medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Exploring the impact of an artificial intelligence-based intraoperative image navigation system in laparoscopic surgery on clinical outcomes: A protocol for a multicenter randomized controlled trial

Daichi Kitaguchi[1,2,3], Nozomu Fuse[4], Masashi Wakabayashi[5], Norihito Kosugi[1,2], Yuto Ishikawa[1,2], Kazuyuki Hayashi[1,2], Hiro Hasegawa[2,3], Nobuyoshi Takeshita[2], Masaaki Ito[2,3*]

[1]AI Digital Device Innovation Office, National Cancer Center Hospital East, Chiba, Japan

[2]Department for the Promotion of Medical Device Innovation, National Cancer Center Hospital East, Chiba, Japan

[3]Department of Colorectal Surgery, National Cancer Center Hospital East, Chiba, Japan

[4]Clinical Research Support Office, National Cancer Center Hospital East, Chiba, Japan

[5]Biostatistics Division, Center for Research Administration and Support, National Cancer Center Hospital East, Chiba, Japan

***Corresponding author:**

Masaaki Ito, MD, PhD

Department for the Promotion of Medical Device Innovation

Department of Colorectal Surgery

National Cancer Center Hospital East

6-5-1, Kashiwanoha, Kashiwa, Chiba 277-8577, Japan

E-mail: maito@east.ncc.go.jp, Tel: +81-4-7133-1111, Fax: +81-4-7130-0199

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Background: In the research field of artificial intelligence (AI) in surgery, there are many open questions that must be clarified. Well-designed randomized controlled trials (RCTs) are required to explore the positive clinical impacts by comparing the use and non-use of AI-based intraoperative image navigation. Therefore, herein, we propose the "ImNavi" trial, a multicenter RCT, to compare the use and non-use of an AI-based intraoperative image navigation system in laparoscopic surgery.

Methods: The ImNavi trial is a Japanese multicenter RCT involving 1:1 randomization between the use and non-use of an AI-based intraoperative image navigation system in laparoscopic colorectal surgery. The participating institutions will include three high-volume centers with sufficient laparoscopic colorectal surgery caseloads (>100 cases/year), including one national cancer center and two university hospitals in Japan. Written informed consent will be obtained from all patients. Patients aged between 18 and 80 years scheduled to undergo laparoscopic left-sided colorectal resection will be included in the study. The primary outcome is the time required for each target organ, including the ureter and autonomic nerves, to be recognized by the surgeon after its initial appearance on the monitor. Secondary outcomes include intraoperative target organ injuries, intraoperative complications, operation time, blood loss, duration of postoperative hospital stay, postoperative complications within 30 days, postoperative male sexual dysfunction 1 month after surgery, surgeon's confidence in recognizing each target organ, and the postoperative fatigue of the primary surgeon.

Discussion: The impact of AI-based surgical applications on clinical outcomes beyond numerical expression will be explored from a variety of viewpoints while evaluating quantitative items, including intraoperative complications and operation time, as

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

secondary endpoints. We expect that the findings of this RCT will contribute to advancing research in the domain of AI in surgery.

Trial registration: The trial was registered at the University Hospital Medical Information Network Center (https://www.umin.ac.jp/ctr/index-j.html) on March 28th, 2023 under trial ID: UMIN000050701.

Keywords

Artificial intelligence, computer vision, semantic segmentation, intraoperative image navigation, computer-assisted surgery

Background

Surgical volumes are large and show an increasing trend worldwide [1]; nevertheless, it is estimated that 143 million additional surgical procedures are required annually to save lives and prevent disabilities [2]. Surgery is an essential aspect of healthcare and is associated with increased life expectancy [1]; however, inter-institutional disparities in surgical skill levels remain. The introduction of minimally invasive surgery (MIS), including endoscopic surgical approaches, has made surgery more effective [3]. However, surgery has become more complex, and achieving basic surgical skill levels, which are directly linked to postoperative clinical outcomes, has become more difficult [4, 5]. Improvements in the safety and efficiency of surgery must be a major component in strengthening healthcare systems, and technological innovation may be one of the solutions to achieve this goal.

Computer vision (CV) based on artificial intelligence (AI), particularly deep learning using convolutional neural networks, is a recent technological innovation.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

AI-based CV is a field of computer science that enables AI to extract meaningful information from digital images and videos, and to process and make recommendations based on that information [6]. In recent years, AI-based CV has achieved remarkable success in image recognition tasks in the field of medical image diagnosis, such as radiology [7-9], pathology [10, 11], gastroenterology [12], and ophthalmology [13-15]. In the field of surgery, several AI-based CV applications for MIS, specifically semantic segmentation-based intraoperative image navigation, have been developed based on the nature of MIS, which relies heavily on the visual information provided by surgical endoscopes [16-18]. These are expected to reduce intraoperative anatomical cognitive errors. However, the majority of studies in this field are still in the proof-of-concept stage.

In the research field of AI in surgery, many open questions regarding how AI-based intraoperative image navigation affects surgeon performance exist, including whether it can achieve some improvement in clinical outcomes, and whether it works properly in diverse environments without overfitting. To address these questions, well-designed randomized controlled trials (RCTs) are required to explore the positive clinical impacts of AI-based intraoperative image navigation. Therefore, we propose the “ImNavi” trial, a multicenter RCT, to compare the use and non-use of an AI-based intraoperative image navigation system in laparoscopic surgery.

Methods/Design

Participants, interventions, and outcomes

Trial design

The ImNavi trial is designed as a Japanese multicenter RCT with 1:1 randomization

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

between the use and non-use of an AI-based intraoperative image-navigation system in laparoscopic surgery. The study protocol was prepared according to the reporting guidelines of the Standard Protocol Items: Recommendations for Interventional Trials - Artificial Intelligence extension (SPIRIT-AI) [19]. The SPIRIT-AI checklist is presented in Additional File 1.

Study setting and recruitment

This hospital-based study will be conducted in Japan. Eligible candidates will be identified from patients referred to a colorectal surgeon or colorectal cancer multidisciplinary team based on colonoscopy, computed tomography, or magnetic resonance imaging findings, and eligibility will be confirmed after a review of the study criteria. The target surgical procedures in this study are left-sided colorectal resections, including left hemicolectomy, sigmoidectomy, anterior resection, intersphincteric resection, Hartmann's procedure, and abdominoperineal resection. The target organs for semantic segmentation are the ureter and autonomic nerves, including the hypogastric nerves and the aortic plexus. The AI-based intraoperative image navigation system used in this study is based on an existing semantic segmentation algorithm that we developed [20, 21]. The accuracy and performance of both the ureter recognition model (UreterNet) and the autonomic nerve recognition model (NerveNet) have been validated in our previous study [20]. Alienware x15 R2 (Dell Technologies Inc., Round Rock, Texas, U.S.) was used as the computational system in this study.

Eligibility criteria

Patients aged between 18 and 80 years scheduled to undergo laparoscopic left-sided

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

colorectal resection and who provided written informed consent will be included in the ImNavi trial. Long-term outcomes will not be included in the endpoints; therefore, patients with any colorectal disease will be eligible. Patients with emergent surgery, expected severe intra-abdominal adhesion, significant anatomical anomaly, and those deemed unsuitable by the doctor in charge will be excluded. The trial flow diagram is illustrated in Figure 1.

Interventions

The time required to recognize each target organ, including the ureter and autonomic nerves, after its initial appearance on the monitor will be compared between the surgeons with and without an AI-based intraoperative image navigation system. The initial time of appearance of each target organ will be retrospectively determined by two or more judges observing via video. The system output, semantic segmentation pixels, will be superimposed on a sub-monitor placed adjacent to the main monitor, as illustrated in Figure 2, without influencing the display of the main monitor. The surgeons will decide the time and frequency of viewing the sub-monitor. The primary surgeon will be asked to orally state that they recognize each target organ during surgery. All intraoperative statements made by the primary surgeon will be recorded by a microphone placed near the operating table. The intraoperative audio will be collated with the intraoperative video, and the difference in time points between the initial appearance on the monitor and the primary surgeon's recognition (vocalization) will be calculated for each target organ.

Outcomes

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The primary outcome is the time required for each target organ, including the ureter and autonomic nerves, to be recognized by the surgeon after its initial appearance on the monitor. Secondary outcomes include intraoperative target organ injuries; intraoperative complications; total operation time; operation time for each surgical step divided by existing definitions [22]; blood loss; duration of postoperative hospital stay; postoperative complications within 30 days; postoperative male sexual dysfunction 1 month post-surgery; the surgeon's confidence in recognising each target organ; and postoperative fatigue of the primary surgeon, according to the Piper Fatigue Scale-12 (PFS-12) [23]. The severity of all intraoperative and postoperative complications, according to the Common Terminology Criteria for Adverse Events and the Clavien–Dindo classification, will also be collected as secondary outcomes.

Participant timelines

Before surgery, written informed consent will be obtained from all patients and their primary surgeons, and patients will be allocated to their respective groups using an electronic data capture (EDC) system. A preoperative report, including patient, disease, and primary surgeon background information, will be provided using case report forms (CRFs). After surgery, an intraoperative report on the primary surgeon's surgical information and postoperative fatigue, according to the PFS-12, will be provided using CRFs and questionnaires, respectively. One month after surgery, the postoperative course and outcome reports, including postoperative complications and date of discharge, will be provided using the CRFs. This study will not regulate modalities or intervals between examinations for each participating institution, or intervene in each institution's routine clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Sample size

Based on non-published data from previous studies at our institution, the mean time (± standard deviation) required by surgeons to recognize each target organ without using the AI-based intraoperative image navigation system after its initial appearance on the monitor is estimated to be 6.0 (± 9.0) s. To detect a 2-s decrease with the AI-based system, using a power of 80% and a one-sided significance level of 0.15 in the Wilcoxon rank-sum test, 38 patients in each group (76 patients in total) would be required for this RCT. After accounting for dropouts and ineligibility, an additional 14 patients will be included. Thus, the target sample size would be 90 patients.

Assignment of interventions

Once eligibility is established, patients will be allocated to either the use or non-use of AI-based intraoperative image navigation groups. Randomization, performed by computers through the Internet (https://www2.epoc-ncc.net/), will be adjusted using the minimization method with a random component to balance the groups, participating institution, sex (male versus female), and experience of the primary surgeon (qualified versus unqualified by the Endoscopic Surgical Skill Qualification System of the Japan Society for Endoscopic Surgery) [24]. Patients will be randomized in a 1:1 ratio. Data will be analyzed on an ‘intention-to-treat’ basis when patients are not subjected to a randomized treatment modality.

Data collection, management, and analysis

Data collection will be carried out using CRFs and a validated PFS-12

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

questionnaire. All data will be entered into an EDC system. Regular data quality checks will be performed annually in accordance with the Quality Management Plan. All data will be handled following the Ethical Guidelines for Medical and Biological Research Involving Human Subjects. Data backups will be stored in secure fireproof locations, and test restorations will be performed regularly. After completion of the trial, all essential trial documentation and source documents, including signed informed consent forms and copies of CRFs, will be securely retained for at least 5 years.

Statistical analysis

Baseline numerical data will be described as the mean (± standard deviation) or median (range). Baseline categorical data will be presented as percentages. All comparative analyses will be conducted on an 'intention-to-treat' basis. In addition to the intention-to-treat analysis, per-protocol and as-treated analyses will be applied as sensitivity analyses. Primary outcomes will be compared using the Wilcoxon's rank sum test. Secondary outcomes will be compared using the Fisher's exact test or the Wilcoxon's rank sum test, as necessary.

Monitoring

Regular data quality checks will be performed yearly, per the Quality Management Plan. Central monitoring will be conducted by the data center based on the CRF data collected via the EDC system. The monitoring staff will prepare a monitoring report after central monitoring and report it to the principal investigator.

Ethics and dissemination

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The participating institutions will include three high-volume centers with sufficient laparoscopic colorectal surgery caseloads (>100 cases/year), including one national cancer center and two university hospitals in Japan. The protocol for this RCT has been reviewed and approved by the ethics committees of each participating institution, and the trial has been registered at University Hospital Medical Information Network Center (https://www.umin.ac.jp/ctr/index-j.html) under the trial ID UMIN000050701. Written informed consent will be obtained from all patients after a thorough oral explanation by the doctor in charge at each center participating in the ImNavi trial. All procedures will be conducted in accordance with the ethical standards of the Declaration of Helsinki.

Discussion

Recently, analogous digital solutions have been translated and clinically implemented for diagnostic applications in gastrointestinal endoscopy [25] and radiology [26]. In the field of surgery, several deep learning-based CV solutions have been developed by academic and industry groups, mostly for MIS. Significant work has also been performed on potential use cases [3]; however, no CV tools are widely used for diagnostic or therapeutic applications in surgery. Recently, an intraoperative AI system identifying anatomical landmarks for laparoscopic cholecystectomy has been developed and reported [18]; similarly, in the urological field, an AI alert system that can predict the occurrence of intraoperative bleeding events during robotic surgery has been developed [27]. Although both systems could potentially be used as tools to improve the safety of the intervention, they were only tested in a trial involving 10 patients at a single center. Establishing solid evidence through RCTs is essential before AI can be applied in surgery to move beyond the proof-of-concept step to clinical

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

implementation. In addition, a multicenter study design is desirable to eliminate bias due to overfitting and to demonstrate generalization performance. This study is the first multicenter RCT to be designed to explore the clinical significance of AI-based intraoperative image navigation.

The target CV task of this study is the semantic segmentation of the ureter and autonomic nerves in laparoscopic colorectal surgery. Iatrogenic ureteral injury in colorectal surgery is a rare but potentially devastating complication of colorectal surgery. The incidence is approximately 0.5%–1.5% [28-30]; however, a significant increase has been observed with laparoscopy compared to open colectomies [29, 30] due to the lack of tactile information. In addition, trauma to the aortic plexus may occur during high ligation of the inferior mesenteric artery, and the superior hypogastric nerves may be injured along the sacral promontory or presacral region [31]. Injuries to the sympathetic nerves at this location affect the ability to ejaculate, including retrograde ejaculation, whereas injuries to the parasympathetic nerves result in erectile dysfunction [31]. A previous study reported a decrease in male sexual function following colorectal surgery, with a decrease from 78 to 32% in ejaculatory function, 71 to 24% in erectile function, and 82 to 57% in sexual activity [32]; these dysfunctions are often overlooked but should be strictly avoided.

The time required for recognition of each target organ by surgeons after its initial appearance on the monitor was set as the primary endpoint of this study, with the expectation that the AI-based intraoperative image navigation system would contribute to the more rapid recognition of target organs. Intraoperative target organ injury is a more straightforward primary endpoint; however, the number of events was considered too small for analysis. In addition, we believe that more rapid recognition of key

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

anatomical landmarks during surgery can lead to several benefits other than prevention of organ injury, such as identification of the correct dissection plane, and confidence in proceeding with the procedure. The impact of AI-based surgical applications on clinical outcomes beyond numerical expression will be explored from various facets while evaluating quantitative items, including intraoperative complications and operation time, as secondary endpoints. Furthermore, we expect that the findings of this RCT will contribute to advancing research in the domain of AI in surgery.

Strengths and limitations of this study

- The ImNavi trial is a Japanese multicenter RCT with 1:1 randomization between the use and non-use of an artificial intelligence-based intraoperative image-navigation system in laparoscopic surgery.
- Eligible candidates will be identified from patients referred to a colorectal surgeon or colorectal cancer multidisciplinary team based on colonoscopy, computed tomography, or magnetic resonance imaging findings, and eligibility will be confirmed following a review of the criteria.
- The primary outcome is the time required for each target organ to be recognized by the surgeon after its initial appearance on the monitor.
- Secondary outcomes include intraoperative target organ injuries; intraoperative complications; total operation time; operation time for each surgical step divided by existing definitions; blood loss; duration of postoperative hospital stay; postoperative complications within 30 days; postoperative male sexual dysfunction 1 month after surgery; surgeon's confidence in recognition of each target organ; and postoperative fatigue of the primary surgeon, according to the Piper Fatigue Scale-12.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

- The severity of all intraoperative and postoperative complications, measured according to Common Terminology Criteria for Adverse Events and the Clavien–Dindo classification, will also be collected as secondary outcomes.

Trial Status

This trial represents the first version of the study protocol. This trial was registered at the University Hospital Medical Information Network Center (https://www.umin.ac.jp/ctr/index-j.html) on March 28th, 2023 (trial ID: UMIN000050701). Recruitment will begin on June 2023 and is scheduled to be completed by December 2024.

LIST OF ABBREVIATIONS

AI: artificial intelligence

RCT: randomized controlled trials

MIS: minimally invasive surgery

CV: computer vision

EDC: electronic data capture

CRF: case report forms

DECLARATIONS

Ethics approval and consent to participate

The protocol for this RCT has been reviewed and approved by the ethics committees of each participating institution, and the trial has been registered at the University Hospital Medical Information Network Center (https://www.umin.ac.jp/ctr/index-j.html) under the trial ID UMIN000050701. Written informed consent will be obtained from all

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

patients after a thorough oral explanation by the doctor in charge at each center participating in the ImNavi trial. All procedures will be conducted in accordance with the ethical standards of the Declaration of Helsinki.

Consent for publication

Not applicable

Availability of data and materials

Not applicable

Competing interests

The authors declare that they have no competing interests.

Funding

This research is supported by the National Cancer Center Japan Research and Development Fund, grant number 2022-A-11.

Authors' contributions

All authors made substantial contributions to the study concept or data analysis and interpretation; drafted the manuscript or revised it critically for important intellectual content; approved the final version of the manuscript to be published; and agreed to be accountable for all aspects of the work.

Acknowledgements

We would like to thank all the members of the Clinical Research Support Office of the National Cancer Center Hospital East for their continuous support during the planning of this study.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

1. Weiser TG, Haynes AB, Molina G, Lipsitz SR, Esquivel MM, Uribe-Leitz T, et al. Estimate of the global volume of surgery in 2012: an assessment supporting improved health outcomes. Lancet. 2015;385 Suppl 2:S11.
2. Meara JG, Leather AJ, Hagander L, Alkire BC, Alonso N, Ameh EA, et al. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. Lancet. 2015;386:569-624.
3. 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;5:163.
4. Birkmeyer JD, Finks JF, O'Reilly A, Oerline M, Carlin AM, Nunn AR, et al. Surgical skill and complication rates after bariatric surgery. N Engl J Med. 2013;369:1434-42.
5. Curtis NJ, Foster JD, Miskovic D, Brown CSB, Hewett PJ, Abbott S, et al. Association of surgical skill assessment with clinical outcomes in cancer surgery. JAMA Surg. 2020;155:590-8.
6. Kitaguchi D, Takeshita N, Hasegawa H, Ito M. Artificial intelligence-based computer vision in surgery: Recent advances and future perspectives. Ann Gastroenterol Surg. 2022;6:29-36.
7. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020;577:89-94.
8. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

computed tomography. Nat Med. 2019;25:954-61.

9. Huang P, Lin CT, Li Y, Tammemagi MC, Brock MV, Atkar-Khattra S, et al. Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method. Lancet Digit Health. 2019;1:e353-62.

10. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med. 2019;25:1054-6.

11. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. Nat Rev Clin Oncol. 2019;16:703-15.

12. Zhou D, Tian F, Tian X, Sun L, Huang X, Zhao F, et al. Diagnostic evaluation of a deep learning model for optical diagnosis of colorectal cancer. Nat Commun. 2020;11:2961.

13. Liu H, Li L, Wormstone IM, Qiao C, Zhang C, Liu P, et al. Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs. JAMA Ophthalmol. 2019;137:1353-60.

14. Milea D, Najjar RP, Zhubo J, Ting D, Vasseneix C, Xu X, et al. Artificial intelligence to detect papilledema from ocular fundus photographs. N Engl J Med. 2020;382:1687-95.

15. Xie Y, Nguyen QD, Hamzah H, Lim G, Bellemo V, Gunasekeran DV, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. Lancet Digit Health. 2020;2:e240-9.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

16. Madani A, Namazi B, Altieri MS, Hashimoto DA, Rivera AM, Pucher PH, et al. Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. Ann Surg. 2022;276:363-9.

17. 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. Ann Surg. 2022;275:955-61.

18. Nakanuma H, Endo Y, Fujinaga A, Kawamura M, Kawasaki T, Masuda T, et al. An intraoperative artificial intelligence system identifying anatomical landmarks for laparoscopic cholecystectomy: a prospective clinical feasibility trial (J-SUMMIT-C-01). Surg Endosc. 2023;37:1933-42.

19. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020;2:e549-60.

20. Kitaguchi D, Harai Y, Kosugi N, Hayashi K, Kojima S, Ishikawa Y, et al. Artificial intelligence for the recognition of key anatomical structures in laparoscopic colorectal surgery. Br J Surg. 2023;110:1355-8.

21. Kojima S, Kitaguchi D, Igaki T, Nakajima K, Ishikawa Y, Harai Y, et al. Deep-learning-based semantic segmentation of autonomic nerves from laparoscopic images of colorectal surgery: an experimental pilot study. Int J Surg. 2023;109:813-20.

22. Kitaguchi D, Takeshita N, Matsuzaki H, Oda T, Watanabe M, Mori K, et al. Automated laparoscopic colorectal surgery workflow recognition using artificial

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

intelligence: Experimental research. Int J Surg. 2020;79:88-94.

23. Reeve BB, Stover AM, Alfano CM, Smith AW, Ballard-Barbash R, Bernstein L, et al. The Piper Fatigue Scale-12 (PFS-12): psychometric findings and item reduction in a cohort of breast cancer survivors. Breast Cancer Res Treat. 2012;136:9-20.

24. Mori T, Kimura T, Kitajima M. Skill accreditation system for laparoscopic gastroenterologic surgeons in Japan. Minim Invasive Ther Allied Technol. 2010;19:18-23.

25. Hassan C, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. Gastrointest Endosc. 2021;93:77-85.e6.

26. van Leeuwen KG, Schalekamp S, Rutten MJCM, van Ginneken B, de Rooij M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. Eur Radiol. 2021;31:3797-804.

27. Checcucci E, De Cillis S, Amparore D, Gabriele V, Piramide F, Piana A, et al. Artificial Intelligence alert systems during robotic surgery: a new potential tool to improve the safety of the intervention. Urol Vid J. 2023;18:100221.

28. Halabi WJ, Jafari MD, Nguyen VQ, Carmichael JC, Mills S, Pigazzi A, et al. Ureteral injuries in colorectal surgery: an analysis of trends, outcomes, and risk factors over a 10-year period in the United States. Dis Colon Rectum. 2014;57:179-86.

29. Palaniappa NC, Telem DA, Ranasinghe NE, Divino CM. Incidence of iatrogenic ureteral injury after laparoscopic colectomy. Arch Surg. 2012;147:267-71.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

30. Marcelissen TA, Den Hollander PP, Tuytten TR, Sosef MN. Incidence of iatrogenic ureteral injury during open and laparoscopic colorectal surgery: A single center experience and review of the literature. Surg Laparosc Endosc Percutan Tech. 2016;26:513-5.

31. Giglia MD, Stein SL. Overlooked long-term complications of colorectal surgery. Clin Colon Rectal Surg. 2019;32:204-11.

32. Adam JP, Denost Q, Capdepont M, et al. Prospective and longitudinal study of urogenital dysfunction after proctectomy for rectal cancer. Dis Colon Rectum. 2016;59:822-30.

Figure legends

Figure 1: Trial flow diagram

Figure 2: Example of the monitor layout in an operation room

The system output (semantic segmentation pixels) is superimposed on a sub-monitor placed adjacent to the main monitor, and does not affect the main monitor.

Additional files

Additional file 1: Ethical approval document (English translation version)

Additional file 2: Copy of the original funding documentation (English translation version)

Additional file 3: A completed SPIRIT checklist

Additional file 4: A SPIRIT figure

SPIRIT-AI Checklist: Recommended items to address in a protocol and related documents for clinical trials evaluating AI interventions

Section		SPIRIT 2013 Item[a]	SPIRIT-AI Item		Addressed on Page No [b]
Administrative Information					
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	SPIRIT-AI Elaboration 1(i)	Indicate that the intervention involves artificial intelligence / machine learning and specify the type of model.	1
			SPIRIT-AI Elaboration 1(ii)	Specify the intended use of the AI intervention.	1
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry			4
	2b	All items from the World Health Organization Trial Registration Data Set			4
Protocol version	3	Date and version identifier			N/A
Funding	4	Sources and types of financial, material, and other support			17
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors			1
	5b	Name and contact information for the trial sponsor			1-2
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities			17
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)			17
Introduction					

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	SPIRIT-AI 6a (i) Extension	Explain the intended use of the AI intervention in the context of the clinical pathway, including its purpose and its intended users (e.g. healthcare professionals, patients, public).	6-7
			SPIRIT-AI 6a (ii) Extension	Describe any pre-existing evidence for the AI intervention.	6-7
	6b	Explanation for choice of comparators			6-7
Objectives	7	Specific objectives or hypotheses			7
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)			7-8
Methods: Participants, Interventions and Outcomes					
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	SPIRIT-AI 9 Extension	Describe the onsite and offsite requirements needed to integrate the AI intervention into the trial setting.	8
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	SPIRIT-AI 10 (i) Elaboration	State the inclusion and exclusion criteria at the level of participants.	9
			SPIRIT-AI 10 (ii) Extension	State the inclusion and exclusion criteria at the level of the input data.	9
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	SPIRIT-AI 11a (i) Extension	State which version of the AI algorithm will be used.	9-10
			SPIRIT-AI 11a (ii) Extension	Specify the procedure for acquiring and selecting the input data for the AI intervention.	9-10
			SPIRIT-AI 11a (iii) Extension	Specify the procedure for assessing and handling poor quality or unavailable input data.	9-10
			SPIRIT-AI 11a	Specify whether there is human-AI interaction in the	9-10

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

			(iv) Extension	handling of the input data, and what level of expertise is required for users.	9-10
			SPIRIT-AI 11a (v) Extension	Specify the output of the AI intervention.	
			SPIRIT-AI 11a (vi) Extension	Explain the procedure for how the AI intervention's output will contribute to decision-making or other elements of clinical practice.	9-10
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)			9-10
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)			9-10
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial			9-10
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended			10
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)			10-11
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions			11-12

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

		supporting any sample size calculations			
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size			8
Methods: Assignment of Interventions (For Controlled Trials)					
Sequence generation	16A	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions			12
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned			12
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions			12
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how			12
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial			12
Methods: Data Collection, Management, And Analysis					
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms			12-13

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

		can be found, if not in the protocol			
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols			12-13
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol			12-13
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol			13
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)			13
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)			13
Methods: Monitoring					
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed			13-14
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision			13-14

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

		to terminate the trial			
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	SPIRIT-AI Extension 22	Specify any plans to identify and analyse performance errors. If there are no plans for this, explain why not.	13-14
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor			13-14
Ethics and Dissemination					
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval			14-15
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)			14-15
Consent or ascent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)			14-15
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable			14-15
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial			14-15
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site			14-15
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	SPIRIT-AI Extension 29	State whether and how the AI intervention and/or its code can be accessed, including any restrictions to access or re-use.	14-15
Ancillary and	30	Provisions, if any, for ancillary and post-trial care, and for compensation			14-15

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

post-trial care		to those who suffer harm from trial participation			
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions			14-15
	31b	Authorship eligibility guidelines and any intended use of professional writers			14-15
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code			14-15
Appendices					
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates			N/A
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable			N/A

[a] It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items.

[b] Indicates page numbers to be completed by authors during protocol development

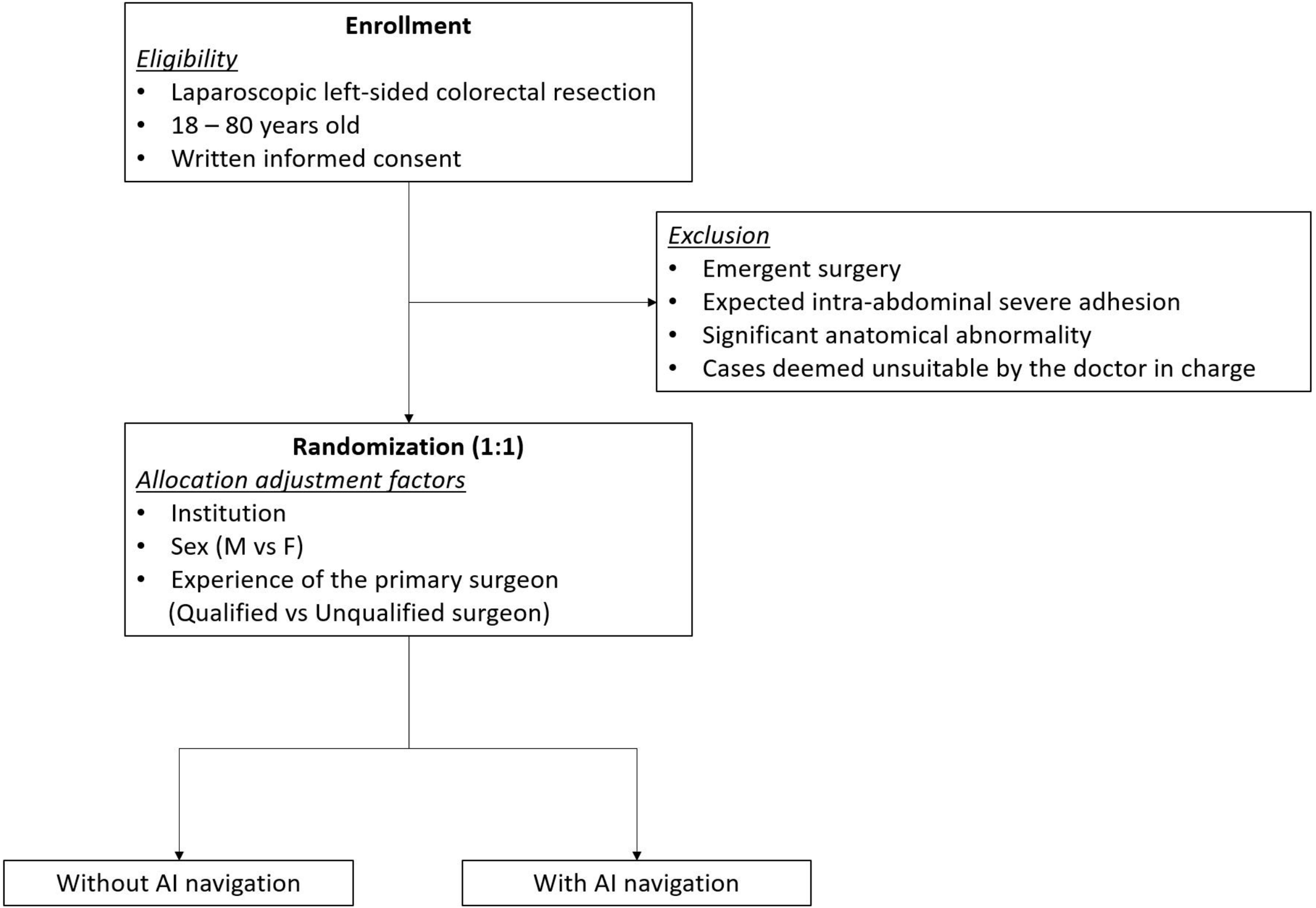
medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.05.24310603; this version posted August 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

SPIRIT figure: The schedule of enrolment, interventions, and assessments.

	STUDY PERIOD				
	Enrolment	**Allocation**	**Post-allocation**		**Close-out**
TIMEPOINT	***-t_1***	**0**	***OP****	***Post-OP***	***POM1*****
ENROLMENT:					
Eligibility screen	X				
Informed consent	X				
Allocation		X			
INTERVENTIONS:					
Without AI navigation			X		
With AI navigation			X		
ASSESSMENTS:					
Baseline variables	X	X			
Postoperative variables				X	X
Follow-up variables					X

*OP: date of operation
**POM1: postoperative month 1


Enrollment
Eligibility
• Laparoscopic left-sided colorectal resection
• 18 – 80 years old
• Written informed consent
Exclusion
• Emergent surgery
• Expected intra-abdominal severe adhesion
• Significant anatomical abnormality
• Cases deemed unsuitable by the doctor in charge
Randomization (1:1)
Allocation adjustment factors
• Institution
• Sex (M vs F)
• Experience of the primary surgeon
(Qualified vs Unqualified surgeon)
Without AI navigation
With AI navigation