The American Journal of Surgery 246 (2025) 116257

Contents lists available at ScienceDirect

The American Journal of Surgery

journal homepage: www.americanjournalofsurgery.com

My Thoughts / My Surgical Practice

Artificial intelligence and the education of future surgeons

ARTICLE INFO

Keywords
Artificial intelligence
Machine learning
Simulation
Surgical education
Surgery

ABSTRACT

Artificial intelligence (AI) has the potential to reshape surgical education by enabling personalized feedback, advanced competency evaluations, and enhancing resident selection processes. Through AI-driven simulations and real-time feedback systems, surgical trainees can engage in adaptive learning environments that promote deliberate practice and accelerated skill acquisition. Moreover, intraoperative AI tools may soon offer decision support, guiding surgeons during complex procedures. However, integrating AI into surgical education and practice comes with significant challenges. These include the need for high-quality datasets, the transition of AI systems from simulated environments to actual surgeries, and the ethical implications of data privacy, algorithmic bias, and surgeon autonomy. Overreliance on AI could de-skill surgeons, while biased algorithms may perpetuate disparities in resident selection and performance evaluations. To address these issues, regulatory frameworks must be developed to ensure responsible AI use, focusing on transparency, validation, and augmentation rather than replacement of human expertise. Surgeons must decide where AI's use is appropriate, questioning whether capability alone justifies adoption. With careful consideration of these challenges, AI has the potential to revolutionize surgical education and foster a new generation of highly skilled and competent surgeons.

1. Introduction

Artificial intelligence (AI) is revolutionizing the healthcare landscape, and the realm of surgical education is no exception. In an age where precision and expertise are paramount, AI has emerged as a transformative ally with the potential to enhance how surgeons are educated and trained. By enabling advanced feedback mechanisms and data-driven insights, AI is paving the way for personalized and adaptive learning environments, creating opportunities to bridge gaps in competency and performance evaluation.[1,2] This technological advancement aligns with the ongoing quest for improved surgical outcomes and patient safety as educators seek innovative methods to better prepare surgical trainees for the challenges of the operating room.[1]

Demonstrations of how AI-driven tools can affect the field of surgery in multiple ways include intraoperative guidance (e.g., AI-based systems could be used during laparoscopic cholecystectomy to identify critical anatomical zones and issue warnings when dissection is carried out in a dangerous area[3,4]), intraoperative context awareness (e.g., prediction of bleeding and alert the surgical team beforehand[5]), and enhanced surgical education (e.g., automated feedback).[6,7] Such advancements could make surgical interventions safer and enhance the learning experience for surgical trainees. These systems, designed to prevent fatal errors, represent the future direction of surgery and will likely become a major component of surgical training.

In this narrative review, we explore the current applications of AI that are likely to shape the education of future surgeons. We delve into the multifaceted role of AI in technical skills education, such as through the enhancement of feedback in surgical simulation and intraoperative applications. We then cover the nascent but growing literature on the use of AI, particularly large language models, in resident selection and evaluation while also addressing the challenges and ethical considerations that accompany the implementation of such powerful technologies.

2. AI-driven feedback in surgical simulation

The incorporation of AI into surgical simulators has initiated advancements in precision and functionality, but its current role in general surgery remains more limited than in other fields, such as ophthalmology.[8] While AI-powered simulators are being developed, most existing simulators focus on preoperative planning and enhancing basic skills training, particularly in laparoscopy and endoscopy. While these techniques are mandatory components of general surgery training,[9] it is crucial to note that many existing simulators do not fully utilize AI-driven feedback or assessment mechanisms. Virtual reality (VR) simulators, which offer immersive environments and adaptive feedback based on user performance, are promising vessels for AI-driven feedback but are not yet essential components for all surgical training programs.[10–12]

AI has the potential to enhance surgical training through quantitative performance assessment and tailored feedback based on individual skill levels. Researchers at McGill University developed a machine learning

This article is part of a special issue entitled: AI_My Thoughts 2024 published in The American Journal of Surgery.

https://doi.org/10.1016/j.amjsurg.2025.116257
Received 7 October 2024; Received in revised form 18 January 2025; Accepted 18 February 2025

algorithm that classifies participants' proficiency during virtual reality-based simulations of hemilaminectomy and brain tumor resection.[12,13] Their "Virtual Operative Assistant" (VOA) assesses skill levels and offers real-time, expert-informed feedback tailored to individual performance benchmarks. While this advancement could transform surgical education in the future, current applications remain limited to specific scenarios. By providing real-time feedback, AI allows trainees to practice independently and learn from their mistakes without constant supervision. This autonomy benefits trainees and experienced surgeons: trainees can hone their skills at their own pace, while experienced surgeons can focus their efforts on higher-level teaching during actual surgeries. In this evolving dynamic, experienced surgeons could observe how trainees incorporate AI feedback into their techniques, enriching the learning process and ensuring that valuable operating room time is used efficiently for deeper mentorship.

A recent cohort study investigated the unintended effects of AI-enhanced competency-based curricula on surgical skill acquisition, revealing both positive and negative outcomes. Fazlollahi et al. (2023) developed and tested the first AI-selected competencies in surgical simulation and categorized the affected metrics into three groups: intrinsic, implicit, and extrinsic.[6] Intrinsic metrics, directly taught by the intelligent system, demonstrated the effectiveness of AI-driven feedback in helping trainees focus on specific learning objectives. Implicit metrics, such as instrument divergence, showed significant improvements even without direct feedback due to their alignment with intrinsic learning goals. Extrinsic metrics, such as healthy tissue removal rates and dominant hand usage,[13,14] were not directly related to learning objectives (in this study, force and acceleration used to remove a brain tumor) and showed both improvement toward and divergence from experienced surgeon benchmarks. Study participants improved their control of instruments but also became slower with their dominant hand. Despite quantitative metrics demonstrating slower movement, human evaluators judged these participants as being more efficient in their movements on their Objective Structured Assessment of Technical Skill rating. This finding raises questions on the relationship between highly granular, AI-driven performance metrics and global assessment of surgical skill. Surgical educators will need to decide whether AI assessment in simulation should prioritize granular metrics or global assessment and weigh how individual decision-making and varying surgical contexts impact these trade-offs across different learning environments.

AI-driven performance evaluation and feedback generation come with limitations. A systematic review of various ML techniques revealed significant challenges, including that simulated tasks may not accurately reflect performance in the operating room, given the complexity of real surgical environments and the difficulty in creating high-fidelity simulation models or situations.[15] Additionally, there is a lack of validity data to evaluate performance, as most of these models are in an early stage of development. There is a need for a substantial amount of data needed to train and validate the algorithms.[16] This highlights the need to standardize surgical data collection and cross-institutional open-source data-sharing policies.

Given the ongoing challenges, limitations, and outstanding questions, careful consideration and thorough cost-benefit analyses are thus necessary when deciding whether to implement intelligent tutoring systems, which are still in their infancy, in surgical training.

3. Intraoperative applications

AI's role in the intraoperative realm is rapidly expanding, with the possibility of offering real-time support and feedback during surgical procedures.[17] Modern AI tools have the ability to monitor and assess surgeons' performance in laparoscopic and endoscopic surgeries, providing concurrent analysis to refine technical skills.[18,19] These systems aim to optimize the alignment between human expert evaluations and AI assessments, offering objective evaluations in critical areas such as suturing and knot tying.[20–22] However, as AI models are often trained on a limited number of expert evaluations, biases can arise, necessitating broader validation efforts to ensure fairness and accuracy in assessments.

AI has enabled the development of visual question-answering (VQA) systems that allow surgical trainees to ask real-time questions during operations. These models, which combine computer vision and language processing, can answer broader questions, such as identifying procedural steps or the state of instruments, including their position, angles, and proximity to anatomical structures.[23,24] As these systems advance, they may provide more detailed, context-specific feedback, becoming valuable tools in environments where resident-to-faculty ratios are high.[25] VQA systems offer the potential to score performance and also provide explanations, helping trainees understand why certain actions are correct or incorrect and how they can improve.

AI has also shown promise in enhancing intraoperative safety by providing alerts and warnings during surgery. AI tools for laparoscopic cholecystectomy can identify safe and unsafe zones to guide trainees.[3] A similar tool utilizes deep learning algorithms to provide guidance to avoid parathyroid gland ischemia during endoscopic thyroid surgery.[26] Additionally, AI-driven systems that monitor vital signs to predict adverse events have been implemented to anticipate complications and provide early warnings.[27,28] Such systems may, in the future, intervene directly during robotic surgeries, preventing potentially harmful maneuvers. While these technologies reduce the burden on trainees and mentors, the potential risks of relying too heavily on AI must be considered. For instance, halting a robotic movement based on AI prediction might interfere with necessary but seemingly risky surgical steps, underscoring the need to balance human judgment and AI intervention.

Many of these real-time AI guidance systems use augmented reality (AR) technology to overlay visual guidance directly onto the surgical field. The development of AI-driven 3D reconstruction technologies, capable of visualizing surgical procedures in real-time, adds another layer of sophistication to intraoperative training. AI-driven segmentation tools offer AR overlays 3D renderings of axial imaging during robotic liver surgeries, enhancing the surgeon's understanding of anatomical structures to assist in targeting lesions and guiding resection margins.[29] This form of immediate spatial orientation allows trainees to view surgeries from multiple angles, enhancing their comprehension of complex procedures. Future innovations, such as dynamic 4D reconstructions, could offer even more detailed insights, enabling trainees to evaluate entire surgical sequences intraoperatively and postoperatively. With advancements in AR glasses and other devices, trainees can expect increasingly immersive learning environments where AI seamlessly integrates into the surgical workflow, providing continuous guidance and support.[30–32]

In summary, AI's contributions to intraoperative safety and competency evaluation will likely influence the future of surgical training. These tools will not only help trainees refine their technical skills but may also ensure safer operating environments by providing real-time guidance and intervention. However, reliance on AI-driven systems raises concerns, particularly the potential for erosion of decision-making and de-skilling of surgeons. Striking a balance between AI assistance and human oversight will be key to enhancing the skills essential for achieving surgical excellence.

4. AI in resident selection

The evolution of resident selection has yielded a paradigm shift from a heavy reliance on quantitative measures, such as Step 1 and Step 2 Clinical Knowledge (CK) scores, toward a more holistic review of applications.[33] The transition to a pass/fail grading system for Step 1 has diminished the weight of this examination in the selection process. While Step 2 CK scores still correlate with in-training examinations and board certification, their predictive power is limited. Studies indicate that although lower Step 2 CK scores are associated with lower performance on the American Board of Surgery In-Training Examination (ABSITE), this metric alone cannot accurately forecast a resident's overall success in residency or clinical practice.[33]

In response to these limitations, governing bodies have recommended a holistic review process for evaluating applicants in all realms of medicine. This approach balances academic attributes with personal experiences to assess an applicant's potential as a future physician. Notably, implementing this approach at a large academic institution led to significant increases in the proportion of women and underrepresented minorities in medicine (UIM) among applicants, alongside improvements in traditional performance metrics like ABSITE scores.[34] AI may contribute to a holistic review process by employing natural language processing (NLP) to analyze personal statements for personality traits and potential fit within surgical teams.[35] AI tools may help identify gender bias in the selection process, as highlighted in reviews examining bias in letters of recommendation for general surgery residency candidates.[36]

Using AI algorithms to screen for personality traits and team compatibility may risk perpetuating existing biases inherent in the selection process. These biases, embedded within historical data and subjective evaluations, can be perpetuated by AI systems unless carefully monitored and audited. A recent study **demonstrated** that AI-selected applicants differed substantially from those chosen by a program director (PD), with only a 7 % overlap between the two groups. While AI expedited the review process, analyzing over 1200 applications in under 12 hours, its reliance on training data introduced biases, such as favoring applicants from US medical schools with higher Step 2 CK scores and more publications. Such findings underscore the risk of perpetuating historical biases in the data used to train these algorithms, which could inadvertently sustain homogeneity in resident cohorts.[37]

AI has also demonstrated the ability to generate personal statements that mimic human-authored content. Studies show that AI-generated statements, particularly those created by large language models like ChatGPT (OpenAI, San Francisco, CA), were not detected as AI by reviewers 56 % of the time.[38] AI-based personal statement review systems may inherit biases in their training data, favoring specific demographics, schools, or linguistic styles while penalizing others.[39] Avoiding such biases requires deliberate human oversight to align selection processes with institutional values. Emphasizing robust auditing systems and active human involvement is essential to mitigate these risks and foster diverse, inclusive, and effective surgical teams.

5. AI in resident evaluation

AI aligns with several priorities set forth by the American Board of Surgery's (ABS) Blue Ribbon Committee II on Surgical Education (BRCII), which recently published a list of competency-based goals to advance surgical education.[40] Integrating defined training goals into daily practice workflows offers an opportunity to leverage AI in evaluating residents' performance. The BRCII highlighted the need for predictive assessments of cognitive, technical, and behavioral competencies. Studies have demonstrated that ML models can match or surpass expert evaluations in determining skill based on video analysis of surgical procedures.[41,42] AI can also identify residents at risk for substandard performance, as demonstrated by work using deep learning to analyze behavioral and motivational characteristics. Such information can enable tailored learning plans and improve trainee performance.[43] These advancements not only provide quantitative measures of technical skill but also streamline feedback mechanisms for trainees. However, as is usually the case with AI, this can only be achieved with a collective effort to collect the necessary data, as current examples are limited to small datasets at a small number of institutions.

There has been a recent shift in the surgical training paradigm toward Entrustable Professional Activities (EPAs), competency-based milestones defining the core components of surgical training. An initial set of 18 core EPAs has been tested and implemented by the American Board of Surgery in response to concerns about traditional certification paradigms that prioritize case volume over demonstrated competence.[44] Even though it is very early in the EPA era, it has been shown that AI could be used further to refine the competencies and identify which themes relate to resident entrustment and, therefore, resident autonomy.[45]

The integration of data-driven systems, such as the AMA-NYU Precision Education Data Lab and the AMA Graduate Profile Report, demonstrates the potential of big data informatics in surgical education. Analyses from the AMA Graduate Profile Report identified discrepancies in opioid prescribing patterns among medical school graduates, leading to a redesign of the pharmacology curriculum at one institution.[46] These tools highlight how proximal outcome measures can validate competency-based training and certification frameworks in real-time, ensuring that educational objectives align with clinical and societal needs. AI integration with such systems promises to enhance resident assessment processes while supporting a national infrastructure for data sharing and evidence-informed policymaking, effectively linking training outcomes to broader public health goals.

Integrating AI into surgical education offers opportunities to enhance trainee assessment and align program outcomes with public health needs. However, this progress is not without significant challenges. Achieving these goals requires addressing data privacy concerns, ensuring equitable access to AI technologies, and fostering collaboration across institutions to build data-sharing frameworks. While the potential benefits are promising, AI's successful implementation will demand sustained efforts, rigorous validation, and a commitment to addressing the ethical, logistical, and technical hurdles that lie ahead.

6. Challenges and ethical considerations

As previously highlighted, AI's integration into surgical education can transform the training paradigm as we know it. However, implementing these advancements comes with significant challenges. In particular, there is a need for large, high-quality datasets for training AI algorithms and ensuring that AI tools can transition seamlessly from simulated environments to the complexities of actual surgeries.[47] The ethical concerns surrounding data privacy and the responsible use of personal health information must be addressed. As AI continues to evolve, regulatory frameworks must be developed to ensure the responsible use of AI in surgical education and practice, thus permitting a future where AI augments rather than replaces the human element in surgery.[47,49]

It is particularly important to balance AI assistance and human oversight. Overreliance on AI for decision-making risks de-skilling surgeons and reducing their ability to manage unforeseen intraoperative events. AI must augment rather than replace human judgment, with surgeons maintaining ultimate control over critical decisions. AI systems should be transparently designed with explainability in mind. The "black-box" nature of some AI algorithms complicates transparency, raising concerns about surgeons' ability to understand, accept, or disagree with AI guidance systems. Without proper oversight and evidence-based validation, these systems could inadvertently endanger patient safety.[48] Additionally, concerns about bias in AI algorithms, especially in resident selection and performance evaluations, require robust validation and transparency to avoid unfair assessments.[49]

While AI presents numerous opportunities, its unscientific application poses significant dangers. Unvalidated or poorly tested AI systems could provide inaccurate recommendations, leading to harmful decisions during surgery. Reliance on AI-generated predictions without thorough clinical validation could result in suboptimal surgical outcomes if critical anatomical features or steps are misidentified.[50,51] By addressing these challenges with comprehensive ethical and regulatory guidelines, AI can transform surgical education while safeguarding human expertise and patient safety.

7. Next steps

To effectively integrate AI into surgical education, ensuring access to high-quality, comprehensive data is essential for driving algorithm development and validation. Key data requirements include video and performance analytics from a wide range of surgical procedures, as well as detailed evaluations of resident behaviors, cognitive skills, and technical proficiency. Data from diverse demographics and educational backgrounds will help address potential algorithmic bias.[52]

There is a pressing need to identify or develop metrics for surgical performance that are educationally and clinically meaningful to guide the development and validation of models.[53] The BRCII recommended establishing a multidisciplinary surgical education council to provide consensus on technology implementation and prospective assessment,[40] and such a council should also offer an organization-wide endorsement of key metrics that AI developers should target to unlock AI-enabled assessment that is meaningful.

Perhaps the most significant challenge lies not in acquiring this data but in determining the extent to which AI should be trusted to perform the tasks envisioned. While AI has the potential to accelerate assessment processes, enhance feedback, and improve selection criteria, the question remains: should AI fully assume these responsibilities (or be utilized at all)? Concerns about transparency, accountability, and the preservation of human judgment are significant.[54]

The ethical implications of allowing AI to influence decisions regarding residency selection, performance evaluation, and patient care require careful consideration. With personal statements and recommendation letters being the few areas of an application where assessment of non-quantitative qualities is possible, we must ask ourselves whether we will abdicate evaluation of those qualities to algorithms and if these sections will offer any evaluation value in the future if applicants and letter writers rely on generative AI. Therefore, comprehensive frameworks are necessary to guide the responsible implementation of AI, balancing innovation with caution.[55]

The decision to expand AI's role in surgical education should be driven by a careful, iterative evaluation process involving diverse stakeholders such as educators, clinicians, engineers, and ethicists. Transparent data collection, validation of algorithms, and ongoing human oversight are crucial to ensuring that AI tools align with the goals of fostering competent, equitable, and compassionate surgeons.[56]

8. Conclusion

AI can potentially transform the delivery of surgical education but should not be seen as a panacea for its complexities. Instead, AI is a complementary tool that supports human engagement, knowledge transfer, and supervision. A key challenge lies in determining the appropriate boundaries of AI's role. Surgeons must identify where AI's use is appropriate and question whether an AI's capability alone justifies its adoption.

A balanced, transparent approach is crucial for responsibly harnessing AI in surgical training to ensure fairness, accountability, and equity. Rigorous validation, human oversight, and transparent data practices can help mitigate algorithmic biases and uphold surgical education's core values. All stakeholders are responsible for shaping regulations that enhance, rather than undermine, training outcomes. The future of surgical education lies not in unquestioningly embracing AI but in integrating it strategically, addressing both its vast potential and risks. With careful, ethical implementation, AI can play a pivotal role in shaping the next generation of surgeons, fostering a training environment that is both innovative and centered on patient care and professional integrity.

CRediT authorship contribution statement

Sebastian Leon: Writing – review & editing, Writing – original draft, Conceptualization. **Sangjoon Lee:** Writing – review & editing, Writing – original draft, Conceptualization. **Juan Esteban Perez:** Writing – review & editing, Writing – original draft, Project administration, Conceptualization. **Daniel A. Hashimoto:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Conceptualization.

Disclosures

DAH was previously a consultant for Johnson and Johnson and Activ Surgical. DAH is vice chair of the board of directors of the Global Surgical AI Collaborative, a non-profit that promotes the democratization of surgical care through the intersection of education, innovation, and technology. He is co-inventor on three pending patents on automated analysis of operative video (US 20240112809, US 20230334868, US 20200170710). The other authors declare no potential conflicts of interest.

Grammarly (San Francisco, CA) was used for spelling, grammar, and word choice.

Funding

DAH is supported by the University of Pennsylvania Linda Pechenik Montague Award.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Daniel Hashimoto reports a prior relationship with Johnson and Johnson Ltd that includes: consulting or advisory. Daniel Hashimoto reports a prior relationship with Activ Surgical that includes: consulting or advisory. Daniel Hashimoto has patent #US 20240112809 pending to Massachusetts General Hospital. Daniel Hashimoto has patent #US 20230334868, pending to Massachusetts General Hospital. Daniel Hashimoto has patent #US 20200170710 pending to Massachusetts General Hospital. DAH is vice chair of the board of directors of the Global Surgical AI Collaborative, a non-profit that promotes the democratization of surgical care through the intersection of education, innovation, and technology. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Thanawala R. The impact of AI on surgery residency programs: improving competency, performance, and the future. *Global Surgical Education - Journal of the Association for Surgical Education*. 2024 2024/05/04;3(1):55.
2. Rubalcava NGC, Kapani N, Martinez Quinones P. How artificial intelligence is expected to transform surgical training. Available from: https://www.facs.org/for-medical-professionals/news-publications/news-and-articles/bulletin/2023/august-2023-volume-108-issue-8/how-artificial-intelligence-is-expected-to-transform-surgical-training/; 2023.
3. Madani A, Namazi B, Altieri MS, et al. Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. *Ann Surg*. 2022 Aug 1;276(2):363–369. PubMed PMID: 33196488. Pubmed Central PMCID: PMC8186165. Epub 20201113.
4. Laplante S, Namazi B, Kiani P, et al. Validation of an artificial intelligence platform for the guidance of safe laparoscopic cholecystectomy. *Surg Endosc*. 2023 Mar;37(3): 2260–2268. PubMed PMID: 35918549. Epub 20220802.
5. Checcucci E, De Cillis S, Amparore D, et al. Artificial Intelligence alert systems during robotic surgery: a new potential tool to improve the safety of the intervention. *Urology Video Journal*. 2023;18:100221, 2023/06/01/.
6. Fazlollahi AM, Yilmaz R, Winkler-Schwartz A, et al. AI in surgical curriculum design and unintended outcomes for technical competencies in simulation training. *JAMA Netw Open*. 2023 Sep 5;6(9):e2334658. PubMed PMID: 37725373. Pubmed Central PMCID: PMC10509729. Epub 20230905. eng.
7. Park JJ, Tiefenbach J, Demetriades AK. The role of artificial intelligence in surgical simulation. *Front Med Technol*. 2022;4:1076755. PubMed PMID: 36590155. Pubmed Central PMCID: PMC9794840. Epub 20221214. eng.
8. Bakshi SK, Lin SR, Ting DSW, et al. The era of artificial intelligence and virtual reality: transforming surgical education in ophthalmology. *Br J Ophthalmol*. 2021 Oct; 105(10):1325–1328. PubMed PMID: 32816750. Epub 20200814. eng.
9. General surgery training requirements - American board of surgery [Accessed October 3, 2024]; Available from: https://www.absurgery.org/get-certified/general-surgery/training-requirements/; August 13, 2024.
10. Bravo J, Wali AR, Hirshman BR, et al. Robotics and artificial intelligence in endovascular neurosurgery. *Cureus*. 2022 Mar;14(3):e23662. PubMed PMID: 35371874. Pubmed Central PMCID: PMC8971092. Epub 20220330.
11. 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. PubMed PMID: 32106247. Pubmed Central PMCID: PMC7046231. Epub 20200227.
12. Winkler-Schwartz A, Bissonnette V, Mirchi N, et al. Artificial intelligence in medical education: best practices using machine learning to assess surgical expertise in virtual reality simulation. *J Surg Educ*. 2019 Nov-Dec;76(6):1681–1690. PubMed PMID: 31202633. Epub 20190613.
13. Yilmaz R, Ledwos N, Sawaya R, et al. Nondominant hand skills spatial and psychomotor analysis during a complex virtual reality neurosurgical task-A case series study. *Oper Neurosurg (Hagerstown)*. 2022 Jul 1;23(1):22–30. PubMed PMID: 35726926. Epub 20220420.
14. Nieboer TE, Sari V, Kluivers KB, et al. A randomized trial of training the non-dominant upper extremity to enhance laparoscopic performance. *Minim Invasive Ther Allied Technol*. 2012 Jul;21(4):259–264. PubMed PMID: 21939399. Epub 20110923.
15. Lam K, Chen J, Wang Z, et al. Machine learning for technical skill assessment in surgery: a systematic review. *NPJ Digit Med*. 2022 Mar 3;5(1):24. PubMed PMID: 35241760. Pubmed Central PMCID: PMC8894462. Epub 20220303.
16. Maier-Hein L, Eisenmann M, Sarikaya D, et al. Surgical data science - from concepts toward clinical translation. *Med Image Anal*. 2022 Feb;76:102306. PubMed PMID: 34879287. Pubmed Central PMCID: PMC9135051. Epub 20211118.
17. Mascagni P, Alapatt D, Sestini L, et al. Computer vision in surgery: from potential to clinical value. *NPJ Digit Med*. 2022 Oct 28;5(1):163. PubMed PMID: 36307544. Pubmed Central PMCID: PMC9616906. Epub 20221028.
18. Gruter AAJ, Van Lieshout AS, van Oostendorp SE, et al. Video-based tools for surgical quality assessment of technical skills in laparoscopic procedures: a systematic review. *Surg Endosc*. 2023 Jun;37(6):4279–4297. PubMed PMID: 37099157. Pubmed Central PMCID: PMC10234871. Epub 20230426.
19. Fukuda H, Ishihara R, Kato Y, et al. Comparison of performances of artificial intelligence versus expert endoscopists for real-time assisted diagnosis of esophageal squamous cell carcinoma (with video). *Gastrointest Endosc*. 2020 Oct;92(4):848–855. PubMed PMID: 32505685. Epub 20200604.
20. Uemura M, Tomikawa M, Miao T, et al. Feasibility of an AI-based measure of the hand motions of expert and novice surgeons. *Comput Math Methods Med*. 2018;2018: 9873273. PubMed PMID: 29686724. Pubmed Central PMCID: PMC5857335. Epub 20180304.
21. Alonso-Silverio GA, Perez-Escamirosa F, Bruno-Sanchez R, et al. Development of a laparoscopic box trainer based on open source hardware and artificial intelligence for objective assessment of surgical psychomotor skills. *Surg Innov*. 2018 Aug;25(4): 380–388. PubMed PMID: 29809097. Epub 20180529.
22. Wang Z, Fey AM. SATR-DL: improving surgical skill assessment and task recognition in robot-assisted surgery with deep neural networks. *Annu Int Conf IEEE Eng Med Biol Soc*. 2018 Jul;2018:1793–1796. PubMed PMID: 30440742.
23. Bai L, Islam M, Seenivasan L, Ren H, eds. *Surgical-vqla: Transformer with Gated Vision-Language Embedding for Visual Question Localized-Answering in Robotic Surgery*. IEEE: IEEE International Conference on Robotics and Automation (ICRA); 2023, 2023.
24. Surgical-vqa: visual question answering in surgical scenes using transformer. In: Seenivasan L, Islam M, Krishna AK, Ren H, eds. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer; 2022.
25. Wang G, Bai L, Nah WJ, et al. Surgical-LVLM: learning to adapt large vision-language model for grounded visual question answering in robotic surgery. *arXiv preprint arXiv: 240510948*; 2024. https://doi.org/10.48550/arXiv.2405.10948.
26. Wang B, Yu JF, Lin SY, et al. Intraoperative AI-assisted early prediction of parathyroid and ischemia alert in endoscopic thyroid surgery. *Head Neck*. 2024 Aug; 46(8):1975–1987. PubMed PMID: 38348564. Epub 20240213.
27. A health 4.0 integrated system for monitoring and predicting patient's health during surgical procedures. In: Arpaia P, Cicatiello M, Benedetto ED, et al., eds. *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. 2020 25-28 May 2020.
28. Singhal M, Gupta L, Hirani K. A comprehensive analysis and review of artificial intelligence in anaesthesia. *Cureus*. 2023 Sep;15(9):e45038. PubMed PMID: 37829964. Pubmed Central PMCID: PMC10566398. Epub 20230911.
29. Giannone F, Felli E, Cherkaoui Z, et al. Augmented reality and image-guided robotic liver surgery. *Cancers*. 2021 Dec 14;13(24). PubMed PMID: 34944887. Pubmed Central PMCID: PMC8699460. Epub 20211214.
30. Samset E, Schmalstieg D, Vander Sloten J, et al. Augmented reality in surgical procedures. *Proc SPIE*. 2008;6806:6806, 2008 03/01.
31. Heinrich F, Huettl F, Schmidt G, et al. HoloPointer: a virtual augmented reality pointer for laparoscopic surgery training. *Int J Comput Assist Radiol Surg*. 2021 Jan; 16(1):161–168. PubMed PMID: 33095424. Pubmed Central PMCID: PMC7822765. Epub 20201023.
32. Zorzal ER, Campos Gomes JM, Sousa M, et al. Laparoscopy with augmented reality adaptations. *J Biomed Inf*. 2020 Jul;107:103463. PubMed PMID: 32562897. Epub 20200618.
33. Jones A, Benns M, Farmer R. Using resident performance on Step 2 to predict surgical residency success. *Surgery*. 2024 Sep 19. PubMed PMID: 39304445. Epub 20240919.
34. Nehemiah A, Roberts SE, Song Y, et al. Looking beyond the numbers: increasing diversity and inclusion through holistic review in general surgery recruitment. *J Surg Educ*. 2021 May-Jun;78(3):763–769. PubMed PMID: 32950431. Epub 20200917.
35. John A, Kavic S. Leveraging artificial intelligence for resident recruitment: can the dream of holistic review be realized? *Artificial Intelligence Surgery*. 2022;12/25(2): 195–206.
36. Sarraf D, Vasiliu V, Imberman B, Lindeman B. Use of artificial intelligence for gender bias analysis in letters of recommendation for general surgery residency candidates. *Am J Surg*. 2021 Dec;222(6):1051–1059. PubMed PMID: 34674847. Epub 20211002.
37. Hassan M, Ayad M, Nembhard C, et al. Artificial intelligence compared to manual selection of prospective surgical residents. *J Surg Educ*. 2025 Jan;82(1):103308. PubMed PMID: 39509905. Epub 20241106.
38. Lum ZC, Guntupalli L, Saiz AM, et al. Can artificial intelligence fool residency selection committees? Analysis of personal statements by real applicants and generative AI, a randomized, single-blind multicenter study. *JB JS Open Access*. 2024 Oct-Dec;9(4). PubMed PMID: 39450246. Pubmed Central PMCID: PMC11498924. Epub 20241024.
39. Mehrabi N, Morstatter F, Saxena N, et al. A survey on bias and fairness in machine learning. *ACM Comput Surv*. 2021;54(6):1–35.
40. Mellinger JD, Brasel K, Elster E, et al. Key issues in surgical residency education: recommendations of the Blue Ribbon II committee residency education subcommittee. *Ann Surg*. 2024 Jul 1. PubMed PMID: 38946537. Epub 20240701.
41. Anh NX, Nataraja RM, Chauhan S. Towards near real-time assessment of surgical skills: a comparison of feature extraction techniques. *Comput Methods Progr Biomed*. 2020 Apr;187:105234. PubMed PMID: 31794913. Epub 20191119.
42. Funke I, Mees ST, Weitz J, Speidel S. Video-based surgical skill assessment using 3D convolutional neural networks. *Int J Comput Assist Radiol Surg*. 2019 Jul;14(7): 1217–1225. PubMed PMID: 31104257. Epub 20190518.
43. Yost MJ, Gardner J, Bell RM, et al. Predicting academic performance in surgical training. *J Surg Educ*. 2015 May-Jun;72(3):491–499. PubMed PMID: 25600356. Epub 20150116.
44. Montgomery KB, Mellinger JD, Lindeman B. Entrustable professional Activities in surgery: a review. *JAMA Surg*. 2024 May 1;159(5):571–577. PubMed PMID: 38477902. Pubmed Central PMCID: PMC11260519.
45. Stahl CC, Jung SA, Rosser AA, et al. Natural language processing and entrustable professional activity text feedback in surgery: a machine learning model of resident autonomy. *Am J Surg*. 2021;221(2):369–375, 2021/02/01/.
46. Richardson J, Santen SA, Mejicano GC, et al. Learner assessment and program evaluation: supporting precision education. *Acad Med*. 2024 Apr 1;99(4S Suppl 1): S64–S70. PubMed PMID: 38166211. Epub 20231228.
47. Snyder KB, Stewart RA, Hunter CJ. *Ethical considerations for the application of artificial intelligence in pediatric surgery*. *AI and Ethics*. 2024 2024/07/24.
48. Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ*. 2020 Mar 25;368:m689. PubMed PMID: 32213531. Pubmed Central PMCID: PMC7190037. Epub 20200325.
49. Colborn K, Brat G, Callcut R. Predictive analytics and artificial intelligence in surgery-opportunities and risks. *JAMA Surg*. 2023 Apr 1;158(4):337–338. PubMed PMID: 36576823. Pubmed Central PMCID: PMC10204072.
50. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg*. 2018 Jul;268(1):70–76. PubMed PMID: 29389679. Pubmed Central PMCID: PMC5995666.
51. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng*. 2018 Oct;2(10):719–731. PubMed PMID: 31015651. Epub 20181010.
52. Satapathy P, Hermis AH, Rustagi S, et al. Artificial intelligence in surgical education and training: opportunities, challenges, and ethical considerations - correspondence. *Int J Surg*. 2023 May 1;109(5):1543–1544. PubMed PMID: 37037597. Pubmed Central PMCID: PMC10389387. Epub 20230501.
53. Hashimoto DA, Dimick JB, Pugh CM. Practical guide to use of simulation and video data. *JAMA Surg*. 2025 Jan 8. PubMed PMID: 39774631. Epub 20250108.
54. Power D. Ethical considerations in the era of AI, automation, and surgical robots: there are plenty of lessons from the past. *Discover Artificial Intelligence*. 2024 2024/ 10/01;4(1):65.
55. Collins JW, Marcus HJ, Ghazi A, et al. Ethical implications of AI in robotic surgical training: a Delphi consensus statement. *Eur Urol Focus*. 2022 Mar;8(2):613–622. PubMed PMID: 33941503. Epub 20210430.
56. Vedula SS, Ghazi A, Collins JW, et al. Artificial intelligence methods and artificial intelligence-enabled metrics for surgical education: a multidisciplinary consensus. *J Am Coll Surg*. 2022 Jun 1;234(6):1181–1192. PubMed PMID: 35703817. Pubmed Central PMCID: PMC10634198. Epub 20220511.

Sebastian Leon[a], Sangjoon Lee[a], Juan Esteban Perez[a], Daniel A. Hashimoto[a,b,*]
[a] *Department of Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA*
[b] *Department of Computer and Information Science, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA*

[*] Corresponding author. University of Pennsylvania Department of Surgery 3400 Spruce Street, 4 Silverstein Pavilion Philadelphia, PA, 19104, USA.
E-mail address: daniel.hashimoto@pennmedicine.upenn.edu (D.A. Hashimoto).