Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Trends in the Use of Artificial Intelligence in Surgical Simulation

María Inés Gaete

Pontificia Universidad Católica de Chile: Pontificia Universidad Catolica de Chile

Julian Varas

jevaras@uc.cl

Pontificia Universidad Catolica de Chile Escuela de Medicina https://orcid.org/0000-0002-5828-9623

Research Article

Keywords: artificial intelligence, surgical simulation, machine learning, technical skills, assessment, feedback

Posted Date: February 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8937038/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Trends in the Use of Artificial Intelligence in Surgical Simulation

Abstract

Technical skill is a significant determinant of postoperative outcomes, yet traditional approaches to teaching and assessing surgical performance are resource-intensive, subjective, and challenging to scale. Simulation-based training was introduced to address these limitations; however, most programs still rely on human observation and expert availability. Artificial intelligence (AI) has recently emerged as a powerful complement to simulation, offering automated performance analysis, personalized feedback, and large-scale educational analytics.

This narrative review describes current trends in the use of AI in surgical simulation. We first summarize innovations in machine learning and deep learning, natural language processing, and skills and performance analysis. We then outline key educational opportunities related to technical skills training, assessment, debriefing, and quality improvement. Implementation strategies and challenges are discussed, including institutional buy-in, funding and infrastructure, integration into existing curricula, and faculty development. Finally, we highlight future directions such as multimodal foundation models, generative simulation, global expansion into low and middle-income settings, and research priorities for rigorous validation and equity.

Across multiple studies, AI-based systems can accurately classify expertise, segment procedures, and predict scores on validated rating scales. Randomized trials also show that AI tutoring can match or exceed expert instruction in simulated tasks. However, unintended effects, algorithmic bias, and limited evidence of long-term educational impact underscore the need for thoughtful implementation and sustained human oversight. AI, at least over the next 5 years, should be understood as an augmenting technology that can enhance, but not replace, expert surgical educators.

Keywords: artificial intelligence; surgical simulation; machine learning; technical skills; assessment; feedback.

Introduction

Technical skill is strongly associated with patient outcomes. Peer-rated assessments of bariatric surgeons showed that those in the lowest quartile of technical skill had significantly higher rates of complications and mortality compared with those in the highest quartile [1]. Similar associations have been demonstrated in colorectal and other procedures, where variation in surgeon skill accounts for a substantial proportion of variation in postoperative complications, reoperations, and serious morbidity [2,3]. These data have reinforced the need for structured training and objective assessment of operative performance.

Simulation-based education has become a cornerstone of competency-based surgical training. Simulators and video-based assessments allow deliberate practice, repeated performance, and the safe acquisition of technical skills outside the operating room. Nevertheless, most simulation programs still rely on human raters to score performance and provide feedback. Faculty time is limited, inter-rater reliability can be modest, and scaling individualized feedback to large cohorts of learners is challenging.

Artificial intelligence (AI) has been proposed to address these gaps. Early work explored the use of computer vision and machine learning to extract metrics from virtual reality simulators and classify levels of expertise [4]. The broader surgical literature has highlighted both the promises and the perils of AI, including opportunities for decision support and training, as well as concerns regarding bias, transparency, and regulation [5]. Within simulation specifically, recent reviews emphasize that modern AI techniques can support simulators in providing personalized feedback, automating assessment, and creating immersive learning environments [6].

Recent narrative overviews of AI in surgical training describe emerging applications across simulation, automated assessment, and intraoperative decision support, and highlight how AI may help mitigate reduced operative exposure and growing service pressures [7,8].

Methods

This article is a narrative review that synthesizes recent literature on artificial intelligence in surgical simulation. We drew on peer-reviewed empirical studies, systematic and narrative reviews, consensus statements, and professional society reports that addressed AI in relation to simulation-based training, technical skills assessment, debriefing, implementation, or future directions. Evidence from both surgical and closely related health professions domains was included when it directly informed the use of AI in surgical simulation. The goal was to provide a conceptual overview of current trends rather than an exhaustive systematic review or meta-analysis.

The objective of this narrative review is to describe current trends in the use of AI in surgical simulation, following a structure that includes technological innovations, educational opportunities, implementation strategies and challenges, and future directions for research.

Innovations

Machine learning and deep learning

Machine learning (ML) and deep learning are the most mature AI technologies in surgical simulation. A systematic review of 66 studies found that ML algorithms applied to motion data or video achieved accuracies above 80 percent when classifying technical skill, although tasks and datasets varied widely [9].

One of the earliest high-impact examples is the work by Winkler-Schwartz et al., who used a virtual reality neurosurgical simulator to capture detailed psychomotor metrics and trained algorithms to classify participants as novices, intermediates, or experts [4]. The models identified specific operative factors

associated with expertise, such as economy of motion and bimanual coordination, demonstrating that ML could uncover features of technical skill not readily apparent to human raters.

Building on these foundations, Mirchi et al. introduced the Virtual Operative Assistant, an explainable AI framework that analyzes performance in a tumor resection simulation and provides an overall skill classification, along with a breakdown of metric performance [10]. Deep neural networks have also been used to monitor expertise continuously. Yilmaz et al. developed the Intelligent Continuous Expertise Monitoring System, which uses recurrent and convolutional architectures to estimate bimanual skill every 0.2 seconds in a virtual reality task, allowing fine-grained, real-time assessment [11].

Other groups have integrated novel data streams. An eye-movement-based deep learning system distinguished laparoscopic novices, intermediates, and experts with over 80% accuracy using gaze coordinates and image data from a standard peg transfer task [12]. Natheir et al. combined electroencephalography with ML models to differentiate skilled from less skilled performance during simulated neurosurgery, suggesting that cognitive workload signatures could complement motor performance metrics [13].

More recently, a large-scale laparoscopic surgical performance dataset was introduced to support deep learning-based assessment of simulated tasks. Power et al. used three-dimensional convolutional networks trained on hundreds of annotated videos to classify expertise levels with high discriminative performance, providing a benchmark for future work in automated video-based assessment [14]. In parallel, AI-assisted laparoscopic box trainers have been developed that use object detection and tracking to generate validated metrics during basic tasks, enabling automated evaluations at scale [15].

In addition to virtual reality and synthetic environments, box trainer platforms have begun to integrate computer vision architectures such as YOLO to automatically track instruments, identify key steps, and calculate task times. In basic laparoscopic tasks, these systems have demonstrated high agreement with human raters, providing a scalable solution for automated assessment in low-cost simulators [16].

In robotic surgery, expert groups anticipate that the next generation of simulators will combine high-fidelity haptics, cloud-based performance tracking, and AI-driven guidance, thereby tightening the link between console practice and clinical performance [17].

Industry reports map a rapidly expanding ecosystem of surgical AI companies developing video-analytics platforms, digital surgery dashboards, telepresence systems, and VR-based simulation tools, underscoring the pace of commercial innovation and the need for educators to critically appraise vendor offerings [18].

Natural language processing

Natural language processing (NLP) is beginning to influence surgical simulation by transforming how feedback is generated and delivered. While much of the literature on large language models in surgery is still conceptual, several reviews outline the potential of NLP for summarizing performance, translating quantitative metrics into narrative comments, and supporting reflection [5,19].

In practice, early systems have coupled ML-based scoring engines with templated or semi-free-text feedback modules. The Virtual Operative Assistant, for example, provides textual explanations of which aspects of performance were above or below expert benchmarks, reinforcing positive behavior and indicating specific targets for improvement [10]. Emerging prototypes, including institutional platforms for asynchronous video review, are starting to explore the use of large language models to draft feedback that faculty can then edit. These tools promise to reduce the time required to produce high-quality feedback while preserving the educator's voice.

Early empirical work has begun to test these concepts in other health professions. In physiotherapy skills training, a large language model integrated into a video-based practice platform was used to classify the

quality of instructors' comments and generate clearer reformulations, supporting feedback uptake without replacing human oversight [20].

In surgery, conversational agents such as ChatGPT have also been used to generate narrative feedback on intraoperative scenarios. In a recent study comparing surgeon-written and AI-generated comments, trainees rated the usefulness and clarity of ChatGPT feedback as comparable to faculty feedback, while still emphasizing the importance of human validation and contextualization [21].

Skills and performance analysis

AI-driven performance analysis extends beyond classification to examine how skills evolve over time and how assessment metrics are related to competency frameworks. The consensus statement from Vedula et al. defined priorities for AI-enabled metrics in surgical education, including objective quantification of learning curves, automated detection of unsafe behaviors, and standardized thresholds for entrustment decisions [22].

Scoping reviews of automated video analysis and AI-based skill assessment have highlighted substantial variability in tasks, metrics, and validation strategies, but they consistently show that AI systems can approximate human ratings on widely used global rating scales such as the Objective Structured Assessment of Technical Skill (OSATS) and the Global Evaluative Assessment of Robotic Skills (GEARS) in minimally invasive and robotic settings [23,24].

At the same time, new evaluation paradigms are emerging. AI-based systems can identify discrete errors, such as tissue collisions or excessive force, and link them to safety outcomes in simulation. Some models combine multiple data modalities to derive higher-level constructs like efficiency, precision, and respect for tissue, inching closer to a holistic representation of technical competence.

Opportunities

Training skills

Evidence that AI can enhance skills acquisition in simulation is accumulating. In a randomized clinical trial involving 70 medical students, an AI tutor based on the Virtual Operative Assistant produced higher performance scores and better skill transfer than remote expert instruction or no feedback, with similar cognitive and affective responses across groups [25]. A subsequent cohort study showed that an AI-designed curriculum improved safety metrics, such as reduced tissue damage and lower applied forces, although it also unintentionally slowed hand movements, highlighting the need for balanced training objectives [26].

Beyond specific AI tools, reviews of surgical simulation and training technologies in general surgery education have confirmed that structured simulation curricula improve technical performance and confidence, particularly when paired with objective assessment and feedback mechanisms [27].

A randomized controlled trial comparing real time AI-feedback, in-person expert instruction, and no feedback found that trainees receiving AI-guidance achieved superior composite performance while maintaining OSATS scores comparable to those taught by experts [28]. More recently, a crossover trial demonstrated that the most effective sequence paired initial expert coaching with subsequent AI-feedback, reinforcing the view that AI and human instruction are complementary rather than competing [29].

A subsequent commentary on this trial further explored the implications of AI-augmented human instruction for curriculum design, arguing that algorithms should be framed as tools for structuring expert coaching rather than as stand-alone tutors [30].

An accompanying editorial from a managed-care perspective emphasized that the greatest gains are achieved when AI augments, rather than replaces, human coaching and stressed that preserving the "human touch" remains essential for learner motivation and professionalism [31].

Beyond isolated tasks, AI-driven systems are beginning to reshape entire curricula. A systematic review of AI-based laparoscopic suturing training reported consistent improvements in precision, efficiency, and long-term retention compared with traditional methods, particularly when deep learning and motion capture were used to drive real-time feedback on needle handling, knot tying, and grip control [32].

These developments align with institutional efforts to expand access to remote and asynchronous skills training. Video-centric platforms that allow trainees to upload procedures, receive structured expert feedback, and increasingly combine that feedback with AI-generated analytics, such as C1DO1 and Teach1, illustrate how AI-enabled simulation can be integrated into everyday clinical workflows without requiring constant faculty presence.

At the undergraduate level, large-scale skills programs built around asynchronous video upload and structured feedback have shown that thousands of students can be trained and assessed with relatively modest faculty time, while achieving sustained improvements in technical performance. These initiatives illustrate how video-based workflows create an infrastructure that can later support AI-enabled assessment [33].

Similarly, multi-institutional distance simulation programs for advanced laparoscopic skills have demonstrated that asynchronous video-based training can be scaled across centers and countries while maintaining high learner satisfaction and measurable gains in performance [34].

Assessment and debriefing

AI can reduce assessment burden while improving consistency. In a multi-institutional study using three hospitals on two continents, Kiyasseh et al. showed that an AI system trained with human explanations produced feedback on suturing videos that aligned more closely with expert comments and reduced bias compared with standard training approaches [22].

Some platforms combine automated scoring with analytics on the feedback itself. In the Teach1 system, AI tools first pre-screen videos to flag potential technical issues and then analyse the feedback written by instructors, scoring it on dimensions such as specificity and actionability and offering suggestions for improvement. Higher-quality feedback as rated by the AI system has been associated with greater learner performance gains, suggesting that AI can support both student learning and faculty development [35].

Automated video analysis systems can segment surgical procedures into phases and steps, quantify time spent in each segment, and flag deviations from best practice. Scoping reviews report that such metrics differentiate between experience levels and correlate with human ratings in both simulation and clinical settings, although longitudinal data linking AI-based assessments to actual improvement remain limited [36].

In debriefing, AI-generated dashboards and visualizations can summarize key events, errors, and trends, allowing educators to focus on higher-level cognitive aspects of performance. Trials of AI-augmented human instruction show that when experts have access to AI-derived metrics, their coaching becomes more targeted, and trainee performance improves beyond what either AI or human tutoring can achieve alone [29].

Learning analytics dashboards that aggregate AI-derived metrics across multiple attempts can further enhance debriefing. By visualizing trends in errors, task times, and feedback uptake over time, these dashboards help learners plan deliberate practice and enable faculty to identify trainees who may benefit from targeted coaching [37].

Quality improvement

Simulation has long been viewed as a tool for quality improvement, and AI amplifies this role by enabling large-scale and objective analyses. When simulation performance can be measured automatically and

consistently, programs can monitor cohorts over time, identify persistent weaknesses, and evaluate the impact of curricular interventions.

Although most AI simulation studies focus on educational outcomes, their importance is underscored by clinical evidence. Peer-rated technical skill and structured video assessment have been linked to complication rates and mortality in bariatric surgery, colectomy, and other operations [1,2,3]. If AI-based assessment can reliably approximate or enhance these ratings, it may become a powerful surrogate for tracking technical quality and supporting surgeon remediation, with downstream effects on patient safety.

Experiences from orthopedic surgery describe how AI-enabled planning tools, simulators, and postoperative analytics are already being linked to quality-improvement pipelines, offering a template for connecting simulation data with real-world outcomes in other surgical disciplines [38].

In bariatric and metabolic surgery, an international Delphi consensus has outlined core AI-related competencies for surgeons and highlighted the need to integrate topics such as data governance, algorithmic bias, and model interpretability into training, framing AI adoption as part of a broader quality and safety agenda [39].

Global initiatives also highlight quality improvement at the systems level. For example, AI-supported training programs for basic emergency surgery in low and middle-income countries aim to mitigate shortages of surgical educators by using low-cost simulators combined with automated phase recognition to ensure that essential steps of standard procedures are performed correctly before trainees operate on patients [40]. Recent position statements from professional societies such as the American College of Surgeons argue that AI-enabled simulation and performance analytics could help standardize training and support quality improvement programs across diverse health systems [41].

Implementation Strategies and Overcoming Challenges

Institutional buy-in

Successful adoption of AI in simulation requires institutional commitment. Decision-makers must balance enthusiasm for innovation with evidence of educational benefits, cost considerations, and regulatory obligations. Reviews of AI in surgical simulation and education emphasize that current evidence is promising but largely confined to single-center studies or short-term trials, and that rigorous, multi-site evaluations are needed before high-stakes decisions rely on AI outputs [6,19]. Recent reports describing AI deployment in the operating room highlight how surgeons are already navigating ambient documentation, predictive scheduling, and real-time guidance tools, underscoring the need to align educational uses of AI with broader clinical workflow changes [42].

The multidisciplinary consensus, led by the American College of Surgeons, provides a roadmap for AI methods and metrics in surgical education, including priority use cases, data requirements, and milestones over two, five, and ten years [22]. This type of guidance can help institutions frame AI-projects within broader educational strategies rather than isolated technology pilots. Recent roadmaps for educational technologies in surgery and conceptual models such as the Four Domains Model of AI in Surgical Education emphasize aligning AI-enabled simulation with clear educational objectives, governance structures, and competency-based curricula [43,44].

Robust data governance is also critical. Experiences from AI-assisted plastic surgery have underscored challenges related to image anonymization, consent for secondary use of training videos, cross-border data transfer, and the risk of re-identification when models are trained on small datasets, reinforcing the need for explicit privacy protocols in simulation projects [45].

Funding and infrastructure

AI-enabled simulation demands investment in hardware, software, and data infrastructure. High-quality video capture, secure storage, and computing resources are essential. Some commercial simulators now incorporate AI-modules, whereas others require local development or integration with external platforms. Institutions must consider the total cost of ownership, including maintenance, data governance, and technical support. Consortia that share datasets and models may reduce duplication of effort and increase equity by allowing resource-constrained centers to access validated tools.

Analyses of AI adoption in healthcare systems underscore that technical capability alone is insufficient; data fragmentation, lack of clear business models, regulatory uncertainty, and clinician mistrust remain major barriers, all of which also shape the feasibility and sustainability of AI-enabled simulation initiatives [46].

At the hospital level, implementation reports frequently converge on a similar set of barriers: limited technical support, concerns about reliability in edge cases, skepticism from front-line clinicians, and lack of clear return on investment, as illustrated by surveys of hospital leaders on AI deployment in clinical workflows [47].

Integration into training

Embedding AI into existing curricula is more complex than acquiring technology. Educators must determine which competencies will be assessed, how AI generated metrics will be used in both formative and summative assessments, and how to communicate these expectations to learners effectively. Scoping reviews of AI and ML in surgical coaching note that very few studies examine user acceptance, workflow integration, or impact on overall program outcomes [48].

An incremental approach is often prudent. Programs may begin by utilizing AI for low-stakes formative feedback, such as automated metrics on basic tasks or preliminary scoring of video submissions, while maintaining human control over progression decisions. As evidence accumulates and trust increases, AI-

based assessments can be gradually incorporated into formal milestones or microcredentials, with clear policies in place for oversight and appeal.

Faculty development

Faculty trust and competence are critical for safe implementation. Instructors require training to interpret AI-outputs effectively, understand their limitations, and integrate them judiciously into coaching. Studies on AI-guided curricula highlight unintended consequences when systems optimize for a narrow set of metrics, such as slower but safer movements, which may not align with all clinical goals [26].

Randomized trials show that combining expert coaching with AI-feedback yields better outcomes than either approach alone [28,29]. Complementing these learner-focused interventions, an AI-based system (Teach1) that audits the quality of instructors' written feedback in simulated paracentesis training has shown that higher-quality feedback, as scored by the algorithm on established criteria, is associated with greater subsequent gains in OSATS performance, underscoring that AI is most powerful when it strengthens human coaching rather than replaces it [35].

Educators therefore need support not only in using AI but also in rethinking their role as supervisors of AI-guided learning, curating feedback, and helping learners reconcile automated comments with clinical judgment. Broader cultural issues are also relevant. AI can be perceived as threatening if it is seen as a replacement for human expertise. Faculty development programs that frame AI as a tool to enhance mentorship, reduce repetitive scoring tasks, and free time for higher level teaching are more likely to foster engagement.

Future Directions and Research

The next generation of AI in surgical simulation will likely be multimodal, adaptive, and globally networked. Foundation models that integrate video, audio, kinematic data, eye-tracking, and textual descriptors may provide richer representations of operative performance than any single data stream alone. Such models could power intelligent tutors capable of engaging in conversational interactions, providing contextualized explanations, and continuously adapting to learner needs This trajectory closely parallels broader developments in surgical AI, where multimodal models are being tested for clinical decision support and video-based analytics [49]. More broadly, reviews of AI-supported precision surgery suggest that gains in intraoperative accuracy and outcomes will depend on how well training programs integrate these tools into simulation and skills curricula [50].

Generative AI is poised to influence scenario design. Synthetic video and three-dimensional content can be used to augment datasets, simulate rare complications, or create customized cases without increasing patient risk. Early work in computer vision suggests that hand pose estimation and gesture recognition can provide accurate proxies for technical dexterity using standard video alone, which may facilitate AI-assisted training even in low-resource environments [51].

At the program level, AI could support longitudinal tracking of competence across rotations and institutions, enabling portable digital portfolios that document validated performance over time. Reviews focused on global surgical education argue that AI will be essential for aggregating and curating the large volumes of operative video and educational content generated worldwide, helping educators deliver just-in-time resources and identify knowledge gaps [52,53].

Key research priorities include external validation of AI tools across diverse settings, robust evaluation of educational impact beyond short-term task performance, assessment of effects on patient outcomes, and exploration of ethical and legal implications such as data ownership, consent, and algorithmic fairness [3,40,48]. Engagement of trainees and faculty in co-design processes will be vital to ensure that emerging technologies address real educational needs rather than creating additional burdens.

Conclusion

Artificial intelligence is rapidly reshaping the landscape of surgical simulation. Machine learning and deep learning models can classify expertise, continuously monitor performance, and link complex motion or video data to validated rating scales. Natural language processing is beginning to transform how feedback is generated and delivered. Randomized trials indicate that AI tutoring can match or exceed expert instruction in simulated tasks, and consensus statements outline how AI-enabled metrics could support competency-based curricula.

At the same time, significant challenges remain. Evidence is still limited regarding long-term learning, clinical transfer, and equity. Unintended effects, such as overemphasis on specific metrics and the risk of algorithmic bias, highlight the need for careful oversight. Implementation requires institutional commitment, infrastructure, thoughtful integration into assessment systems, and sustained faculty development.

Overall, AI should be understood as an augmenting technology. When used responsibly, it can extend the reach of expert educators, provide consistent and scalable feedback, and support data-informed quality improvement. The goal is not to replace human judgment, but to harness AI to train safer, more competent surgeons and promote high-quality surgical care across diverse settings.

Disclosures: Dr. Julian Varas is the founder of Training Competence, an official spinoff startup from the Pontificia Universidad Católica de Chile. Training Competence develops educational web-based and mobile platforms. The rest of authors have nothing to disclose.

References

1. Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. N Engl J Med. 2013;369(15):1434-1442.

2. Stulberg JJ, Huang R, Kreutzer L, et al. Association between surgeon technical skills and patient outcomes. JAMA Surg. 2020;155(10):960-968.

3. Woods MS, Gawande AA, Hutter MM, et al. Association between surgical technical skills and clinical outcomes: a systematic review. Ann Surg. 2023;278(2):e419-e428.

4. Winkler-Schwartz A, Yilmaz R, Mirchi N, et al. Machine learning identification of surgical and operative factors associated with surgical expertise in virtual reality simulation. JAMA Netw Open. 2019;2(8):e198363.

5. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. Ann Surg. 2018;268(1):70-76.

6. Park JJ, Tiefenbach J, Demetriades AK. The role of artificial intelligence in surgical simulation. Front Med Technol. 2022;4:1076755.

7. Guerrero DT, Kewalramani D, Villagrán I, Varas J. Advancing surgical education: the use of artificial intelligence in surgical training. Am Surg. 2023;89(1):49-54.

8. Valente LM, Park AJ, Nahai F. Transforming surgical education with artificial intelligence: opportunities and challenges. Aesthetic Surg J. 2024;44(5):NP1577-NP1586

9. Lam K, Chen J, Wang Z, et al. Machine learning for technical skill assessment in surgery: a systematic review. NPJ Digit Med. 2022;5(1):24.

10. Mirchi N, Bissonnette V, Yilmaz R, et al. The Virtual Operative Assistant: an explainable artificial intelligence tool for simulation based training in surgery and medicine. PLoS One. 2020;15(2):e0229596.

11. Yilmaz R, Winkler-Schwartz A, Mirchi N, et al. Continuous monitoring of surgical bimanual expertise using deep neural networks in virtual reality simulation. NPJ Digit Med. 2022;5(1):54.

12. Kuo RJ, Chen HJ, Kuo YH. The development of an eye movement based deep learning system for laparoscopic surgical skills assessment. Sci Rep. 2022;12:11036.

13. Natheir S, Christie S, Yilmaz R, et al. Utilizing artificial intelligence and electroencephalography to assess expertise on a simulated neurosurgical task. Comput Biol Med. 2023;152:106286.

14. Power D, Burke C, Madden MG, Ullah I. Automated assessment of simulated laparoscopic surgical skill performance using deep learning. Sci Rep. 2025;15(1):13591.

15. Belmar F, Gaete MI, Escalona G, et al. Artificial intelligence in laparoscopic simulation: a promising future for large scale automated evaluations. Surg Endosc. 2023;37(6):4942-4946.

16. Cruz E, Selman R, Figueroa U, Belmar F, Jarry C, Sanhueza D, et al. A scalable solution: effective AI implementation in laparoscopic simulation training assessments. Glob Surg Educ. 2025;4:46. https://doi.org/10.1007/s44186-025-00355-9.

17. Newhook TE, Ellison EC. Robotics: emerging trends that herald the future of robotic surgical simulation. AUA News. 2024 Mar;29(3) Available from: https://auanews.net/issues/articles/2024/march-2024/robotics-emerging-trends-that-herald-the-future-of-robotic-surgical-simulation.

18. IntuitionLabs. AI-assisted surgery: a guide to key companies and technology [Internet]. 2024 [cited 2025 Dec 8] Available from: https://intuitionlabs.ai/articles/surgical-ai-companies.

19. Leon S, Lee S, Perez JE, Hashimoto DA. Artificial intelligence and the education of future surgeons. Am J Surg. 2025;246:116257.

20. Villagran I, Hernandez R, Schuit G, Neyem A, Fuentes-Cimma J, Miranda C, et al. Implementing Artificial Intelligence in Physiotherapy Education: A Case Study on the Use of Large Language Models (LLM) to Enhance Feedback. IEEE Trans Learn Technol. 2024;17:2079-2090. https://doi.org/10.1109/TLT.2024.3450210.

21. Jarry Trujillo C, Vela Ulloa J, Escalona Vivas G, Grasset Escobar E, Villagran Gutierrez I, Achurra Tirado P, et al. Surgeons vs ChatGPT: Assessment and Feedback Performance Based on Real Surgical Scenarios. J Surg Educ. 2024;81(7):960-966. https://doi.org/10.1016/j.jsurg.2024.03.012.

22. Kiyasseh D, Li Z, Haque TF, et al. A multi institutional study using artificial intelligence to provide reliable and fair feedback to surgeons. Commun Med. 2023;3(1):42.

23. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg. 1997;84(2):273-278.

24. Goh AC, Goldfarb DW, Sander JC, Miles BJ, Dunkin BJ. Global Evaluative Assessment of Robotic Skills (GEARS): a valid tool for assessment of robotic surgical skills. J Urol. 2012;187(1):247-252.

25. Fazlollahi AM, Bakhaidar M, Alsayegh A, et al. Effect of artificial intelligence tutoring vs expert instruction on learning simulated surgical skills among medical students: a randomized clinical trial. JAMA Netw Open. 2022;5(2):e2149008.

26. Fazlollahi AM, Yilmaz R, Winkler-Schwartz A, et al. Artificial intelligence in surgical curriculum design and unintended outcomes for technical competencies in simulation training. JAMA Netw Open. 2023;6(9):e2334658.

27. Shahrezaei M, et al. Medical education trends for future physicians in the era of advanced technology and artificial intelligence: an integrative review. BMC Med Educ. 2019;19:460.

28. Yilmaz R, Bakhaidar M, Alsayegh A, et al. Real time multifaceted artificial intelligence vs in person instruction in teaching surgical technical skills: a randomized controlled trial. Sci Rep. 2024;14:15130.

29. Giglio B, Albeloushi A, Alhaj AK, et al. Artificial intelligence–augmented human instruction and surgical simulation performance: a randomized clinical trial. JAMA Surg. 2025;160(9):993-1003.

30. Fried GM, Feldman LS. Artificial intelligence–augmented human instruction and surgical simulation performance. JAMA Surg. 2025;160(8):e252345.

31. Klein HE. How the human touch boosts AI surgical training. Am J Manag Care. 2025 Aug 6. Available Available from: https://www.ajmc.com/view/how-the-human-touch-boosts-ai-surgical-training.

32. Ogbonnaya CN, Li S, Tang C, et al. Exploring the role of artificial intelligence driven training in laparoscopic suturing: a systematic review of skills mastery, retention, and clinical performance in surgical education. Healthcare (Basel). 2025;13(5):571.

33. Figueroa U, Jarry C, Inzunza M, Montero I, Garrido F, Villagran I, et al. Innovation Meets Practice: A Scalable Simulation-based Methodology for Massive Paracentesis Training. Gastroenterology. 2025;168(5):865-869.e2. https://doi.org/10.1053/j.gastro.2024.12.015.

34. Jarry C, Varas Cohen J. Distance simulation in surgical education. Surgery. 2025;180:109097. https://doi.org/10.1016/j.surg.2024.109097

35. Kewalramani D, Sanhueza Roman D, Abedrapo Lagos S, Rammsy F, Villagran I, Escalona G, et al. Quality matters: Artificial intelligence-based assessment of feedback quality predicts technical skill improvement. Surgery. 2025;187:109625. https://doi.org/10.1016/j.surg.2025.109625.

36. Dick L, Boyle CP, Skipworth RJE, et al. Automated analysis of operative video in surgical training: scoping review. BJS Open. 2024;8(5):zrae124.

37. Villagran I, Hernandez R, Schuit G, Neyem A, Fuentes J, Larrondo L, et al. Enhancing Feedback Uptake and Self-Regulated Learning in Procedural Skills Training: Design and Evaluation of a Learning Analytics Dashboard. J Learn Anal. 2024;11(2):138-156. https://doi.org/10.18608/jla.2024.8195.

38. Feng E, Jayasuriya N, Kirby D. Artificial intelligence in orthopedic surgery: a practical overview of current applications and trends. SurgiColl. 2025;3(3).

39. Kermansaravi M, Chiappetta S, Shahabi Shahmiri S, Varas J, Parmar C, Lee Y, et al. International expert consensus on the current status and future prospects of artificial intelligence in metabolic and bariatric surgery. Sci Rep. 2025;15:9312. https://doi.org/10.1038/s41598-025-94335-0.

40. Bhatia MB, Namazi B, Matthews J, et al. Use of artificial intelligence to support surgical education personnel shortages in low and middle income countries: developing a safer surgeon. Global Surg Educ. 2023;2:64.

41. American College of Surgeons. AI has potential to transform global surgical systems. Bull Am Coll Surg. 2024 Jun;109(6) Available from: https://www.facs.org/for-medical-professionals/news-publications/news-and-articles/bulletin/2024/june-2024-volume-109-issue-6/ai-has-potential-to-transform-global-surgical-systems/.

42. American College of Surgeons. AI transforms the OR as surgeons navigate complex challenges. Bull Am Coll Surg. 2025 Sep;110(8) Available from: https://www.facs.org/for-medical-professionals/news-publications/news-and-articles/bulletin/2025/september-2025-volume-110-issue-8/ai-transforms-the-or-as-surgeons-navigate-complex-challenges/.

43. Tewari AR, Raghuwanshi J, Labaran LA, Li X. The Four Domains Model of Artificial Intelligence in Surgical Education: adapting applications to trainee level. J Innov Health Sci Educ. 2024;1(3):3.

44. Fried GM, Varas J, Telem DA, Greenberg CC, Hashimoto DA, Paige JT, et al. Opportunities and Applications of Educational Technologies in Surgical Education and Assessment. Ann Surg. 2025;281(1):34-39. https://doi.org/10.1097/SLA.0000000000006367.

45. Anderson JL, Mitchell SR, Thompson BK, Harper ED, Andrew J. Data privacy in AI-assisted plastic surgery. 2025 Aug 6 Available from: https://www.researchgate.net/publication/397330067_Data_Privacy_in_AI-Assisted_Plastic_Surgery.

46. Tian M, Khullar D. Why is AI adoption in health care lagging? Brookings Institution. 2024 Jan 4 Available from: https://www.brookings.edu/articles/why-is-ai-adoption-in-health-care-lagging/.

47. Meddare. Five barriers slowing AI adoption in hospitals [Internet]. 2025 [cited 2025 Dec 8] Available from: https://meddare.ai/barriers-to-ai-adoption-in-hospitals/.

48. Vedula SS, Ghazi A, Collins JW, et al. Artificial intelligence methods and AI enabled metrics for surgical education: a multidisciplinary consensus. J Am Coll Surg. 2022;234(6):1181-1192.

49. Hossain I, Madani A, Laplante S, et al. Machine learning perioperative applications in visceral surgery: a narrative review. Front Surg. 2024;11:1393898.

50. Ingale P, Kulkarni AP, Dargari VM, Shukla T. The role of artificial intelligence in enhancing surgical precision and outcomes. J Sci Adv Surg. 2019;6(3):1559.

51. Bkheet E, D'Angelo AL, Goldbraikh A, Laufer S. Using hand pose estimation to automate open surgery training feedback. arXiv. 2022.

52. Brian R, Murillo A, Gomes C, Alseidi A. Artificial intelligence and robotic surgical education. Global Surg Educ. 2024;3:60.

53. Kitaguchi D, Ito M. The power of AI on aggregating, managing, and curating medical knowledge for global surgical education. Global Surg Educ. 2024;3:65.