00032-8
64. Reinke A, Li ZO, Tizabi MD, André P, Knopp M, Rother MM, et al. Current validation practice undermines surgical AI development. arXiv [Preprint]. 2025 Nov 6:arXiv:2511.03769. doi:10.48550/arXiv.2511.03769
65. Lin JC, Jain B, Iyer JM, et al. Benefit-risk reporting for FDA-cleared artificial intelligence–enabled medical devices. JAMA

Health Forum. 2025;6(9):e253351. doi:10.1001/jamahealthforum.2025.3351
66. Lee B, Jay M, Fox H, Padley J, Dai T, Levin AS. FDA-cleared artificial intelligence medical devices in orthopaedic surgery. J Am Acad Orthop Surg Glob Res Rev. 2026;10(2):e25.00170. doi:10.5435/JAAOSGlobal-D-25-00170
67. Eidlisz J, Ajmal E, Awan B. FDA-approved artificial intelligence and machine learning–enabled medical devices in general surgery: A review based on regulatory panel classification. Surgery. 2026;190:109846. doi:10.1016/j.surg.2025.109846
68. Dhawan R, Shauly O, Shay D, Brooks K, Losken A. Growth in FDA-Approved Artificial Intelligence Devices in Plastic Surgery: A Key Look Into the Future. Aesthet Surg J. 2025;45(1):108-111. doi:10.1093/asj/sjae209
69. Sivakumar R, Lue B, Kundu S. FDA approval of artificial intelligence and machine learning devices in radiology: a systematic review. JAMA Netw Open. 2025;8(11):e2542338. doi:10.1001/jamanetworkopen.2025.42338
70. Clark P, Kim J, Aphinyanaphongs Y. Marketing and US Food and Drug Administration clearance of artificial intelligence and machine learning–enabled software in and as medical devices: a systematic review. JAMA Netw Open. 2023;6(7):e2321792. doi:10.1001/jamanetworkopen.2023.21792
71. Windecker D, Baj G, Shiri I, Mohammadi Kazaj P, Kaesmacher J, Gräni C, et al. Generalizability of FDA-approved artificial intelligence–enabled medical devices for clinical use. JAMA Netw Open. 2025;8(4):e258052. doi:10.1001/jamanetworkopen.2025.8052
72. Mohammed ZK. Explainable AI in Health Care: Trust and Transparency in AI-Powered Medical Diagnosis. In: Domínguez-Morales M, Luna-Perejón F, editors. The Latest Advances in the Field of Intelligent Systems. Rijeka: IntechOpen; 2025. doi:10.5772/intechopen.1011279
73. Arjomandi Rad A, Vardanyan R, Athanasiou T, Maessen J, Sardari Nia P. The ethical considerations of integrating artificial intelligence into surgery: a review. Interdiscip CardioVasc Thorac Surg. 2025;40(3):ivae192. doi:10.1093/icvts/ivae192
74. Barrios PA, Meredyth N, Morris R, Haines K, Kim GJ, Peterson CY. Ethical considerations in the deployment of artificial intelligence in surgery. J Surg Res. 2025;315:268-274. doi:10.1016/j.jss.2025.08.019
75. Mittermaier M, Raza MM, Kvedar JC. Bias in AI-based models for medical applications: challenges and mitigation strategies. npj Digit Med. 2023;6:113. doi:10.1038/s41746-023-00858-z
76. Johnson-Mann CN, Loftus TJ, Bihorac A. Equity and Artificial Intelligence in Surgical Care. JAMA Surg. 2021;156(6):509–510. doi:10.1001/jamasurg.2020.7208
77. Ferreres AR. Ethical aspects of artificial intelligence in general surgical practice. Rev Col Bras Cir. 2024;51:e20243762EDIT01. doi:10.1590/0100-6991e-20243762EDIT01-en
78. Pressman SM, Borna S, Gomez-Cabello CA, Haider SA, Haider C, Forte AJ. AI and Ethics: A Systematic Review of the Ethical Considerations of Large Language Model Use in Surgery Research. Healthcare. 2024;12(8):825. doi:10.3390/healthcare12080825
79. Schmidgall S, Opfermann JD, Kim JW, Krieger A. Will your next surgeon be a robot? Autonomy and AI in robotic surgery. Sci Robot. 2025;10(104):eadt0187. doi:10.1126/scirobotics.adt0187
80. Gumbs AA, Croner R, Abu-Hilal M, Bannone E, Ishizawa T, Spolverato G, et al. Surgomics and the Artificial intelligence, Radiomics, Genomics, Oncopathomics and Surgomics (AiRGOS) Project. Art Int Surg. 2023;3(3):180–185. doi:10.20517/ais.2023.24
81. Gumbs AA, Diana M, Rawicz-Pruszyński K, Spolverato G, Abu-Hilal M, Frigerio I, et al. AIONS Consensus Conference on Definitions of Artificial Intelligence Surgery, Surgomics and Robotics. Art Int Surg. 2026;6:98–113. doi:10.20517/ais.2025.113

82. Yelchuri H, Singh DK, Gnani NK, Prabhakar TV, Singh C. RoboTwin: A Robotic Teleoperation Framework Using Digital Twins. arXiv:2506.01027 [Preprint]. 2025 Jun 1 [cited 2026 Mar 22]. Available from: https://arxiv.org/abs/2506.01027. doi:10.48550/arXiv.2506.01027
83. Kewalramani D, Loftus TJ, Mayol J, Narayan M. Artificial intelligence in surgery: a global balancing act. Br J Surg. 2024;111(3):znae062. doi:10.1093/bjs/znae062
84. Kenig N, Monton Echeverria J, Muntaner Vives A. Artificial Intelligence in Surgery: A Systematic Review of Use and Validation. J Clin Med. 2024;13(23):7108. doi:10.3390/jcm13237108
85. Ruiz NI, Cardona Salazar I, Naranjo Palacio L, et al. Accuracy and reliability of artificial intelligence in surgical decision-making: a literature review. Cureus. 2025;17(10):e95337. doi:10.7759/cureus.95337
86. Arboit L, Schneider DN, Collins T, Hashimoto DA, Perretta S, Dallemagne B, et al. Surgeons Awareness, Expectations, and Involvement with Artificial Intelligence: a Survey Pre and Post the GPT Era. arXiv:2506.08258 [Preprint]. 2025 Jun 9 [cited 2026 Mar 22]. Available from: https://arxiv.org/abs/2506.08258. doi:10.48550/arXiv.2506.08258
87. Brandenburg JM, Müller-Stich BP, Wagner M, van der Schaar M. Can surgeons trust AI? Perspectives on machine learning in surgery and the importance of eXplainable Artificial Intelligence (XAI). Langenbecks Arch Surg. 2025;410:53. doi:10.1007/s00423-025-03626-7