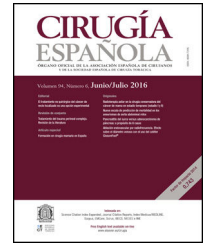




CIRUGÍA ESPAÑOLA

www.elsevier.es/cirugia



Special article

Applications of artificial intelligence in surgery: clinical, technical, and governance considerations

Pietro Mascagni^{a,b,c,*}, Deepak Alapatt^d, Luca Sestini^d, Tong Yu^d, Sergio Alfieri^{b,c}, Salvador Morales-Conde^e, Nicolas Padoy^{a,d}, Silvana Perretta^{a,f,g}

^a IHU Strasbourg, Strasbourg, France

^b Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy

^c Università Cattolica del Sacro Cuore, Rome, Italy

^d University of Strasbourg, CNRS, INSERM, ICube, UMR7357, Strasbourg, France

^e University Hospital Virgen Macarena, University of Sevilla, Seville, Spain

^f IRCAD, Research Institute Against Digestive Cancer, Strasbourg, France

^g Nouvel Hôpital Civil, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

ARTICLE INFO

Article history:

Received 25 April 2024

Accepted 29 April 2024

Keywords:

Artificial intelligence

Surgical data science

Digital surgery

Surgical quality

Quality improvement

ABSTRACT

Artificial intelligence (AI) will power many of the tools in the armamentarium of digital surgeons. AI methods and surgical proof-of-concept flourish, but we have yet to witness clinical translation and value. Here we exemplify the potential of AI in the care pathway of colorectal cancer patients and discuss clinical, technical, and governance considerations of major importance for the safe translation of surgical AI for the benefit of our patients and practices.

© 2024 Published by Elsevier España, S.L.U. on behalf of AEC.

Aplicaciones de la inteligencia artificial en cirugía: consideraciones clínicas, técnicas y de gobernanza

RESUMEN

La inteligencia artificial (IA) impulsará muchas de las herramientas del arsenal de los cirujanos digitales. Los métodos de IA y las pruebas de concepto quirúrgicas están floreciendo, pero aún tenemos que presenciar la traslación y el valor clínicos. Aquí ejemplificamos el potencial de la IA en la vía de atención de los pacientes con cáncer colorrectal y

Palabras clave:

Inteligencia artificial

Ciencia de datos quirúrgicos

Cirugía digital

* Corresponding author.

E-mail address: pietro.mascagni@ihu-strasbourg.eu (P. Mascagni).

<https://doi.org/10.1016/j.ciresp.2024.04.006>

0009-739X/© 2024 Published by Elsevier España, S.L.U. on behalf of AEC.

Calidad quirúrgica
Mejora de la calidad

analizamos consideraciones clínicas, técnicas y de gobernanza de gran importancia para la traslación segura de la IA en cirugía en beneficio de nuestros pacientes y prácticas.

© 2024 Publicado por Elsevier España, S.L.U. en nombre de AEC.

Introduction

The ability to timely model the vast amount of digital data produced during surgical care well sets artificial intelligence (AI) to improve surgical care. Sharing this vision, in the last years several surgeons and computer scientists have joined forces in academic and industrial surgical data science groups to demonstrate the potential of data and analytics AI to positively impact the whole surgical care pathway – from screening to home-monitoring.¹⁻³

Probably one of the most advanced use cases is colorectal cancer (CRC) (Fig. 1). Level one evidence on the value of systems that detect (computer-assisted detection, CAdE) and characterize (CADx) colorectal polyps during screening colonoscopy is mounting, not only with respect to the increase in key performance indicators such as adenoma detection rate⁴ but also through modelling studies showing the potential of these systems to decrease costs and prevent CRC induced deaths.⁵ AI is also being investigated to predict the depth of invasion prior to endoscopic resection⁶ and the risk of lymph node metastasis after endoscopic resection of early CRC,⁷ so as to better identify patients that would benefit from local, endoscopic, resection and surveillance versus patients requiring surgery. In the case of endoscopic resection, AI could help safely performing endoscopic submucosal dissection, as recently shown by a proof-of-concept system automatically segmenting colonic layers and vessels to guide endoscopists.⁸

When surgery is needed, AI could help us stratify patients to tailor the surgical approach. For instance, models analyzing patient and disease data to predict the status of inferior mesenteric artery lymph nodes in left-sided CRC⁹ could inform on whether to perform a high ligation with oncological benefits or to preserve the left colic artery to reduce the risk of anastomotic leakages. Intraoperatively, AI systems approaching commercialization promise to guide surgeons toward the right plane of dissection,^{10,11} potentially helping preserve nerves for better functional outcomes or preventing feared ureteral injuries. Collaterally, AI systems already in clinical use might help anesthetists predict and prevent intraoperative hypotensive events.¹² Postoperatively, advanced patient monitoring could provide early warnings of patients deterioration needing urgent care or fast-tracking home discharge, were patient recovery could be monitored by wearables.^{13,14}

However, it is still the early days of AI in surgery and several technical, clinical and governance considerations still need to be addressed before data and analytics show their full value to patients, surgeons, and healthcare systems.¹⁵

Here we briefly analyze open challenges and discuss possible solutions.

Clinical considerations

We are only just starting to study surgical quality, its impact on patients' outcomes, and how to improve it, especially for

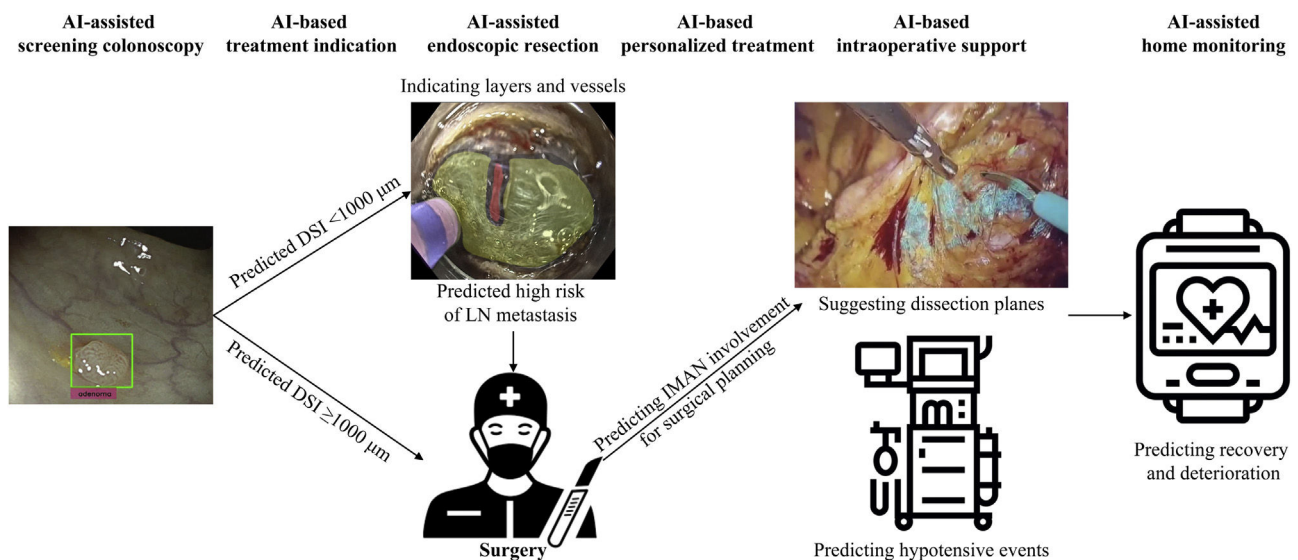


Fig. 1 – AI for CRC care pathway. The AI-assisted screening colonoscopy image is adapted from Hassan et al.³⁶ The AI-assisted endoscopic resection image indicating layers and vessels is adapted from Ebigo et al.⁸ The AI-based intraoperative support image suggesting dissection planes is adapted from a live demo of Eureka (Anaut, Japan) performed by Salvador Morales Conde at the XLII Curso de Actualización en Cirugía, February 7-9, 2024, in Seville, Spain. DSI: depth of submucosal invasion; LN: lymph node; IMAN: inferior mesenteric artery nodes.

Clinical considerations	
Reality	Possible solutions
Surgery is still a black-box	OR data with outcomes AI-derived insights
Robust surgical evidence is hard to obtain	Pragmatic trial designs Comprehensive registries
Surgery resist to change	Well-designed UX Implementation support

Fig. 2 – Clinical considerations around AI in surgery.

what regards the intraoperative phase of surgical care. Arguably, the need to collect and consistently analyze (i.e., annotate) surgical data such as procedural videos and relative outcomes to develop AI models is in itself showing to be the first collateral value induced by the vast interest in surgical AI. For example, while we were collecting and analyzing data to develop an AI for the automatic assessment of the critical view of safety (CVS) in cholecystectomy, we found out that this universally recommended safety step was correctly implemented in as low as 16% of the cases. Such insights led to proposing 5-second rule,¹⁶ a brief time-out to verify the CVS with assistants. In a before vs after we found out that such a simple quality improvement intervention triplicates the rate of clinical implementation of the CVS,¹⁶ suggesting a strategy to fill the guidelines to surgical practice implementation gap. However, most intraoperative events and dynamics still need to be well studied. AI can certainly help by providing quantitative analysis of data at scale, but the surgical community needs to play a central role in defining problems and solutions first. Formalizing key performance measures and indicators, a strategy used in closely related disciplines such as gastrointestinal endoscopy, and defining standard operating procedures, protocols to effectively deal with intraoperative events, could help quantitatively assess and consistently improve our practice. Only then we can develop AI to help solve real surgical problems.

AI is not a pill, and we cannot study its effect on surgical care using the same study designs optimized and popularized by pharmacological research. Most AI-based systems don't act directly on patients, but rather act on how care is delivered. For this reason, if we were to test the effect of an AI-based quality improvement initiative on patient outcomes, performing a classical trial in which the same surgeon receives (intervention group) or doesn't receive (control group) AI-assistance based on the randomization status of the patient would be at high risk of bias. In this specific case, a surgeon that receives AI-support for one patient and no support for the next one will likely still do better with the second patient. The effect of the intervention could spill onto the control leading to a so-called contamination bias, a typical risk of service-delivery studies. More in general, if the intervention acts on the delivery of care rather than the patient, then we should control for baseline surgeon's characteristic and, if necessary, randomize at this stage. For the above reasons cluster randomized trials with baseline periods, a type of pragmatic study design allowing to randomize groups (i.e., clusters) of patients (e.g., all patients treated by the same surgeon, in the same department, at the

same hospital, etc.) and including both a before-after and a parallel control, represent a valid study design option for AI-based interventions.^{17,18} Alternatively, well thought-out registries recording, for instance, also if the AI-based system was utilized or if automated recommendations were followed or not, could provide evidence on the real-world and over-time value of such system.

Being an instrument in the hands of caregivers, the value of an AI-based system is intrinsically dependent on the way clinicians interact with it and use it. For instance, endoscopists might be induced into a false sense of comfort when using AI and decrease the quality of their screening colonoscopy, as suggested by pragmatic, real-world trials.^{19,20} In addition, AI-systems may be used differently from how manufacturers and regulators intended, as recently shown in radiology.²¹ Purposely defined educational initiatives targeting surgical staff willing to implement these new tools into practice could surely help. In addition, the above suggested registries as well as other monitoring initiatives could help pick up and measure usage drifts requiring continuous education or novel regulatory approvals.

Fig. 2 summarizes clinical considerations.

Technical considerations

AI models, especially the most sophisticated ones, are often referred to as black box. Given the high number of parameters and the intricate interplay of those, the inner working of models such as deep neural networks is hard to decode. While being a very interesting research problem, at first explainability doesn't seem an application problem – at the end of the day, doctors and regulators often don't know the exact mechanisms of drugs we use in every day clinical practice. However, the explainability problem is of major importance for the safe application of AI systems in clinical and surgical practices because it limits our ability to define the system's indications and contraindications to use. As a matter of fact, it is currently hard to estimate if an AI system will work for a specific patient population or in a specific setting if the dataset used for training and testing the AI-model doesn't represent it. The current solution is to restrict indications only to the development population. However, this is very hard to control and limits the potential of AI to prevent major but rare and underrepresented conditions. On the one hand, the computer science community is actively searching for solutions to the explainability problem²² and adding guardrails such as

Technical considerations	
Reality	Possible solutions
AI is a black-box	Explainable or self-limiting AI Large, diverse and representative datasets
AI can fail	Sequential, hierarchical modelling using deep neural networks and expert-systems
Hospitals are not AI-ready	System integration AI-native devices

Fig. 3 – Technical considerations around AI in surgery.

systems capable of flagging when they are not likely to deliver an accurate prediction.²³ On the other hand, the clinical community is trying to address this problem collecting and openly sharing large, diverse, and representative datasets on which AI systems can be trained and test. For instance, with the Society of Gastrointestinal and Endoscopic Surgeons (SAGES) we are organizing the Critical View of Safety Challenge, an initiative collecting more than 1000 surgical videos from all over the globe and gathering clinically meaningful annotations so as to release a large and representative dataset to benchmark AI solutions for safe laparoscopic cholecystectomy (<https://www.cvschallenge.org/>).

AI predictions will never be 100% accurate. No matter how well we train a model, this could always fail, especially when analyzing cases not well represented in the development dataset, as discussed above. In surgery as well as in other high-stake applications of AI, inaccuracies can have high human and financial cost and risk to slow the uptake of such a powerful technology. Here, emulating how we humans tackle complex tasks could help developing a robust AI system. When facing nontrivial problems, we usually break them down into a sequence of smaller problems, cross solutions, and reason through them to solve the bigger picture. Similarly, we can design AI systems that solve problems through sequential, hierarchical modelling. Is saying, such systems would be composed by a set of AI models, each solving narrower tasks, crossing information with one another, to solve down-stream tasks. For example, if we want to build an AI to recognize surgical actions, we might want models to detect surgical tools²⁴ and segment anatomy²⁵ first. Surgical tool and anatomy information would then be used to model surgical actions, such as “hook dissecting the cystic duct”.^{26,27} In addition to simplifying the modelling, such a modular approach could allow to repurpose AI models for different surgical use cases, for instance utilizing the above surgical tool detector to infer metrics of surgical skills.²⁸ Finally, sequential modelling could also entail using deep neural networks together with more controllable rule-based expert systems. A combination of deep and rule-based modelling was used in EndoDigest, a computer vision platform to selectively video document critical events of surgical procedures that showed good performance even when some of its AI predictions were wrong²⁹ and when tested on very diverse data coming from 4 centers.³⁰ Altogether, we believe that a sequential, hierarchical modelling strategy using both powerful deep neural

networks and controllable expert systems might render AI system more performant, robust, scalable, and explainable.³¹

Most hospital information technology (IT) infrastructure and medical devices are not AI-ready. AI is computationally intensive and requires specialized hardware (e.g., graphic processing unit, GPU), especially the most powerful deep neural networks powering most of today state-of-the-art image and text applications. Furthermore, AI-powered system needs to interface with diverse hospital IT systems and/or closed, hardly interoperable, medical devices for fishing data and deliver predictions. These challenges are currently being tackled in several ways. One solution would be to add integration and compute to hospital infrastructure, a strategy enabling the full value of an integrated system but requiring major investments and buy-in (e.g., Caresyntax Integrated OR, Karl Storz OR1). On the other end of the spectrum, several players are incorporating AI capabilities in their medical devices or in stand-alone, at times vendor-neutral, units (e.g., Siemens Healthineers YSIO X.pree and Medtronic GI Genius, respectively). This approach is more rapid and could confer a market hedge, but risks hinder the full potential of digital health buy further siloing health information.

Fig. 3 summarizes technical considerations.

Governance considerations

AI models are built using patient data, not only companies' hardware and know-how. This poses novel questions on how to promote innovations that could benefit patients and healthcare systems while preserving stakeholders' rights, starting from patients privacy and ownership. Regulations on the management of patients' data such as the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA) act in the United States of America exist. However, these are not necessarily clear on the requirements for the compliant development, sharing, and use of datasets for AI development, hindering innovation that could benefit society. Researchers and companies are tackling this challenge by developing pipelines and AI methods to de-identify data and/or bypass the need to share data. Endoscopic videos guiding minimally invasive procedures are being de-identified by removing file meta-data (i.e., data included in the video file such as the day and time of the recording) and using AI models capable of

Governance considerations	
Reality	Possible solutions
Surgical data governance is unclear	Data de-identification Distributed learning and evaluation
Regulatory for AI-powered devices is in its infancy	Open dialogue among stakeholders for supporting safe innovation
Medico-legal liability in AI-assisted practice	Pragmatic approach evaluating AI autonomy and design

Fig. 4 – Governance considerations around AI in surgery.

detecting non-endoscopic, out-of-body images that could potentially show identifiable information (e.g., patients and staff faces, operating room clock, etc.).^{32,33} De-centralized approaches such as federated learning and evaluation that share models across centers rather than centralizing data are currently being studied for surgical AI development.³⁴

Most AI-based applications of clinical utility are classified and regulated as software as medical devices (SaMD). This requires defining an intended use (i.e., the clinical scope the application serves) and providing evidence supporting safety and claims based on risk analysis and classification. Regulatory needs shouldn't be seen as a bottleneck but as a necessary step in the clinical translation of safe and effective solutions. However, intended use doesn't often match the actual clinical use of the AI-powered SaMD, as recently shown in radiology,²¹ and regulators as well as manufacturers are often intimidated by the uncertainty around AI. This requires collaborations between regulators and stakeholders to guarantee efficient and transparent processes based on a common understanding of the importance of a sound audit for all stakeholder, patients as well as businesses.

Finally, a major bottleneck to the adoption of AI in clinical and surgical practices is the uncertainty around medico-legal liability. Who's responsible for a suboptimal outcome when an AI-powered application is used? What if a surgeon follows a wrong recommendation or ignores a correct prediction? We need solid answers to these questions if we believe in the potential of AI in healthcare. While certainly law makers will clarify regulations and regulate novel situations introduced by autonomous systems, we believe that a pragmatic approach answers most of the current use cases. Current AI-powered applications in surgery as well as most AI applications in medicine do not act or make decisions autonomously, but rather estimate information with a certain degree of confidence for aiding clinical decision making and actions. This kind of assistance is comparable to that of many other non-AI medical devices routinely used to inform clinical practices. Like for classical medical devices, design choices influence the risk of AI-powered applications. For instance, if we were to use DeepCVS, an AI to assess the achievement of the critical view of safety in laparoscopic cholecystectomy,³⁵ deploying it for intraoperative guidance (i.e., DeepCVS informs when it is safe to divide the cystic duct) or an automated warning (i.e., DeepCVS warns surgeons if they attempt to divide the cystic duct when it is not safe) would entail very different risks. In the first design a wrong prediction (e.g., DeepCVS giving the green light to proceed when the critical view of safety is not

achieved) could induce a bile duct injury while in the second design a wrong prediction (e.g., DeepCVS warning the surgeon when the critical view of safety is actually achieved) could, at worst, slightly increase procedural times.

Fig. 4 summarizes governance considerations.

Re-imaging surgery together

While surely incomplete, we hope that the above clinical, technical, and governance considerations could help spark discussions in the surgical data science community. Together with the continued identification of new surgical use cases, development of AI methods, and proof-of-concept of surgical AI applications, we believe such multidisciplinary discussions will get as closer to revealing the true value of AI applications in surgery.

REFERENCES

1. Maier-Hein L, Vedula SS, Speidel S, Navab N, Kikinis R, Park A, et al. Surgical data science for next-generation interventions. *Nat Biomed Eng.* 2017;1:691–6.
2. Maier-Hein L, Eisenmann M, Sarikaya D, März K, Collins T, Malpani A, et al. Surgical data science – from concepts toward clinical translation. *Med Image Anal.* 2022;76102306.
3. Lam K, Abràmoff MD, Balibrea JM, Bishop SM, Brady RR, Callcut RA, et al. A Delphi consensus statement for digital surgery. *NPJ Digit Med.* 2022;5:1–9.
4. Hassan C, Spadaccini M, Mori Y, Foroutan F, Facciorusso A, Gkolfakis P, et al. Real-time computer-aided detection of colorectal neoplasia during colonoscopy. *Ann Intern Med.* 2023;176:1209–20.
5. Areia M, Mori Y, Correale L, Repici A, Bretthauer M, Sharma P, et al. Cost-effectiveness of artificial intelligence for screening colonoscopy: a modelling study. *Lancet Digital Health.* 2022;4:e436–44.
6. Bai J, Liu K, Gao L, Zhao X, Zhu S, Han Y, et al. Computer-aided diagnosis in predicting the invasion depth of early colorectal cancer: a systematic review and meta-analysis of diagnostic test accuracy. *Surg Endosc.* 2023;37:6627–39.
7. Ichimasa K, Kudo S, Lee JWJ, Nemoto T, Yeoh KG. Artificial intelligence-assisted treatment strategy for T1 colorectal cancer after endoscopic resection. *Gastrointestinal Endoscopy.* 2023;97:1148–52.
8. Ebigbo A, Mendel R, Scheppach MW, Probst A, Shahidi N, Prinz F, et al. Vessel and tissue recognition during third-space endoscopy using a deep learning algorithm. *Gut.* 2022;71:2388–90.

9. Wang X, Zheng Z, Xie Z, Yu Q, Lu X, Zhao Z, et al. Development and validation of artificial intelligence models for preoperative prediction of inferior mesenteric artery lymph nodes metastasis in left colon and rectal cancer. *Eur J Surg Oncol.* 2022;48:2475–86.
10. Kolbinger FR, Bodenstedt S, Carstens M, Leger S, Krell S, Rinner FM, et al. Artificial Intelligence for context-aware surgical guidance in complex robot-assisted oncological procedures: an exploratory feasibility study. *Eur J Surg Oncol.* 2023;106996.
11. Ryu S, Goto K, Imaizumi Y, Nakabayashi Y. Laparoscopic colorectal surgery with anatomical recognition with artificial intelligence assistance for nerves and dissection layers. *Ann Surg Oncol.* 2024;31:1690–1.
12. Wijnberge M, Geerts BF, Hol L, Lemmers N, Mulder MP, Berge P, et al. Effect of a machine learning-derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: the HYPE randomized clinical trial. *JAMA.* 2020;323:1052–60.
13. Knight SR, Ng N, Tsanas A, Mclean K, Pagliari C, Harrison EM. Mobile devices and wearable technology for measuring patient outcomes after surgery: a systematic review. *NPJ Digit Med.* 2021;4:1–14.
14. Wells CI, Xu W, Penfold JA, Keane C, Gharibans AA, Bissett IP, et al. Wearable devices to monitor recovery after abdominal surgery: scoping review. *BJS Open.* 2022;6zrac031.
15. Mascagni P, Alapatt D, Sestini L, Altieri MS, Madani A, Watanabe Y, et al. Computer vision in surgery: from potential to clinical value. *NPJ Digit Med.* 2022;5:1–9.
16. Mascagni P, Rodríguez-Luna MR, Urade T, Felli E, Pessaux P, Mutter D, et al. Intraoperative time-out to promote the implementation of the critical view of safety in laparoscopic cholecystectomy: a video-based assessment of 343 procedures. *J Am Coll Surg.* 2021;S1072-7515(21).
17. McClure GR, Belley-Cote EP, Spence J, Fu Lee S, Whitlock RP. Selecting the optimal level of clustering: an approach to trial design decision making. *J Am Coll Surgeons.* 2020;231:397.
18. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ.* 2015;350:h391.
19. Levy I, Bruckmayer L, Klang E, Ben-Horin S, Kopylov U. Artificial intelligence-aided colonoscopy does not increase adenoma detection rate in routine clinical practice. *Am J Gastroenterol.* 2022;117:1871.
20. Ladabaum U, Shepard J, Weng Y, Desai M, Singer SJ, Mannalithara A. Computer-aided detection of polyps does not improve colonoscopist performance in a pragmatic implementation trial. *Gastroenterology.* 2023;164481-483.e6.
21. van Leeuwen KG, Hedderich DM, Harvey H, Schalekamp S. How AI should be used in radiology: assessing ambiguity and completeness of intended use statements of commercial AI products. *Insights Imaging.* 2024;15:51.
22. van der Velden BHM, Kuijff HJ, Gilhuijs KGA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med Image Anal.* 2022;79102470.
23. Guha Roy A, Ren J, Azizi S, Loh A, Natarajan V, Mustafa B, et al. Does your dermatology classifier know what it doesn't know? Detecting the long-tail of unseen conditions. *Med Image Anal.* 2022;75102274.
24. Nwoye CI, Mutter D, Marescaux J, Padoy N. Weakly supervised convolutional LSTM approach for tool tracking in laparoscopic videos. *Int J Comput Assist Radiol Surg.* 2019;14:1059–67.
25. Murali A, Alapatt D, Mascagni P, Vardazaryan A, Garcia A, Okamoto N, et al. Encoding Surgical Videos as Latent Spatiotemporal Graphs for Object and Anatomy-Driven Reasoning. In: Greenspan H, Madabhushi A, Mousavi P, Salcudean S, Duncan J, Syeda-Mahmood T, et al., editors. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023 Cham: Springer Nature Switzerland*; 2023; p. 647–57.
26. Nwoye CI, Gonzalez C, Yu T, Mascagni P, Mutter D, Marescaux J, et al. Recognition of Instrument-Tissue Interactions in Endoscopic Videos via Action Triplets. In: Martel AL, Abolmaesumi P, Stoyanov D, Mateus D, Zuluaga MA, Zhou SK, et al., editors. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020 Cham: Springer International Publishing*; 2020; p. 364–74.
27. Nwoye CI, Yu T, Sharma S, Murali A, Alapatt D, Vardazaryan A, et al. CholecTriplet2022: show me a tool and tell me the triplet — An endoscopic vision challenge for surgical action triplet detection. *Med Image Anal.* 2023;89102888.
28. Lavanchy JL, Zindel J, Kirtac K, Twick I, Hosgor E, Candinas D, et al. Automation of surgical skill assessment using a three-stage machine learning algorithm. *Sci Rep.* 2021;11:5197.
29. Mascagni P, Alapatt D, Urade T, Vardazaryan A, Mutter D, Marescaux J, et al. A computer vision platform to automatically locate critical events in surgical videos: documenting safety in laparoscopic cholecystectomy. *Ann Surg.* 2021;274:e93.
30. Mascagni P, Alapatt D, Laracca GG, Guerriero L, Spota A, Fiorillo C, et al. Multicentric validation of EndoDigest: a computer vision platform for video documentation of the critical view of safety in laparoscopic cholecystectomy. *Surg Endosc.* 2022.
31. Mascagni P, Alapatt D, Urade T, Vardazaryan A, Mutter D, Marescaux J, et al. Response to comments on: a computer vision platform to automatically locate critical events in surgical videos: documenting safety in laparoscopic cholecystectomy. *Ann Surg.* 2022;276:e637.
32. Lavanchy JL, Vardazaryan A, Mascagni P, Mutter D, Padoy N. Preserving privacy in surgical video analysis using a deep learning classifier to identify out-of-body scenes in endoscopic videos. *Sci Rep.* 2023;13:9235.
33. De Backer P, Simoons J, Mestdagh K, Hofman J, Eckhoff JA, Jobczyk M, et al. Privacy-proof live surgery streaming: development and validation of a low-cost, real-time robotic surgery anonymization algorithm. *Ann Surg.* 2024. <http://dx.doi.org/10.1097/SLA.0000000000006245>.
34. Kassem H, Alapatt D, Mascagni P, Karargyris A, Padoy N. Federated cycling (FedCy): semi-supervised federated learning of surgical phases. *IEEE Transactions on Medical Imaging.* 2023;42:1920–31.
35. Mascagni P, Vardazaryan A, Alapatt D, Urade T, Emre T, Fiorillo C, et al. Artificial intelligence for surgical safety: automatic assessment of the critical view of safety in laparoscopic cholecystectomy using deep learning. *Annals of Surgery.* 2022;275:955–61.
36. Hassan C, Balsamo G, Lorenzetti R, Zullo A, Antonelli G. Artificial intelligence allows leaving-in-situ colorectal polyps. *Clin Gastroenterol Hepatol.* 2022;20:2505–13. e4.