

Artificial Intelligence in Surgical Training: A Comprehensive Update on Simulation, Assessment, and Emerging Innovations

Rushabh Shah^{1,*}, Yash Verma¹, Daniyal Ashraf², Umar Shifas², Kai Yuen Wong¹

¹Department of Plastic and Reconstructive Surgery, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, CB2 0QQ Cambridge, UK

²School of Clinical Medicine, University of Cambridge, CB2 0SP Cambridge, UK

*Correspondence: rushabh.shah1@nhs.net (Rushabh Shah)

Academic Editor: John Sandars

Submitted: 2 January 2026 Revised: 10 February 2026 Accepted: 3 April 2026 Published: 22 June 2026

Abstract

Artificial intelligence (AI) is reshaping surgical education. This article summarises recent developments across simulation, assessment and recruitment. It also examines the barriers and governance challenges of implementation, and considers emerging technologies such as digital twins, generative AI and intra-operative decision-support that are likely to shape the future of surgical training.

Keywords: artificial intelligence; surgical training; simulation; skill assessment; machine learning

1. Introduction

Surgical training has moved away from the traditional apprenticeship model towards competency-based and simulation-supported curricula [1,2]. This shift reflects both opportunity and necessity. Various factors, such as an ageing population and increasing burden of surgical disease versus the need for high-quality care within restricted theatre time, have challenged conventional experiential learning. In the United Kingdom, the introduction of the European Working Time Directive significantly changed the delivery of surgical training with increased emphasis on accountability and transparency [3]. The duration of surgical training is now shorter with increased service provision and significantly reduced training time. Thus, there has been a need to change the way surgical training is delivered [4].

Despite subsequent reforms, contemporary studies continue to highlight persistent constraints on operative exposure, increasing service pressures, and the need for more efficient and structured approaches to surgical training [5,6].

Simulation centres, skills laboratories, and structured assessment frameworks such as the Objective Structured Assessment of Technical Skills (OSATS), alongside workplace-based assessments within the Intercollegiate Surgical Curriculum Programme (ISCP), have brought greater consistency and objectivity to surgical training [1]. However, challenges to the effectiveness of workplace-based assessments include adequate time, adapting to different learning needs and effective feedback. In this context, artificial intelligence (AI) offers unique potential. Through machine learning, computer vision, and natural language processing, AI can analyse complex procedural data, quantify performance, and deliver real-time, standardised feedback at scale [7,8]. AI operationalises this by con-

verting rich multimodal data (video, kinematics, force) into interpretable metrics and targeted feedback.

Although AI has already demonstrated transformative value in diagnostic imaging, pathology, and clinical decision support, its integration into surgical education has been more measured [9]. This caution arises from various factors such as algorithmic transparency, governance, inherent difficulty of modelling human tissue handling and intra-operative decision-making [10,11]. Nevertheless, early applications of AI in simulation, automated skill assessment, and intelligent tutoring have shown promising results, delivering consistent, objective feedback that augments rather than replaces expert mentorship [12,13].

This article is intended as a narrative education and training update rather than a formal systematic review. The literature was identified through targeted searches of PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) and Google Scholar (<https://scholar.google.com/>), focusing on English-language publications from approximately 2015 to 2025 relating to artificial intelligence in surgical education, simulation, assessment, and training governance. Key search terms included combinations of “artificial intelligence”, “machine learning”, “surgical training”, “simulation”, “skill assessment”, and “education”. Reference lists of relevant reviews and seminal articles were also screened to identify additional pertinent studies. The aim was to provide a balanced, representative overview of established and emerging applications rather than an exhaustive synthesis of all available evidence.

2. Understanding AI in Surgical Training

AI refers to computer systems that can analyse information, recognise patterns, and make decisions or recommendations with varying degrees of autonomy, often emu-



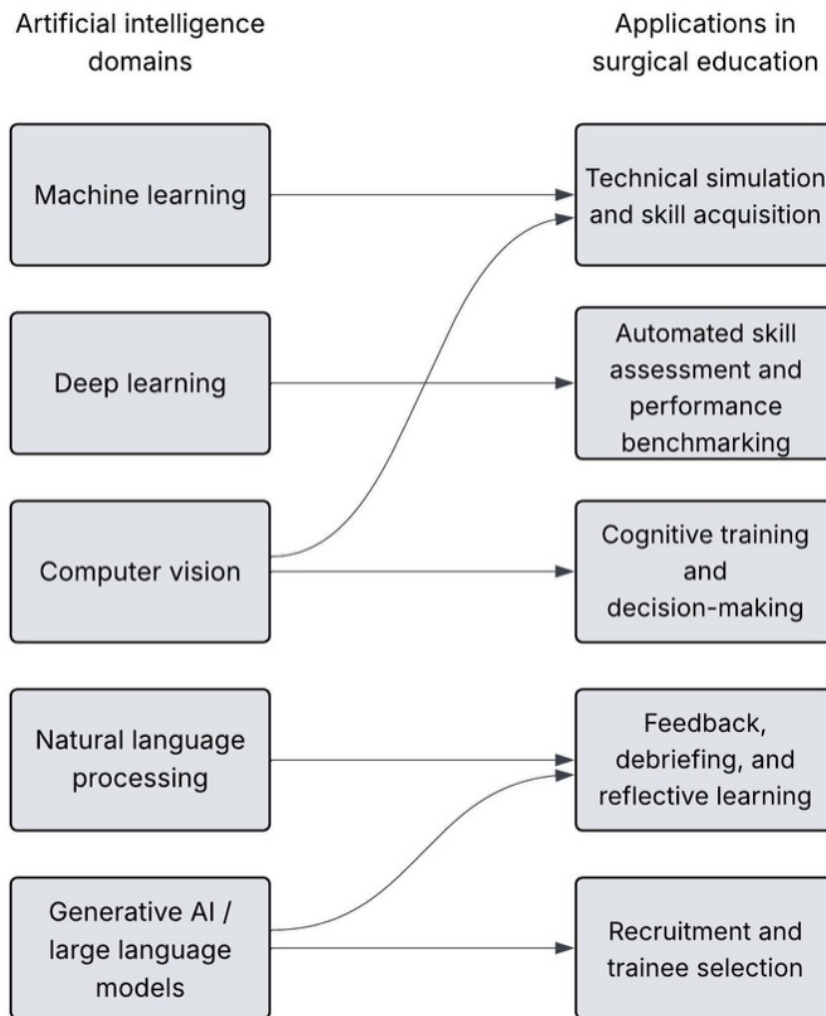


Fig. 1. Core artificial intelligence (AI) domains and their applications in surgical education. Figure created using Lucidchart (Lucid Software Inc., South Jordan, UT, USA).

lating aspects of human reasoning and learning. In surgical education, AI is increasingly used to analyse operative performance, enhance simulation-based learning, and provide structured, objective feedback in real-time.

2.1 Components of AI in Surgical Training

Several key domains of AI are particularly relevant to surgical training (Fig. 1). These areas and their specific applications are summarised in Table 1 (Ref. [6,14–18]).

2.2 Algorithmic Feedback and Explainability in AI Models

AI models can be broadly categorised by how they process data and communicate their reasoning. Black-box AI models generate outputs without revealing their reasoning, limiting interpretability and user trust. In contrast, explainable AI models reveal the variables influencing a decision [10,19]. For example, real-world applications, such as the Virtual Operative Assistant, provide interpretable dashboards that map instrument motion and feedback metrics to specific surgical actions [13]. Fig. 2 illustrates how such

models enhance educational value by linking algorithmic feedback to observable surgical resident actions.

Surveys suggest that residents perceive interpretable feedback as more credible and motivating than opaque numeric scores [20]. Indeed, this transparency supports reflection, accountability, and sustained skill development.

2.3 AI Applications for Surgical Training

AI is now integrated across diverse surgical training settings, each characterised by distinct data sources, fidelity, and analytical potential [21,22]. Virtual and augmented-reality simulators recreate operative environments in which algorithms can quantify motion metrics and adjust task difficulty in real-time. Robotic and minimally invasive platforms capture high-resolution kinematic and haptic feedback data ideal for machine-learning analysis, enabling continuous performance monitoring. In contrast, open surgical training increasingly relies on wearable sensors, optical tracking, and video capture to assess ergonomics and manual dexterity. These complementary en-

Table 1. Core domains of AI and their applications in surgical education.

AI domain	Function	Application in surgical training	Example references
Machine learning	Learns patterns from structured data to classify or predict outcomes	Gesture recognition, skill scoring, performance benchmarking	Azari et al., 2019 [14]
Deep learning	Uses neural networks to analyse complex visual/motion data	Automatic skill grading from endoscopic video or motion sensors	Ward et al., 2021 [16]
Computer vision	Enables interpretation of images and videos	Phase recognition, instrument tracking, motion analysis	Kitaguchi et al., 2021 [18]; Igaki et al., 2023 [15]
Natural language processing	Processes and generates human language	Automated feedback, viva simulation, report analysis	Leon et al., 2025 [6]
Generative AI/ Large language models	Produces new data (text, images, dialogue) from learned patterns	Interactive tutoring, personalised debriefing, training chatbots	Mason et al., 2025 [17]

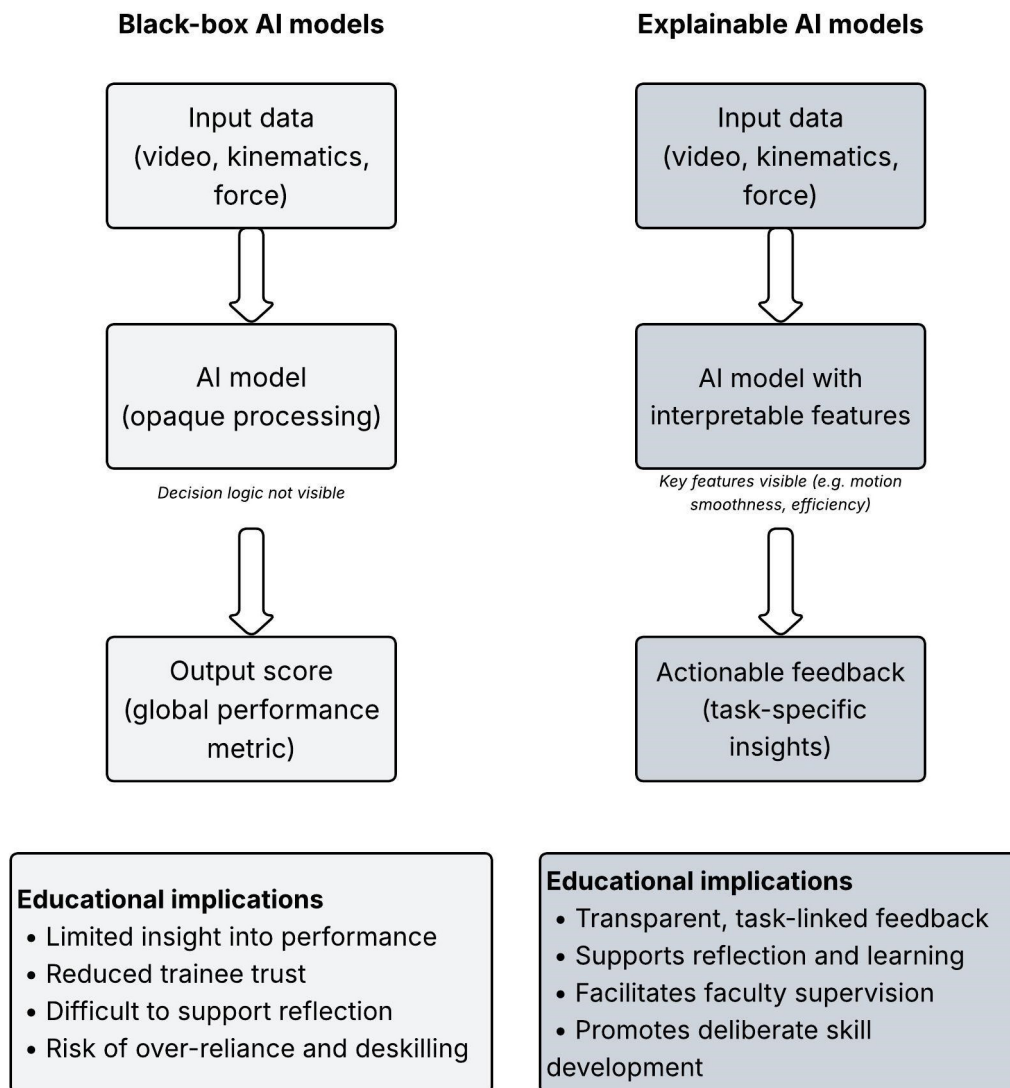


Fig. 2. Comparison between black-box and explainable AI models in surgical education. Figure created using Lucidchart (Lucid Software Inc., South Jordan, UT, USA).

vironments collectively generate the multimodal datasets that underpin AI-based feedback and assessment systems.

While the technical foundations and applications of AI are increasingly well described, successful adoption within surgical education requires careful translation from research-

Table 2. AI applications in surgical training: data sources, educational function, evidence maturity, and limitations.

AI application area	Surgical context	Primary data source	Educational function	Evidence maturity	Key limitations	Key references
Technical skill assessment	Robotic surgery	Console telemetry, operative video	Automated skill scoring, formative feedback	Advanced	Proprietary platforms; limited cross-system generalisability	Moglia et al., 2021 [24]; Pedrett et al., 2023 [8]; Igaki et al., 2023 [15]
Technical simulation	Laparoscopic surgery	Endoscopic video, motion tracking	Real-time feedback, proficiency-based training	Moderate	Variable camera views; annotation burden	Kirubarajan et al., 2022 [7]; Cacciamani et al., 2021 [26]; Yilmaz et al., 2024 [12]
Open surgery assessment	Open procedures	Optical tracking, wearable sensors, force data	Motion analysis, ergonomic feedback	Early	Limited datasets; lack of standardisation	Hamza et al., 2025 [23]; Singh et al., 2024 [27]
Cognitive and decision-making simulation	High-risk surgery (e.g., neurosurgery, cardiac surgery)	Imaging, physiological data, scenario logs	Decision-making, crisis management training	Moderate	Scenario generalisability; complex validation	Mirchi et al., 2020 [13]; Loftus et al., 2020 [25]; Leon et al., 2025 [6]
Video-based surgical analytics	Multi-specialty	Operative video	Phase recognition, benchmarking, coaching	Moderate	Labour-intensive annotation; context dependence	Kitaguchi et al., 2021 [18]; Ward et al., 2021 [16]

Evidence maturity reflects the level of educational evaluation reported in the literature (early: feasibility or single-centre studies; moderate: construct validity or novice–expert discrimination; advanced: external validation or widespread educational deployment).

Table 3. Examples of AI-supported cognitive and decision-making simulation in surgical education.

Context/training scenario	AI technology/model	Primary function	Educational application
Resuscitation and crisis management	Reinforcement-learning algorithms integrated with high-fidelity manikins	Analyses real-time physiological data and trainee responses to guide optimal interventions	Dynamic resuscitation drills (e.g., massive haemorrhage, cardiac arrest) with adaptive feedback on timing and prioritisation
Peri-operative decision-making	Probabilistic/predictive AI models	Estimates intra-operative risk and suggests management pathways	Teaching intra-operative judgement, anticipating complications, and balancing risk
Diagnostic reasoning	Deep-learning models trained on imaging and case data	Generates differential diagnoses and outcome probabilities	Case-based simulation to train analytical reasoning and pattern recognition
Team communication and cognitive load	Natural-language processing (NLP) and sentiment analysis	Analyses verbal exchanges and response latency during simulated crises	Feedback on teamwork, closed-loop communication, and stress response
Interactive debriefing and reflection	Large-language-model (LLM)-based conversational agents (e.g., generative pre-trained transformer 5 (GPT-5))	Transcribes, summarises, and contextualises trainer feedback	Post-simulation debriefs and viva-style questioning integrating cognitive and communication skills

ch development to real-world training environments. Fig. 3 outlines a staged implementation pathway for AI integration into surgical education, highlighting key decision points, stakeholder roles, and the importance of continuous outcome monitoring.

3. AI in Surgical Simulation

Simulation is now an essential component of surgical education, providing a safe environment for deliberate practice, error correction, and assessment. This encompasses a range of modalities, including virtual and augmented reality systems, mixed-reality and haptic interfaces, and physical or “organic” models that replicate operative anatomy to varying degrees of fidelity. In contrast to traditional simulations, AI platforms not only facilitate adaptations to learner performance but may more consistently quantify behaviour automatically, alongside providing real-time targeted feedback. Table 2 (Ref. [6–8,12,13,15,16,18,23–27]) summarises key applications of AI in surgical training, highlighting surgical context, primary data sources, educational function, evidence maturity, and current limitations.

3.1 Technical Simulation and Skill Acquisition

Over the past five years, interest in AI-enhanced surgical simulation has increased substantially, with over a dozen primary studies and several systematic reviews evaluating feasibility, accuracy, and educational value [5,7,8]. In a randomised controlled trial of 97 trainees performing simulated brain tumour resection, Yilmaz et al. [12] found that participants receiving AI feedback outperformed the expert-instructed group ($p = 0.005$) while achieving significantly higher OSATS scores than those without feedback ($p < 0.001$). Similarly, Fazlollahi et al. [28] demonstrated accelerated proficiency in laparoscopic suturing with struc-

tured AI guidance. Systematic reviews confirm the reliability of AI-based feedback, showing moderate-to-high accuracy for skill grading and differentiation between novice and expert performance [7,8].

AI applications in surgical simulation vary by procedural setting, influenced by the type and quality of data available for analysis. Minimally invasive and robotic surgery remain the most studied due to readily accessible kinematic and video inputs, whereas open surgery still faces challenges in data acquisition and standardisation.

The educational advantage of using AI lies in its objectivity and adaptability. Continuous, quantitative feedback eliminates inter-observer variability and supports tailored progression so that trainees neither plateau nor advance prematurely while performing simple tasks [6]. Simulation fidelity is also important. Park et al. [22] demonstrated that combining realistic haptic interfaces with adaptive AI feedback improved operative skill transfer, while Hamza et al. [23] noted that tactile realism influences trainee trust in AI-guided systems. However, the design of such systems is critical. AI-supported simulation should prioritise formative feedback that encourages reflection and independent problem-solving, rather than continuous prescriptive guidance that risks promoting cognitive offloading.

AI platforms should complement, not replace, human mentorship. Hybrid models, where AI delivers continuous formative feedback and instructors focus on higher-order reasoning, offer the optimal balance [21]. Nakawala et al. [29] demonstrated this synergy in a thoracentesis simulator where AI guided technical steps while faculty provided contextual discussion, resulting in the greatest performance and confidence gains. These hybrid frameworks preserve key aspects of mentorship, including clinical judgement, decision-making, and professionalism.

Table 4. Commercial platforms incorporating AI into surgical simulation and assessment.

System	Primary function	AI features	Educational application
Touch surgery	Mobile and web-based surgical simulation	Uses AI to analyse user performance, track progress, and deliver real-time adaptive feedback	Step-by-step procedural training across multiple specialties
Osso VR	Virtual reality (VR) surgical simulation	AI-powered real-time performance metrics and personalised feedback	Immersive simulation for technical skill acquisition in a risk-free environment
Mimic technologies	Robotic skills simulator (da Vinci Skills Simulator)	AI algorithms evaluate dexterity, precision, and efficiency	Objective performance assessment in robotic surgery training
Surgical science (Lap-Sim/EndoSim)	Laparoscopic and endoscopic simulation	Real-time AI feedback and performance tracking	Objective assessment and progress monitoring in minimally invasive surgery
C-SATS (Johnson & Johnson)	Cloud-based video review and assessment	Combines AI analytics with peer review to score surgical performance	Continuous learning and benchmarking across surgical specialties
Theator	Video-based surgical analytics	Computer vision annotates operative phases and benchmarks performance	AI-driven feedback and quality improvement through surgical video analysis

C-SATS, Crowd-Sourced Assessment of Technical Skills.

One of the most promising frontiers is the development of digital twins; dynamic, data-driven models that mirror a real surgical environment or individual learner [30]. In education, digital twins could replicate patient anatomy, procedural flow, and intra-operative responses to create fully personalised simulations. Real-time integration of operative video, motion data, and performance analytics would allow adaptive scenarios that evolve according to trainee decisions, effectively creating a “living simulator” [31]. These systems could eventually predict learning trajectories, highlight emerging skill gaps, and tailor training before deficiencies impact clinical performance.

3.2 Cognitive and Decision-Making Simulation

Beyond technical skill, surgical excellence depends on cognitive abilities such as reasoning, anticipation, and decision-making under pressure. AI now enables structured cognitive simulation that integrates multimodal data such as imaging, vital signs, and operative videos to create dynamic and adaptive scenarios. These systems can model clinical uncertainty, prompt management decisions, and provide outcome-based feedback, supporting skills in crisis management and situational awareness. Table 3 summarises current and emerging examples of AI-supported cognitive and decision-making training across surgical and peri-operative contexts.

Advances in generative AI are also set to transform how surgical knowledge is delivered. Natural-language models can generate interactive case discussions, personalised feedback dialogues, and context-specific teaching materials. When combined with operative datasets, generative systems could summarise surgical videos, extract key learning points, and simulate viva-style questioning, offering scalable mentorship beyond the operating room [6]. However, careful validation is essential to prevent misin-

formation, ensure alignment with accredited curricula, and maintain factual accuracy.

3.3 Commercial Platforms

Several commercial platforms have now incorporated AI into surgical simulation, demonstrating its transition from research to practical education. Table 4 summarises key commercial platforms that have incorporated AI into surgical simulation and assessment, illustrating how these technologies are beginning to move from research validation to practical educational deployment.

Simulation provides a controlled environment for developing technical and cognitive skills. Its educational value, however, depends on how performance is measured and interpreted. As surgical training moves toward competency-based models, there is a growing need for assessment tools that can objectively quantify skill and track progression over time. This underpins the expanding role of AI in surgical assessment.

4. AI in Surgical Assessment

4.1 Assessment Tools

Objective evaluation of surgical performance is central to competency-based training. Traditional frameworks such as the Objective Structured Assessment of Technical Skills (OSATS) and the Global Evaluative Assessment of Robotic Skills (GEARS) rely on expert observation and may be limited by subjectivity, inter-rater variability, and the time burden on faculty [1,2]. AI offers a scalable, data-driven adjunct capable of continuously recording, quantifying, and benchmarking performance from video, motion, and haptic data. Its key advantages include greater objectivity, consistency, and reproducibility. AI also has the ability to visualise individual learning curves and detect latent performance markers such as tremor frequency or motion en-

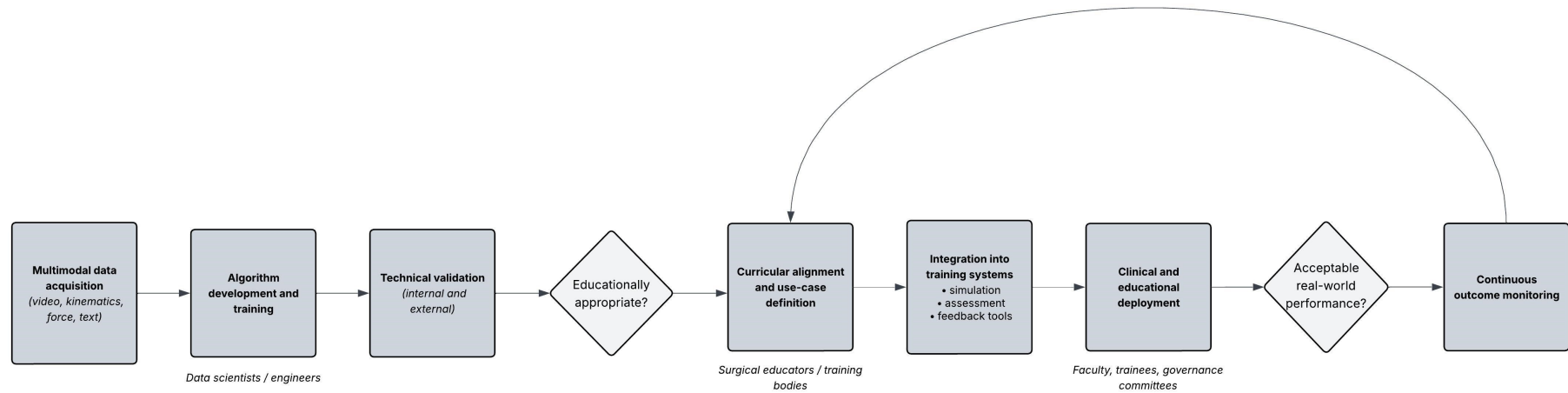


Fig. 3. Conceptual framework illustrating the implementation pathway for AI in surgical education, from multimodal data acquisition and algorithm development through technical validation, curricular alignment, and clinical deployment. Figure created using Lucidchart (Lucid Software Inc., South Jordan, UT, USA).

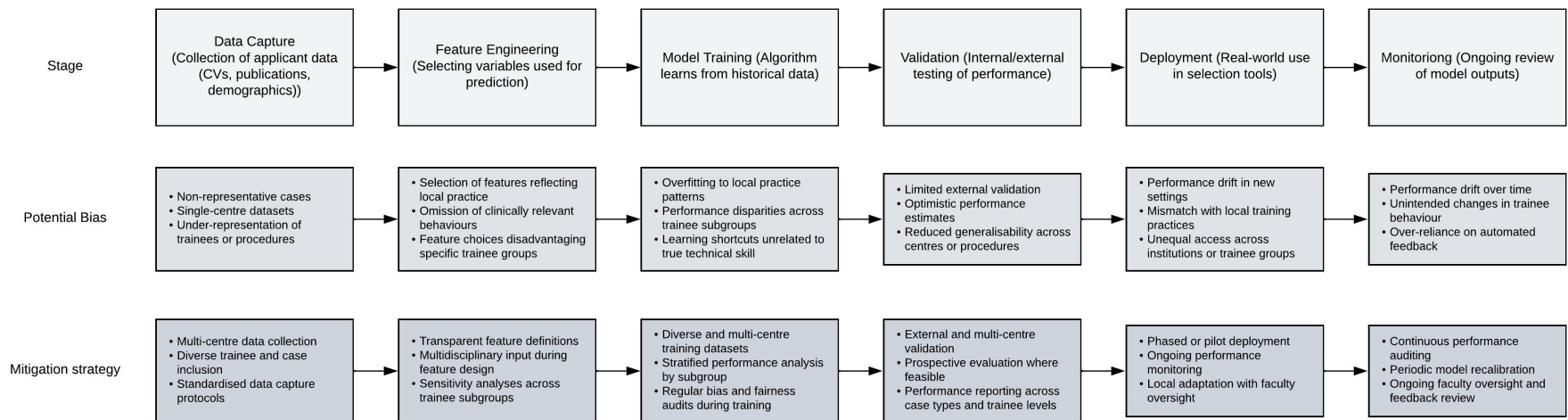


Fig. 4. AI implementation pipeline in surgical education and training, highlighting sources of bias and mitigation strategies. Figure created using Lucidchart (Lucid Software Inc., South Jordan, UT, USA).

tropy, that may capture technical dexterity more accurately than checklist-based assessments [21]. However, important challenges remain. Algorithms trained on small or homogeneous datasets may fail to generalise across instruments, procedures, or institutions. This underscores the need for multi-centre validation, standardised benchmarking frameworks, and transparent models that can be safely interpreted by educators and trainees.

AI-based systems typically use computer vision or kinematic analytics to evaluate operative performance. Computer vision models interpret video to identify gestures, instrument trajectories, and tissue-handling quality, while kinematic systems use robotic telemetry or sensorised instruments to measure path length, velocity, and motion smoothness. Early studies demonstrated that such algorithms can reliably distinguish expert from novice surgeons. Azari et al. [14] reported that an automated OS-ATS surrogate achieved a correlation coefficient of $R^2 \approx 0.81$, exceeding human inter-rater reliability. On the other hand, Fard et al. [32] classified expertise across suturing and knot-tying tasks with 82–90% accuracy. Subsequent deep-learning approaches have extended these findings. Yanagida et al. [33] achieved over 80% accuracy for gesture recognition in laparoscopic cholecystectomy, and Igaki et al. [15] developed a “standardised surgical field” confidence score in colorectal surgery that correlated strongly with expert ratings ($\rho = 0.81$) and detected suboptimal cases with over 90% sensitivity.

Kirubarajan et al. [7] reviewed 49 studies showing consistent success in automated skill discrimination, while Pedrett et al. [8] evaluated 50 minimally invasive models and found high construct validity but limited external validation, with only 8% tested outside their original datasets. Over the past five years, over 50 papers and several systematic reviews have explored automated surgical skill assessment across general surgery, urology, neurosurgery, and cardiovascular surgery [5,7,8,23]. Most research focuses on minimally invasive and robotic platforms, where rich kinematic and video data are readily available, whereas applications in open surgery remain comparatively limited.

4.2 Feedback and Application

AI-driven assessment can provide real-time, interpretable feedback that supports both individual and cohort-level learning. By quantifying performance metrics, these systems generate actionable insights that can guide self-directed improvement and targeted faculty feedback. For instance, Kasa et al. [34] evaluated 72 general-surgical trainees performing laparoscopic simulation tasks and found a strong correlation between AI-generated and expert ratings (mean Spearman $r = 0.67$), particularly for end-product quality. Such analytics enable educators to monitor progression longitudinally, identify outliers, and direct tailored support where needed.

Beyond individual performance review, AI systems are now being explored for group-level benchmarking and formative progression monitoring. Several centres have trialled AI-integrated modules, such as the da vinci research kit, which uses robotic telemetry and video data for automated assessment within structured curricula [24]. However, national training bodies have yet to adopt AI-generated assessments for certification purposes, and most systems remain primarily formative under faculty supervision.

Similarly, in assessment settings, AI tools should be used to support reflection and targeted improvement. They should not replace critical judgment. Excessive reliance on automated scoring may risk deskilling and reduced situational awareness.

While early outcomes are encouraging, successful integration of AI-based feedback into surgical education will depend on transparent validation, explainable outputs, and robust ethical standards governing data security, consent, and feedback use. The transition from formative to summative application will require alignment between technological capability and regulatory oversight to ensure fairness, accountability, and educational benefit.

The next stage of development will be real-time human-AI collaboration. Intra-operative decision-support tools are already assisting with anatomical recognition, instrument tracking, and error detection. These same technologies could be repurposed for education, overlaying expert annotations or safety prompts during supervised procedures to create context-aware, in-the-moment learning experiences [35]. As these systems mature, they may enable trainees to observe and internalise expert reasoning in real-time.

AI is increasingly used to assess performance within surgical training programmes. Similar analytic approaches are now being explored earlier in the educational pathway. This includes their potential role in recruitment and trainee selection.

5. AI in Recruitment and Selection

Beyond assessing performance within training, AI is increasingly being explored to support recruitment and trainee selection. These applications aim to enhance efficiency, transparency, and fairness, though they introduce new ethical and practical challenges.

Algorithms can automatically process large applicant pools, extract quantitative metrics such as academic performance, examination results, and publication history, and assist in ranking or shortlisting candidates [25]. However, AI-generated and human-derived rankings may differ substantially. Hassan et al. [36] compared an AI-assisted residency ranking with selections made by programme directors and found only 7% consensus. The AI system weighted objective metrics more heavily, inadvertently favouring candidates from research-intensive institutions. This diver-

gence underscores the potential for both bias and over-standardisation if models are trained on incomplete or historically skewed datasets [11]. As illustrated in Fig. 4, bias can arise at any stage of the recruitment data pipeline, underscoring the need for transparency, validation, and continuous human oversight. Equally, core attributes to surgical training, such as teamwork, professionalism, empathy, and adaptability, remain difficult to measure [6].

In parallel, the increasing availability of generative AI tools introduces additional challenges for recruitment processes. A recent concern is the use of generative AI by applicants themselves. Large language model tools can now draft personal statements, optimise written assessments, and simulate interview answers, creating inequities linked to digital literacy and access [37]. This is a significant concern as many recruitment processes now occur online. Several UK training bodies, such as the General Medical Council and the UK Foundation Programme Office, have issued guidance to candidates discouraging undisclosed use of generative AI in applications and emphasising the importance of structured interviews that assess real-time reasoning [5].

Ethical and legal considerations are integral to the responsible use of AI in recruitment. Black-box screening tools complicate contestability and undermine candidate confidence in the selection process [10]. Institutions must therefore ensure compliance with equal opportunity, data protection, and employment law frameworks [38]. There should also be transparency in how algorithms are utilised within the recruitment process, as well as ongoing governance [39,40].

Ultimately, AI should serve as a decision-support tool rather than a decision-maker. Transparent algorithms, routine bias audits, applicant disclosure policies, and retained human accountability will be essential to ensure that AI-assisted recruitment complements the core values of surgical education - equity, integrity, and professional judgement.

6. Ethical, Legal, and Governance Considerations in AI-Driven Surgical Education

The integration of AI into surgical education offers significant opportunities for objective, personalised learning but also raises essential concerns regarding transparency, fairness, and accountability [41]. A major limitation lies in the datasets on which these systems are trained. Small, homogeneous, or Western-centric data risk propagating bias that disadvantages certain operative techniques, trainee demographics, or institutional contexts. As highlighted by Johnson-Mann et al. [42], most existing surgical datasets originate from high-income countries, restricting model generalisability. Incorporating diverse data from low- and middle-income settings is therefore crucial to avoid embedding structural inequities and to ensure AI-

driven training tools remain equitable, representative, and applicable across the global surgical workforce [11].

Several practical approaches are being used to address concerns around dataset homogeneity in surgical AI. One strategy involves multi-institutional collaboration, where data from different hospitals and training centres are combined to capture wider variation in case mix, technique, and clinical context, thereby improving generalisability [43]. Another approach is federated learning, which allows AI models to be trained across multiple institutions without sharing raw patient data, helping to improve model performance while maintaining data governance and confidentiality [44]. In addition, the development of open, annotated surgical video and kinematic datasets has supported transparent benchmarking and external validation of AI systems used for training and assessment [45]. Together, these approaches offer realistic pathways toward more diverse and representative datasets while remaining consistent with ethical and regulatory requirements. In practice, these data strategies are most effective when embedded within institutional governance structures that provide multidisciplinary oversight, define educator accountability for AI-generated feedback, and ensure transparent consent and audit processes prior to educational deployment.

The growing adoption of commercial AI platforms within surgical training further underscores the importance of ethical oversight. As Varas et al. [46] note, systems such as Theator, Touch Surgery, and Crowd-Sourced Assessment of Technical Skills (C-SATS) now employ automated video analytics and algorithmic scoring at scale, raising critical questions regarding consent, data ownership, and accountability for generated feedback. Without transparent governance and clearly defined responsibilities, algorithmic bias or misclassification could propagate into real-world trainee assessment. These examples highlight the need for proactive regulation and standardised validation frameworks before widespread curricular adoption.

Over-reliance on automation poses another ethical challenge. Continuous AI guidance may encourage “cognitive offloading”, reducing independent reasoning and situational awareness [47]. Studies in endoscopy have shown performance decline when AI assistance is withdrawn, underscoring the need for deliberate AI-free practice and reflective debriefing to preserve manual and cognitive skill [48].

Governance frameworks must evolve in parallel with innovation. Institutions should implement robust policies on data ownership, consent, and security to protect both trainee and patient information [45,49]. Algorithms used for assessment should be used in conjunction with other assessment tools and human oversight, with final accountability resting on educators and training committees [48]. Oversight by dedicated AI ethics committees and alignment with emerging professional guidance—such as from the Royal College of Surgeons of England [50]—can en-

sure transparency, validation, and fairness as these systems move from research to clinical curricula.

Ultimately, the ethical deployment of AI in surgical training must rely on human-centred governance: systems ought to remain transparent, explainable, fair, and reversible. When coupled with responsible oversight, AI can enhance—rather than erode—the moral and professional foundations upon which surgical education is built.

Importantly, the adoption and educational utility of AI are not uniform across surgical specialties. Fields characterised by high procedural standardisation and digitally rich operative platforms, such as robotic urology and minimally invasive surgery, have been early adopters of AI-enabled simulation and automated performance assessment, supported by high-fidelity video and kinematic telemetry [8, 26,34]. In contrast, specialties dominated by open techniques or complex, variable anatomy face greater challenges in data capture, annotation, and model generalisability, limiting the maturity of AI-based assessment tools [23]. Image-intensive and high-risk disciplines, including neurosurgery and cardiac surgery, have prioritised applications in pre-operative planning, cognitive simulation, and decision-support, where validation thresholds and safety requirements are particularly stringent [12,13,25]. These examples are illustrative rather than exhaustive but highlight the need for specialty-sensitive development, validation, and governance frameworks when integrating AI into surgical training [7,16].

The adoption of artificial intelligence in surgical training is influenced not only by technical performance but also by educator and trainee perceptions. Survey-based studies suggest that both groups recognise the potential benefits of AI-enabled simulation and assessment, particularly in improving objectivity, standardisation, and access to feedback. However, concerns persist regarding algorithmic transparency, over-reliance on automated scoring, and the risk of deskilling if AI tools are not appropriately integrated within existing mentorship structures [13,20]. Educators have highlighted uncertainty around accountability for AI-generated feedback and the need for faculty development in AI literacy, while trainees tend to value consistency and immediacy of feedback but express lower trust in opaque systems [6,16]. These findings emphasise that successful implementation requires not only technical validation but also attention to human factors, clear communication of system limitations, and alignment with curricular goals.

7. Conclusion

AI is becoming an integral component of surgical training, offering new ways to analyse performance, structure feedback, and enhance simulation-based learning. Its successful integration, however, depends on transparent algorithms, diverse datasets, and robust governance to ensure fairness and educational value. Human mentorship, professional judgement and contextual decision-making remain

central to training and must continue to frame where and how AI is used. Key questions now concern which tasks should be augmented, who should oversee AI-generated feedback, and how these tools can be implemented across different training contexts. Future developments in surgical education are likely to be shaped by the convergence of AI with immersive technologies and digital connectivity. Integration of AI with extended reality platforms may enable adaptive simulation environments that respond dynamically to trainee performance, supporting personalised rehearsal and targeted skill development. In parallel, advances in low-latency connectivity and remote telementoring offer opportunities for AI-assisted supervision and feedback beyond the physical training centre, extending access to expertise while preserving local oversight. As longitudinal performance data accumulate across simulations and clinical activity, AI systems may increasingly support personalised learning trajectories that adapt training intensity, content, and feedback to individual progress over time. Realising this potential will depend on careful validation, governance, and alignment with educational objectives, ensuring that innovation enhances rather than fragments surgical training pathways. Continued collaboration between educators, trainees, technologists, and regulators will be essential to ensure safe and equitable adoption.

Key Points

- AI is redefining surgical education by enabling objective, data-driven, and personalised learning that enhances rather than replaces traditional mentorship models.
- AI-driven simulation, integrating both technical and cognitive training, provides adaptive, real-time feedback that accelerates skill acquisition and supports tailored progression across varied surgical environments.
- Automated skill and trainee assessment systems demonstrate expert-level accuracy and scalability but require broader validation and standardisation before integration into formal certification frameworks.
- Ethical and governance safeguards - including explainable algorithms, diverse datasets, and transparent oversight - are essential to maintain fairness, accountability, and trust in educational use.
- Emerging technologies such as digital twins, generative AI, and intra-operative decision-support systems will link training with live surgical practice, fostering adaptive and continuous professional learning.

Availability of Data and Materials

Not applicable.

Author Contributions

RS, YV, DA, US and KYW designed the work. RS drafted the manuscript. All authors contributed to the important editorial changes in the manuscript. All authors read

and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflicts of interest. Kai Yuen Wong is serving as one of the Editorial Board members of this journal. We declare that Kai Yuen Wong had no involvement in the peer review of this article and has no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to John Sandars.

References

- [1] Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, *et al.* Objective structured assessment of technical skill (OSATS) for surgical residents. *The British Journal of Surgery.* 1997; 84: 273–278. <https://doi.org/10.1046/j.1365-2168.1997.02502.x>.
- [2] Aghazadeh MA, Jayaratna IS, Hung AJ, Pan MM, Desai MM, Gill IS, *et al.* External validation of Global Evaluative Assessment of Robotic Skills (GEARS). *Surgical Endoscopy.* 2015; 29: 3261–3266. <https://doi.org/10.1007/s00464-015-4070-8>.
- [3] Goddard AF. The European Working Time Directive and the impact on training: the current evidence. *Clinical Medicine.* 2010; 10: 317–318. <https://doi.org/10.7861/clinmedicine.10-4-317>.
- [4] Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *The New England Journal of Medicine.* 2006; 355: 2664–2669. <https://doi.org/10.1056/NEJMr054785>.
- [5] Guerrero DT, Asaad M, Rajesh A, Hassan A, Butler CE. Advancing Surgical Education: The Use of Artificial Intelligence in Surgical Training. *The American Surgeon.* 2023; 89: 49–54. <https://doi.org/10.1177/00031348221101503>.
- [6] Leon S, Lee S, Perez JE, Hashimoto DA. Artificial intelligence and the education of future surgeons. *American Journal of Surgery.* 2025; 246: 116257. <https://doi.org/10.1016/j.amjsurg.2025.116257>.
- [7] Kirubarajan A, Young D, Khan S, Crasto N, Sobel M, Sussman D. Artificial Intelligence and Surgical Education: A Systematic Scoping Review of Interventions. *Journal of Surgical Education.* 2022; 79: 500–515. <https://doi.org/10.1016/j.jsurg.2021.09.012>.
- [8] Pedrett R, Mascagni P, Beldi G, Padoy N, Lavanchy JL. Technical skill assessment in minimally invasive surgery using artificial intelligence: a systematic review. *Surgical Endoscopy.* 2023; 37: 7412–7424. <https://doi.org/10.1007/s00464-023-10335-z>.
- [9] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine.* 2019; 25: 44–56. <https://doi.org/10.1038/s41591-018-0300-7>.
- [10] Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in health-care: a multidisciplinary perspective. *BMC Medical Informatics and Decision Making.* 2020; 20: 310. <https://doi.org/10.1186/s12911-020-01332-6>.
- [11] Weiner EB, Dankwa-Mullan I, Nelson WA, Hassanpour S. Ethical challenges and evolving strategies in the integration of artificial intelligence into clinical practice. *PLOS Digital Health.* 2025; 4: e0000810. <https://doi.org/10.1371/journal.pdig.0000810>.
- [12] Yilmaz R, Bakhaider M, Alsayegh A, Abou Hamdan N, Fazlollahi AM, Tee T, *et al.* Real-Time multifaceted artificial intelligence vs In-Person instruction in teaching surgical technical skills: a randomized controlled trial. *Scientific Reports.* 2024; 14: 15130. <https://doi.org/10.1038/s41598-024-65716-8>.
- [13] Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS ONE.* 2020; 15: e0229596. <https://doi.org/10.1371/journal.pone.0229596>.
- [14] Azari DP, Frasier LL, Quamme SRP, Greenberg CC, Pugh CM, Greenberg JA, *et al.* Modeling Surgical Technical Skill Using Expert Assessment for Automated Computer Rating. *Annals of Surgery.* 2019; 269: 574–581. <https://doi.org/10.1097/SLA.0000000000002478>.
- [15] Igaki T, Kitaguchi D, Matsuzaki H, Nakajima K, Kojima S, Hasegawa H, *et al.* Automatic Surgical Skill Assessment System Based on Concordance of Standardized Surgical Field Development Using Artificial Intelligence. *JAMA Surgery.* 2023; 158: e231131. <https://doi.org/10.1001/jamasurg.2023.1131>.
- [16] Ward TM, Mascagni P, Madani A, Padoy N, Perretta S, Hashimoto DA. Surgical data science and artificial intelligence for surgical education. *Journal of Surgical Oncology.* 2021; 124: 221–230. <https://doi.org/10.1002/jso.26496>.
- [17] Mason K, Lim E, Higham A, de Pennington N. How Large Language Models Can Shape Your Clinical Practice and Make You a More Efficient Clinician. *Bulletin.* 2025; 107: 154–156. <https://doi.org/10.1308/rcsbull.2025.60>.
- [18] Kitaguchi D, Takeshita N, Hasegawa H, Ito M. Artificial intelligence-based computer vision in surgery: Recent advances and future perspectives. *Annals of Gastroenterological Surgery.* 2021; 6: 29–36. <https://doi.org/10.1002/ags3.12513>.
- [19] Deng H, Eftekhari Z, Carlin C, Veerapong J, Fournier KF, Johnston FM, *et al.* Development and Validation of an Explainable Machine Learning Model for Major Complications After Cytoreductive Surgery. *JAMA Network Open.* 2022; 5: e2212930. <https://doi.org/10.1001/jamanetworkopen.2022.12930>.
- [20] Braunstein D, Farooq H, Paolino M, Moynihan A, Cahill RA. Comparative analysis of perceptions on artificial intelligence in surgery: a survey study among surgeons and medical students in Ireland. *Irish Journal of Medical Science.* 2025; 194: 2341–2349. <https://doi.org/10.1007/s11845-025-04079-z>.
- [21] Brian R, Murillo A, Gomes C, Alseidi A. Artificial intelligence and robotic surgical education. *Global Surgical Education.* 2024; 3: 60. <https://doi.org/10.1007/s44186-024-00262-5>.
- [22] Park JJ, Tiefenbach J, Demetriades AK. The role of artificial intelligence in surgical simulation. *Frontiers in Medical Technology.* 2022; 4: 1076755. <https://doi.org/10.3389/fmedt.2022.1076755>.
- [23] Hamza H, Shabir D, Aboumarzouk O, Al-Ansari A, Shaban K, Navkar NV. Automated skills assessment in open surgery: a scoping review. *Engineering Applications of Artificial Intelligence.* 2025; 153: 110893. <https://doi.org/10.1016/j.engappai.2025.110893>.
- [24] Moglia A, Georgiou K, Georgiou E, Satava RM, Cuschieri A. A systematic review on artificial intelligence in robot-assisted surgery. *International Journal of Surgery.* 2021; 95: 106151. <https://doi.org/10.1016/j.ijssu.2021.106151>.

- [25] Loftus TJ, Tighe PJ, Filiberto AC, Efron PA, Brakenridge SC, Mohr AM, *et al.* Artificial Intelligence and Surgical Decision-making. *JAMA Surgery*. 2020; 155: 148–158. <https://doi.org/10.1001/jamasurg.2019.4917>.
- [26] Cacciamani GE, Anvar A, Chen A, Gill I, Hung AJ. How the use of the artificial intelligence could improve surgical skills in urology: state of the art and future perspectives. *Current Opinion in Urology*. 2021; 31: 378–384. <https://doi.org/10.1097/MOU.0000000000000890>.
- [27] Singh SP, Shayan AM, Gao J, Bible J, Groff RE, Singapogu R. Objective and Automated Quantification of Instrument Handling for Open Surgical Suturing Skill Assessment: A Simulation-Based Study. *IEEE Open Journal of Engineering in Medicine and Biology*. 2024; 5: 485–493. <https://doi.org/10.1109/OJEMB.2024.3402393>.
- [28] Fazlollahi AM, Bakhaidar M, Alsayegh A, Yilmaz R, Winkler-Schwartz A, Mirchi N, *et al.* Effect of Artificial Intelligence Tutoring vs Expert Instruction on Learning Simulated Surgical Skills Among Medical Students: A Randomized Clinical Trial. *JAMA Network Open*. 2022; 5: e2149008. <https://doi.org/10.1001/jamanetworkopen.2021.49008>.
- [29] Nakawala H, Ferrigno G, De Momi E. Development of an intelligent surgical training system for Thoracentesis. *Artificial Intelligence in Medicine*. 2018; 84: 50–63. <https://doi.org/10.1016/j.artmed.2017.10.004>.
- [30] Kyeremeh J, Asciak L, Blackmur JP, Luo X, Picard F, Shu W, *et al.* Digital twins assisted surgery: A conceptual framework for transforming surgical training and navigation. *The Surgeon*. 2025. <https://doi.org/10.1016/j.surge.2025.09.007>. (online ahead of print)
- [31] Mekki YM, Luijten G, Hagert E, Belkhair S, Varghese C, Qadir J, *et al.* Digital twins for the era of personalized surgery. *NPJ Digital Medicine*. 2025; 8: 283. <https://doi.org/10.1038/s41746-025-01575-5>.
- [32] Fard MJ, Ameri S, Darin Ellis R, Chinnam RB, Pandya AK, Klein MD. Automated robot-assisted surgical skill evaluation: Predictive analytics approach. *The International Journal of Medical Robotics + Computer Assisted Surgery*. 2018; 14: 10.1002/rcs.1850. <https://doi.org/10.1002/rcs.1850>.
- [33] 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. *Surgical Endoscopy*. 2025; 39: 5018–5026. <https://doi.org/10.1007/s00464-025-11903-1>.
- [34] Kasa K, Burns D, Goldenberg MG, Selim O, Whyne C, Hardisty M. Multi-Modal Deep Learning for Assessing Surgeon Technical Skill. *Sensors*. 2022; 22: 7328. <https://doi.org/10.3390/s22197328>.
- [35] Bakken S. AI in health: keeping the human in the loop. *Journal of the American Medical Informatics Association*. 2023; 30: 1225–1226. <https://doi.org/10.1093/jamia/ocad091>.
- [36] Hassan M, Ayad M, Nembhard C, Hayes-Dixon A, Lin A, Janjua M, *et al.* Artificial Intelligence Compared to Manual Selection of Prospective Surgical Residents. *Journal of Surgical Education*. 2025; 82: 103308. <https://doi.org/10.1016/j.jsurg.2024.103308>.
- [37] Horodyski P. Applicants' perception of artificial intelligence in the recruitment process. *Computers in Human Behavior Reports*. 2023; 11: 100303. <https://doi.org/10.1016/j.chbr.2023.100303>.
- [38] Dadaboyev SM, Abdullayeva J, Abbosova N, Suleymenova A, Mamadjanova K. Role of artificial intelligence in employee recruitment: systematic review and future research directions. *Discover Global Society*. 2025; 3: 99. <https://doi.org/10.1007/s44282-025-00246-w>.
- [39] Tursunbayeva A, Fernandez V, Gallardo-Gallardo E, Moschera L. Artificial intelligence and digital data in recruitment. Exploring business and engineering candidates' perceptions of organizational attractiveness. *European Management Journal*. 2026; 44: 235–244. <https://doi.org/10.1016/j.emj.2025.03.002>.
- [40] Sumner MD, Howell TC, Soto AL, Kaplan S, Tracy ET, Zaas AK, *et al.* The Use of Artificial Intelligence in Residency Application Evaluation—A Scoping Review. *Journal of Graduate Medical Education*. 2025; 17: 308–319. <https://doi.org/10.4300/JGME-D-24-00604.1>.
- [41] Arjomandi Rad A, Vardanyan R, Athanasiou T, Maessen J, Sardari Nia P. The ethical considerations of integrating artificial intelligence into surgery: a review. *Interdisciplinary Cardiovascular and Thoracic Surgery*. 2025; 40: ivae192. <https://doi.org/10.1093/icvts/ivae192>.
- [42] Johnson-Mann CN, Loftus TJ, Bihorac A. Equity and Artificial Intelligence in Surgical Care. *JAMA Surgery*. 2021; 156: 509–510. <https://doi.org/10.1001/jamasurg.2020.7208>.
- [43] Twinanda AP, Shehata S, Mutter D, Marescaux J, de Mathelin M, Padoy N. EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos. *IEEE Transactions on Medical Imaging*. 2017; 36: 86–97. <https://doi.org/10.1109/TMI.2016.2593957>.
- [44] Sheller MJ, Edwards B, Reina GA, Martin J, Pati S, Kotrotsou A, *et al.* Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*. 2020; 10: 12598. <https://doi.org/10.1038/s41598-020-69250-1>.
- [45] Cobianchi L, Verde JM, Loftus TJ, Piccolo D, Dal Mas F, Mascagni P, *et al.* Artificial Intelligence and Surgery: Ethical Dilemmas and Open Issues. *Journal of the American College of Surgeons*. 2022; 235: 268–275. <https://doi.org/10.1097/XCS.0000000000000242>.
- [46] Varas J, Coronel BV, Villagrán I, Escalona G, Hernandez R, Schuit G, *et al.* Innovations in surgical training: exploring the role of artificial intelligence and large language models (LLM). *Revista do Colegio Brasileiro de Cirurgioes*. 2023; 50: e20233605. <https://doi.org/10.1590/0100-6991e-20233605-en>.
- [47] Crowston K, Bolici F. Deskilling and upskilling with AI systems. *Information Research an International Electronic Journal*. 2025; 30: 1009–1023. <https://doi.org/10.47989/ir30iConf47143>.
- [48] Budzyń K, Romańczyk M, Kitala D, Kołodziej P, Bugajski M, Adami HO, *et al.* Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study. *The Lancet. Gastroenterology & Hepatology*. 2025; 10: 896–903. [https://doi.org/10.1016/S2468-1253\(25\)00133-5](https://doi.org/10.1016/S2468-1253(25)00133-5).
- [49] Ferreres AR. Ethical and legal issues regarding artificial intelligence (AI) and management of surgical data. *European Journal of Surgical Oncology*. 2025; 51: 108279. <https://doi.org/10.1016/j.ejso.2024.108279>.
- [50] Royal College of Surgeons of England (RCS England). *Guidance on Artificial Intelligence and Digital Technologies in Surgical Practice*. RCS England: London. 2024.