

Protocol

AI Models to Reduce Surgical Complications Through Intraoperative Video Analysis: Protocol for a Prospective Cohort Study

António Sampaio Soares¹, MD, PhD; Sophia Bano², PhD; Laura T Castro³, MD; Margarida Pascoal⁴, MD; Ricardo Rocha³, MD; Paulo Alves³, MD; Paulo Mira³, MD; Joao Costa¹, MD, PhD; Manish Chand², MD, PhD; Danail Stoyanov², PhD; Catarina Barata⁵, PhD

¹School of Medicine, University of Lisbon, Lisbon, Portugal

²University College London, London, United Kingdom

³Hospital Professor Doutor Fernando Fonseca, Amadora, Portugal

⁴Instituto Português de Oncologia Francisco Gentil, Lisbon, Portugal

⁵Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Lisbon, Portugal

Corresponding Author:

António Sampaio Soares, MD, PhD
School of Medicine, University of Lisbon
Av. Prof. Egas Moniz MB
Lisbon 1649-028
Portugal
Phone: 351 963864555
Email: antoniosoares@medicina.ulisboa.pt

Abstract

Background: Complications following abdominal surgery have a very significant negative impact on the patient and the health care system. Despite the spread of minimally invasive surgery, there is no automated way to use intraoperative video to predict complications. New developments in data storage capacity and artificial intelligence (AI) algorithm creation now allow for this.

Objective: This project aims to develop and validate deep learning models for accurately predicting postoperative complications, classified using the Clavien-Dindo scale. A key objective is to build and share an open-source dataset containing both intraoperative video data and postoperative outcomes.

Methods: This prospective cohort study will collect data reflecting day-to-day surgical practice from 1200 patients, focusing on patient outcomes and intraoperative video. Data will be collected from patients undergoing minimally invasive appendectomy, cholecystectomy, and colorectal resection in the urgent and elective settings. Each video will be annotated at the temporal and semantic level by the study team. Comprehensive data collection will encompass three domains: (1) preoperative variables, including patient demographics, comorbidities, laboratory values, and imaging findings; (2) intraoperative data featuring complete surgical video recordings from laparoscopic or robotic monitors, procedure duration, surgical approach, intraoperative complications, and surgeon-defined technical factors; and (3) 30-day postoperative outcomes classified using the Clavien-Dindo scale (grades I-V). This dataset will be shared under a noncommercial CC BY-NC-SA use license to promote scientific collaboration and innovation, with complete anonymization including metadata removal and out-of-body image blurring. For analysis, the dataset will be split into training, validation, and testing sets. Deep learning algorithms will be developed through supervised learning methodology using 2 parallel approaches: data-derived predictors using fine-tuned surgical video foundational models based on vision transformer architectures and surgeon-defined predictors based on documented intraoperative strategies. Algorithms will be trained on the training set to predict the Clavien-Dindo postoperative complication grade and categorize postoperative outcomes in minimally invasive abdominal surgery. Model performance will be analyzed through sensitivity, specificity, positive and negative predictive values, and area under the receiver operating characteristic curve on the validation and testing sets.

Results: Data collection started in 2024 and is expected to extend throughout 2025. The planned outputs include the publication of a research protocol, main results, and the open-source dataset. Through this initiative, the project seeks to significantly advance the field of AI-assisted surgery, contributing to safer and more effective practice.

Conclusions: Through the creation of an open dataset and the development of state-of-the-art deep learning models, this project seeks to transform the current paradigm in minimally invasive surgery. By providing the surgical AI community with robust, real-world data, the project aspires to catalyze innovations that will enhance surgical safety; refine predictive capabilities; and, ultimately, lead to better clinical outcomes.

International Registered Report Identifier (IRRID): DERR1-10.2196/62734

JMIR Res Protoc 2026;15:e62734; doi: [10.2196/62734](https://doi.org/10.2196/62734)

Keywords: artificial intelligence; AI; surgery; dataset; intraoperative video; appendectomy; cholecystectomy

Introduction

The conventional approach to surgical training is deeply rooted in a model that combines theoretical teaching with supervised practice in the operating environment. However, this model has its flaws. One of the most notable is the difficulty in predicting the variability that occurs during surgery itself, whether due to each patient's unique anatomy or the specific pathology that justifies the intervention. Although surgeons accumulate a wealth of information during each procedure, these data often remain confined within the walls of the operating room. This is largely due to the lack of efficient mechanisms for capturing these data and the subsequent inability to analyze them systematically and rigorously. This situation creates a vicious cycle in which learning becomes fragmented and inefficient, negatively affecting clinical outcomes and the evolution of surgical practice.

The emergence of minimally invasive surgery marked a revolution in clinical practice. In addition to minimizing trauma and accelerating patient recovery, this approach provides an unprecedented opportunity: the ability to record the entire surgical procedure on video. Previous works, such as those by Birkmeyer et al [1] and Curtis et al [2], have already provided valuable insights by demonstrating that the quality of surgical gestures is correlated with clinical outcomes. These works were based on a still very rudimentary human assessment that cannot be scaled. With the advent of new technologies in image analysis and artificial intelligence (AI), it is now possible to automate the evaluation of these videos. The objective of this project is to go beyond simple evaluation, developing AI algorithms that can provide real-time support during the surgical act.

Postsurgical complications, both in minimally invasive cholecystectomy and colectomy, have a very high human and economic cost. They increase patient mortality and morbidity and burden health systems with additional costs in postoperative care and prolonged hospital stays [3]. In this context, it is not only desirable but also ethically imperative to seek new ways to reduce the risk of complications and improve clinical outcomes. Currently, postoperative complications are classified according to the Clavien-Dindo complication scale [4]. This scale classifies outcomes according to 5 grades: I (minor deviation from normal recovery), II (requiring pharmacological treatment), III (requiring surgical, endoscopic, or radiological intervention), IV (life-threatening complication requiring intensive care), and V (death of the patient).

The use of AI algorithms for intraoperative video analysis has the potential to revolutionize how safety in surgery is approached, allowing for the identification and correction of problems in real time and, thus, significantly improving the quality of care provided. This prospective dataset is intended to increase the data available for research on AI in surgery reflecting day-to-day practice. It will include outcome data to allow the use of this knowledge to guide the study of the surgical procedure. The aim of this project is the identification of intraoperative factors that predict postoperative complications based on the Clavien-Dindo classification in minimally invasive cholecystectomies, appendectomies, and colorectal resections and the development of AI algorithms for this purpose.

Methods

Dataset Construction

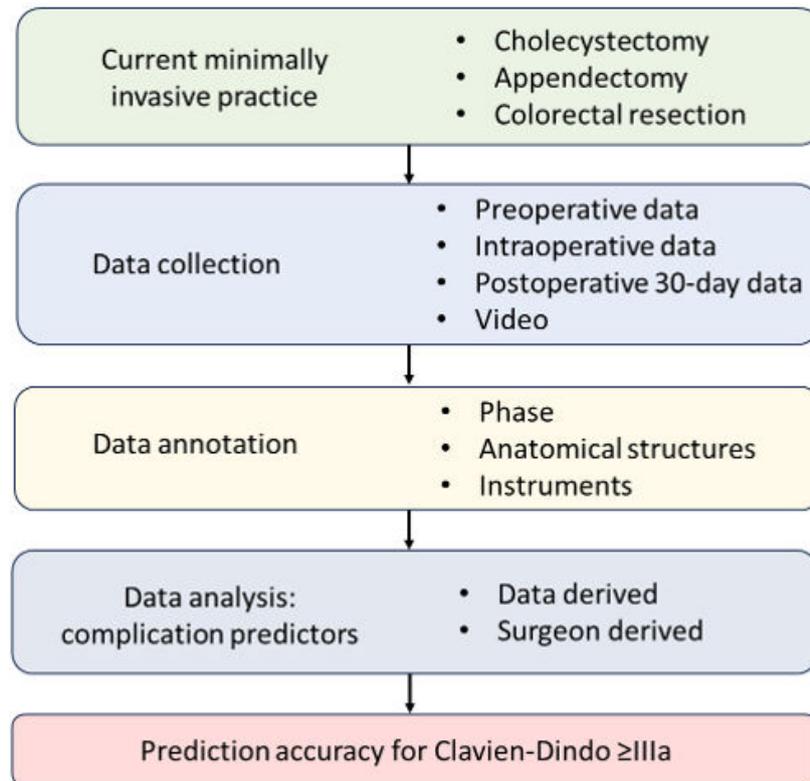
Study Design

This will be a multicentric prospective cohort study for patients undergoing minimally invasive cholecystectomy, appendectomy, or colorectal resection. Recruiting centers will be based in Europe.

Dataset Summary

The dataset will cover both preoperative and intraoperative data, as well as video and 30-day postoperative outcomes, on groups of patients regularly operated in a minimally invasive fashion, as can be seen in [Figure 1](#). These patients will undergo cholecystectomy, appendectomy, and colorectal resection. These are high-volume surgeries, with a significant proportion being performed minimally invasively [5]. Video data will be collected from the minimally invasive device monitor (laparoscopic or robotic). Other variables will be tabular data, as detailed for every variable and procedure in [Multimedia Appendices 1 to 3](#).

The datasets will be called Surg_Cloud Cholecystectomy, Surg_Cloud Appendectomy, and Surg_Cloud Colorectal and will include the version number (eg, V1 and V2). The first release is expected by June 2025. It will be shared through a web host, with log-in credentials available for researchers. The dataset will be made available under a CC BY-NC-SA license. This allows for its use for research purposes but restricts its use for commercial purposes. This balance aims to promote scientific collaboration and innovation while protecting the data from unauthorized commercial exploitation.

Figure 1. Surg_Cloud study flowchart.

Inclusion Criteria

Data will be collected from patients receiving health care under normal circumstances. All patients undergoing cholecystectomy, appendectomy, and colorectal resection performed minimally invasively will be eligible. This applies to patients undergoing surgery in the urgent and elective settings. No changes in normal clinical practice will occur. These procedures are extremely common in a general surgeon's practice, and therefore, most hospitals can potentially enroll patients. Moreover, disease presentation is extremely heterogeneous across these procedures. This provides an opportunity to assess differential strategies in terms of changes in disease presentation. Patients withdrawing informed consent and for whom videos are not recorded adequately will be excluded.

Data Sampling and Shifts Over Time

The dataset will be composed of anonymized patient data from multiple hospitals. The enrollment of patients will take place under day-to-day clinical conditions, reflecting a convenience sample. It is to be expected that sicker patients will not be as represented among the study participants as more clinically stable ones, given that obtaining preoperative consent for unstable patients tends to be more difficult.

Every dataset released will have a corresponding version number where the details of data acquisition will be made explicit. If any significant changes occur between versions of the release, these will also be detailed in the accompanying documentation. It is expected that most variation will occur in terms of different hospitals' approaches to the specific procedures.

Patients will be grouped according to standard practice as mentioned in the data collection forms, which can be found in [Multimedia Appendices 1 to 3](#). Apart from the video and the data mentioned in this protocol, no other data will be collected from patients. These data will be available for every patient.

Groups at Risk of Disparate Health Outcomes

The dataset comprises various attributes that reflect patient interactions in public hospitals. These attributes were selected to provide insights into common health care practices and patient outcomes in these settings. Each attribute is presented at an individual level, ensuring a granular and detailed analysis. However, it is important to note that no patient-identifiable data have been collected to preserve privacy.

Limitations of the Dataset

The dataset will be created from videos collected in a very diverse range of hospitals. The patient selection will be mostly a convenience sample given the constraints of surgical care provision. Data quality will be variable given the range of devices that collect the data. These specifics will be detailed on a per-patient basis.

Modifications Made to the Data

The only data modifications planned are the blurring of out-of-body images to ensure complete anonymization. No information on the hospital where the data were collected will be shared. No further modifications are planned at this stage. No synthetic data will be used. No missing data will be imputed.

Known or Potential Bias Caused or Exacerbated by Data Acquisition and Processing

Given the complete anonymization of patients and the lack of impact on the patients' clinical care, no bias is expected to be exacerbated by the creation of this dataset.

Formal exclusion criteria have not been defined. Patients less likely to be included in the dataset are those undergoing urgent operations, in which the patients are likely to be sicker. This will necessarily impact the ability of the team to ensure proper informed consent. Due to this, a lower number of urgent patients are expected to be included.

The variables included are well standardized. Postoperative complications will be categorized using the Clavien-Dindo scale [4], which has 5 levels of severity, from I (minor deviations from the desired postoperative course) to V (patient death). This scale provides a structured framework for classifying and comparing complications, allowing for more precise analysis.

Ethical Considerations

Patient inclusion will necessitate ethics approval at every hospital and informed consent from every patient recruited. No patient-identifiable data will be shared in the dataset. All videos will be fully anonymized by removing all metadata and blurring out-of-body images if they appear.

This study has already received approval from the Hospital Professor Doutor Fernando Fonseca Ethics Committee (113/2023) to identify predictors of surgical complications, as well as for dataset curation and open-source publication. Every patient approached will be told the aims of this study, the absence of change in routine clinical care, and that only completely anonymized data will be made public in the dataset. No compensation will be provided to participants.

Patient and Public Participation

A formal patient and public involvement initiative is planned to be developed before dataset publication.

Importance of Reproducibility and Generalizability

Reproducibility of results in research is crucial for the credibility and advancement of science. In accordance with the principles of findability, accessibility, interoperability, and reusability [6], this project aims not only to conduct rigorous studies but also to make data and methods available in a way that allows for their verification and reuse by other researchers. Dataset development was conducted according to the STANDING (Standards for Data Diversity, Inclusivity, and Generalisability) Together recommendations [7]. By doing so, we increase the likelihood that the results are generalizable and, consequently, more useful worldwide. The creation of an open dataset follows the example of successful initiatives such as the Heidelberg Colorectal dataset published in the journal *Nature* [8]. This dataset is a reference in the field, consisting of 30 videos of colorectal surgeries

accompanied by annotations and extensive documentation. Other institutions such as the Institut de Recherche contre les Cancers de l'Appareil Digestif (IRCAD) in the case of laparoscopic cholecystectomy [9] and the University of Dresden [10] have also contributed to science by publishing open-source datasets.

Analysis Plan

Objectives

We aim to develop a deep neural network using supervised methodology for predicting the grade of postoperative complications using the Clavien-Dindo scale. The overall aim of this study is the identification of predictors of postoperative complications in the included procedures.

Analysis

In the initial phase, a descriptive statistical analysis of demographic, intraoperative, and postoperative variables will be conducted. This analysis aims to provide a quantitative summary of the data, helping define the profile of the patients and the surgical interventions performed. Measures of central tendency, dispersion, and distribution will be calculated for each variable, such as mean, SD, median, and percentiles.

Subsequently, a regression analysis will be conducted to identify variables that may predict postoperative complications. This regression model will be adjusted to account for potential confounding factors and will allow for the assessment of the independent impact of each predictor variable. The analysis will be conducted for both categorical and quantitative variables. The ability of preoperative variables to predict surgery duration will be assessed.

All statistical tests will be performed with a predefined significance level of $P=.05$, and the results will be interpreted in the light of the clinical context to ensure that the conclusions are not only statistically significant but also clinically relevant.

Algorithm Development

Compliance With Methodological Guidelines

The development of AI algorithms for this project will be rigorously guided by best practices. The TRIPOD-AI (Transparent Reporting of a Multivariable Model for Individual Prognosis or Diagnosis–Artificial Intelligence) [11] guidance will be followed. This ensures that the development, validation, and implementation of AI algorithms in medical contexts follow strict standards to ensure their efficacy and safety. The publication of the STARD-AI (Standards for Reporting Diagnostic Accuracy–Artificial Intelligence) [12] is expected soon, and according to the timeline of this project, their use in publications originating from it is anticipated.

Supervised Learning Approach

The algorithms will be trained using a supervised learning approach, with the categories of the Clavien-Dindo scale serving as the outcome variable or “label,” as can be seen

in [Figure 1](#). Supervised learning was chosen to allow for more focused and accurate training considering the specific and complex characteristics of the medical data involved.

The identification of predictors will follow 2 parallel work streams: data-based predictors and surgeon-defined predictors. For the data-based predictors, we will fine-tune a surgical video foundational model [13], adjusting it to predict the categories of the Clavien-Dindo scale. The foundational model has been pretrained on a vast amount of publicly available surgical videos and is expected to provide a discriminative feature space, which can be further fine-tuned to adjust to the desired task. The surgeon-defined predictors will be based on the exhaustive documentation of intraoperative strategies and their correlation with 30-day complication rates. For surgeon-derived predictors, annotated data will be used to automate their detection via the same algorithm design as for data-derived predictors.

All videos will be made available, with annotations performed in a collaborative fashion. These will identify phase, anatomical structures, and instruments. The annotation process will be standardized using validated protocols when available, such as the protocol by Mascagni et al [14].

The collected video data will be divided into 3 sets: training, validation, and testing. The proportions for this division will be 70%, 15%, and 15%, respectively. This structure aims to ensure that the model is trained on a large amount of data while reserving an adequate portion for validation and testing, thereby minimizing the risk of overfitting.

The Python programming language (Python Software Foundation), widely recognized for its versatility and robustness in AI projects, will be used. The specific infrastructure for training the neural networks will be PyTorch (Meta AI), an open-source machine learning library that is highly effective for implementing convolutional neural networks.

The specific type of neural network to be used will be a vision transformer as this is the backbone of the foundational model to be fine-tuned. Vision transformers have shown remarkable performances in different medical image analysis tasks [15]; thus, they are ideal for the task at hand. Additionally, the ability of transformers to perform a region-based analysis can play a relevant role in identifying key elements in surgical videos that correlate with patient outcomes through the inspection of the output of the self-attention blocks, in particular those associated with the “[class]” token. As we will be leveraging a foundational model to develop the algorithm, we have some level of assurance that training will be easier and lead to better performance.

Evaluation and Continual Improvement

After initial training, the algorithm will undergo several rounds of evaluation and adjustment. Metrics such as accuracy, sensitivity, and specificity will be used to assess the model's performance. The results of these evaluations will inform subsequent iterations of the algorithm, allowing

for continuous improvement of its performance in line with current best practices [16]. Thus, the project aims to develop a highly effective and reliable AI algorithm that can be used to significantly improve surgical practice and patient outcomes.

Data Storage

Data will be uploaded to a secure cloud storage service. This service will include encrypted storage and access controlled via verified credentials, ensuring that data access is traceable. Additionally, the platform allows for the connection with applications to run the code for training and testing the algorithm. Both clinical and video data are uploaded after a period of 30 days after surgery, ensuring complete anonymization of the uploaded data. Video will be preprocessed before storage to remove patient metadata and blur out-of-body frames using an open-source video processing pipeline.

Authorship

Collaborative authorship will be attributed to members of the study group according to the established guidelines [17]. One hospital lead and up to 3 local collaborators will be eligible per center.

Results

Center recruitment is ongoing. Patient recruitment is expected to take place throughout 2026. Early algorithm development should start in June 2025, with the first outputs expected starting from January 2026, as can be seen in [Figure 2](#). As of January 2025, a total of 3 hospitals have joined the study, with 2 working to secure local ethics committee approval. A preliminary assessment of each site has confirmed that data collection—encompassing both minimally invasive surgical video and corresponding patient information—can be integrated into routine workflows without disrupting patient care. Early interactions with hospital leads have indicated strong support for the study's objectives, suggesting that collaboration and data sharing can continue effectively.

Video quality and anonymization processes were confirmed to meet standards, with out-of-body images blurred to protect patient identities and metadata removed before secure upload. All data were subsequently linked via anonymized identifiers, validating the planned protocol for data management. Feedback from this pilot phase has been documented and will guide the standardized workflows for other hospitals as they begin data collection.

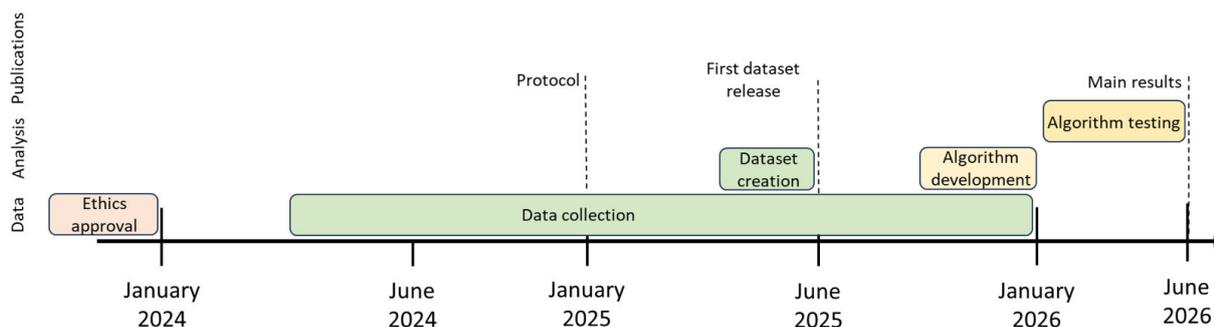
Every patient who participated in the single hospital actively recruiting provided informed consent after receiving clear, concise explanations of the study's aims, potential benefits, and associated risks. The research team worked collaboratively with each hospital's ethics committee to ensure that the consent forms and protocols met local regulatory requirements. Early discussions with patients and clinicians suggested that the consent process is well understood. No significant ethical or logistical challenges were reported during this preliminary testing, and the approach

appears suitable for broader implementation as additional centers come on board.

On the basis of initial enrollment rates and interest from additional sites, it is projected that the dataset will eventually include approximately between 1000 and 1200 patients undergoing minimally invasive cholecystectomy, appendectomy, or colorectal resection. This sample size was determined considering established principles for deep learning model development in medical contexts. For supervised learning tasks using pretrained models, it is recognized that adequate representation of outcome classes is essential for robust model training. Given expected overall complication rates of 15% to 20% across the 3 target procedures (approximately 10% for cholecystectomy, 15% for appendectomy, and 25% for colorectal resection), our planned sample yields approximately 150 to 240 patients with complications classified as Clavien-Dindo grade II or higher. The 70-15-15 training-validation-test split provides sufficient cases across complication grades to support both binary and multiclass classification approaches while maintaining

adequate statistical power to detect clinically meaningful differences in model performance. This comprehensive dataset is expected to comprise preoperative and intraoperative variables, postoperative outcomes up to 30 days, and video recordings of each procedure. All video data will undergo a standardized annotation process to identify surgical phases, instruments, and anatomical structures, which should facilitate subsequent supervised learning tasks. The primary analysis will explore how intraoperative factors correlate with the Clavien-Dindo classification of postoperative complications, with secondary objectives focusing on regression modeling and convolutional neural network training to predict complication severity. These methods are designed to generate actionable insights into surgical risks and lay the groundwork for algorithmic tools that can support clinical decision-making. Feasibility assessment based on pilot data from participating centers indicates average monthly recruitment of 15 to 20 eligible patients per site, and with expansion to 5 to 6 centers, target recruitment remains achievable within the planned timeline.

Figure 2. Surg_Cloud project timeline.



Discussion

Expected Findings

The primary aim of this project is to identify intraoperative factors that predict major postoperative complications following minimally invasive abdominal surgery. By exploring how anatomical structures, surgical instruments, and their interactions contribute to these complications, the project seeks to refine risk stratification and compare these AI-derived insights with existing postoperative complication prediction scores. Moreover, this initiative will result in a shared, open-source dataset that can be interrogated by the wider surgical AI community—maximizing collaborative potential and accelerating advancements in surgical video analysis.

Building on prior efforts—such as those by Birkmeyer et al [1] and Curtis et al [2], who correlated technical performance with clinical outcomes through video review—this initiative adds crucial elements to enhance scalability and applicability. Recent projects, including the Heidelberg Colorectal dataset [6] and open access resources from the Research Institute Against Cancer of the Digestive System [9] and the University of Dresden [10], highlight

the value of systematic video collection, annotation, and analysis. However, many existing datasets have limitations, such as small sample sizes or lack of postoperative outcome data, restricting comprehensive analysis. In contrast, this project integrates the Clavien-Dindo classification [3] for robust tracking of postoperative complications and adopts a multicentric, prospective design to capture a broader range of patient populations and clinical scenarios.

Despite these strengths, some limitations warrant attention. First, reliance on convenience sampling across multiple hospitals may underrepresent urgent or high-risk cases due to limited time for obtaining informed consent. Second, variations in video capture technology and differing levels of surgical expertise introduce heterogeneity that lowers internal validity but also reflect real-world conditions—potentially boosting external validity. Third, the project currently emphasizes algorithm development rather than hypothesis testing, although future iterations will include more rigorous prospective analyses once robust models are established. While other initiatives such as the Operating Room Black Box [18] capture additional operating room data that may influence postoperative complications [19], such technology remains inaccessible to many surgical practices. Consequently, this project focuses on creating a

scalable, multicentric data collection and algorithm development pipeline designed for real-world applicability with a broader user base.

Looking ahead, several directions will further strengthen this endeavor. Plans include expanding dataset storage to accommodate increasing volumes of video and metadata from national and international collaborators guided by established frameworks such as STANDING Together [7] and the TRIPOD-AI guidelines [11]. Additionally, federated learning methodologies [20] will be explored to enable multiple institutions to train algorithms collectively while safeguarding patient privacy. Ongoing dissemination efforts will involve publishing open access protocols, inviting broader research contributions, and fostering a collaborative global surgical AI community. Through these initiatives, the project aspires to accelerate advancements in surgical video analysis and refine the predictive capabilities of AI in accurately forecasting postoperative complications.

Acknowledgments

The authors would like to acknowledge Pietro Mascagni and Fiona Kolbinger for helpful discussion on the topic of complication prediction. The authors used the generative artificial intelligence tool ChatGPT by OpenAI for minor writing edits in certain sections of the manuscript.

Funding

This work has received funding from a Portuguese Society of Coloproctology research grant and Portuguese Foundation for Science and Technology research grant (2024.07248.IACDC) <https://sciproj.ptcris.pt/176816PRJ>.

Data Availability

The dataset created through the implementation of this protocol will be made available to the surgical community through a dedicated publication.

Authors' Contributions

Conceptualization: ASS, SB, MC, DS

Data curation: LTC, MP

Methodology: ASS, SB, JC, CB

Supervision: PA, PM

Writing—original draft: ASS, LTC

Writing—review and editing: ASS, LTC, CB

Conflicts of Interest

DS is employed at Medtronic plc. All other authors declare no conflicts of interest.

Multimedia Appendix 1

Case report form for patients undergoing cholecystectomy.

[\[DOCX File \(Microsoft Word File\), 18 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Case report form for patients undergoing appendectomy.

[\[DOCX File \(Microsoft Word File\), 17 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Case report form for patients undergoing colorectal resection.

[\[DOCX File \(Microsoft Word File\), 19 KB-Multimedia Appendix 3\]](#)

References

1. Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med*. Oct 10, 2013;369(15):1434-1442. [doi: [10.1056/NEJMs1300625](https://doi.org/10.1056/NEJMs1300625)] [Medline: [24106936](https://pubmed.ncbi.nlm.nih.gov/24106936/)]

Conclusions

This project bridges a critical gap in surgical research by integrating high-quality, multicentric intraoperative video data with rigorous postoperative outcome measures. By doing so, it offers a promising avenue for improving the prediction of major complications following minimally invasive abdominal surgery and enhancing surgical safety and quality. The open-source dataset, combined with transparent, reproducible methods, holds the potential to galvanize a broader community of clinicians, data scientists, and AI researchers. Through this collective effort, we anticipate the development of novel AI-driven interventions that will refine surgical practice; reduce complication rates; and, ultimately, improve patient outcomes.

2. Curtis NJ, Foster JD, Miskovic D, et al. Association of surgical skill assessment with clinical outcomes in cancer surgery. *JAMA Surg.* Jul 1, 2020;155(7):590-598. [doi: [10.1001/jamasurg.2020.1004](https://doi.org/10.1001/jamasurg.2020.1004)] [Medline: [32374371](https://pubmed.ncbi.nlm.nih.gov/32374371/)]
3. Healy MA, Mullard AJ, Campbell DA Jr, Dimick JB. Hospital and payer costs associated with surgical complications. *JAMA Surg.* Sep 1, 2016;151(9):823-830. [doi: [10.1001/jamasurg.2016.0773](https://doi.org/10.1001/jamasurg.2016.0773)] [Medline: [27168356](https://pubmed.ncbi.nlm.nih.gov/27168356/)]
4. Clavien PA, Barkun J, de Oliveira ML, et al. The Clavien-Dindo classification of surgical complications: five-year experience. *Ann Surg.* Aug 2009;250(2):187-196. [doi: [10.1097/SLA.0b013e3181b13ca2](https://doi.org/10.1097/SLA.0b013e3181b13ca2)] [Medline: [19638912](https://pubmed.ncbi.nlm.nih.gov/19638912/)]
5. Sheetz KH, Claflin J, Dimick JB. Trends in the adoption of robotic surgery for common surgical procedures. *JAMA Netw Open.* Jan 3, 2020;3(1):e1918911. [doi: [10.1001/jamanetworkopen.2019.18911](https://doi.org/10.1001/jamanetworkopen.2019.18911)] [Medline: [31922557](https://pubmed.ncbi.nlm.nih.gov/31922557/)]
6. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data.* Mar 15, 2016;3:160018. [doi: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)] [Medline: [26978244](https://pubmed.ncbi.nlm.nih.gov/26978244/)]
7. Alderman JE, Palmer J, Laws E, et al. Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations. *Lancet Digit Health.* Jan 2025;7(1):e64-e88. [doi: [10.1016/S2589-7500\(24\)00224-3](https://doi.org/10.1016/S2589-7500(24)00224-3)] [Medline: [39701919](https://pubmed.ncbi.nlm.nih.gov/39701919/)]
8. Maier-Hein L, Wagner M, Ross T, et al. Heidelberg colorectal data set for surgical data science in the sensor operating room. *Sci Data.* Apr 12, 2021;8(1):101. [doi: [10.1038/s41597-021-00882-2](https://doi.org/10.1038/s41597-021-00882-2)] [Medline: [33846356](https://pubmed.ncbi.nlm.nih.gov/33846356/)]
9. Twinanda AP, Shehata S, Mutter D, Marescaux J, de Mathelin M, Padoy N. EndoNet: a deep architecture for recognition tasks on laparoscopic videos. *IEEE Trans Med Imaging.* Jan 2017;36(1):86-97. [doi: [10.1109/TMI.2016.2593957](https://doi.org/10.1109/TMI.2016.2593957)] [Medline: [27455522](https://pubmed.ncbi.nlm.nih.gov/27455522/)]
10. Carstens M, Rinner FM, Bodenstedt S, et al. The Dresden surgical anatomy dataset for abdominal organ segmentation in surgical data science. *Sci Data.* Jan 12, 2023;10(1):3. [doi: [10.1038/s41597-022-01719-2](https://doi.org/10.1038/s41597-022-01719-2)] [Medline: [36635312](https://pubmed.ncbi.nlm.nih.gov/36635312/)]
11. Collins GS, Moons KG, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ.* Apr 16, 2024;385:e078378. [doi: [10.1136/bmj-2023-078378](https://doi.org/10.1136/bmj-2023-078378)] [Medline: [38626948](https://pubmed.ncbi.nlm.nih.gov/38626948/)]
12. Sounderajah V, Ashrafian H, Golub RM, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. *BMJ Open.* Jun 28, 2021;11(6):e047709. [doi: [10.1136/bmjopen-2020-047709](https://doi.org/10.1136/bmjopen-2020-047709)] [Medline: [34183345](https://pubmed.ncbi.nlm.nih.gov/34183345/)]
13. Wang Z, et al. Foundation model for endoscopy video analysis via large-scale self-supervised pre-train. In: Greenspan H, Madabhushi A, Mousavi P, editors. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023.* Springer Nature; 2023:101-111. [doi: [10.1007/978-3-031-43996-4_10](https://doi.org/10.1007/978-3-031-43996-4_10)] ISBN: 9783031439964
14. Mascagni P, Alapatt D, Garcia A, et al. Surgical data science for safe cholecystectomy: a protocol for segmentation of hepatocystic anatomy and assessment of the critical view of safety. *arXiv.* Preprint posted online on Jun 21, 2021. URL: <https://arxiv.org/abs/2106.10916> [Accessed 2025-05-25]
15. Shamshad F, Khan S, Zamir SW, et al. Transformers in medical imaging: a survey. *Med Image Anal.* Aug 2023;88:102802. [doi: [10.1016/j.media.2023.102802](https://doi.org/10.1016/j.media.2023.102802)] [Medline: [37315483](https://pubmed.ncbi.nlm.nih.gov/37315483/)]
16. Jayaraman P, Desman J, Sabounchi M, Nadkarni GN, Sakhuja A. A primer on reinforcement learning in medicine for clinicians. *NPJ Digit Med.* Nov 26, 2024;7(1):337. [doi: [10.1038/s41746-024-01316-0](https://doi.org/10.1038/s41746-024-01316-0)] [Medline: [39592855](https://pubmed.ncbi.nlm.nih.gov/39592855/)]
17. Blencowe N, Glasbey J, Heywood N, et al. Recognising contributions to work in research collaboratives: guidelines for standardising reporting of authorship in collaborative research. *Int J Surg.* Apr 2018;52:355-360. [doi: [10.1016/j.ijssu.2017.12.019](https://doi.org/10.1016/j.ijssu.2017.12.019)]
18. Boet S, Etherington C, Lam S, et al. Implementation of the operating room black box research program at the Ottawa hospital through patient, clinical, and organizational engagement: case study. *J Med Internet Res.* Mar 16, 2021;23(3):e15443. [doi: [10.2196/15443](https://doi.org/10.2196/15443)] [Medline: [33724199](https://pubmed.ncbi.nlm.nih.gov/33724199/)]
19. Suliburk JW, Buck QM, Pirko CJ, et al. Analysis of human performance deficiencies associated with surgical adverse events. *JAMA Netw Open.* Jul 3, 2019;2(7):e198067. [doi: [10.1001/jamanetworkopen.2019.8067](https://doi.org/10.1001/jamanetworkopen.2019.8067)] [Medline: [31365107](https://pubmed.ncbi.nlm.nih.gov/31365107/)]
20. Teo ZL, Jin L, Li S, et al. Federated machine learning in healthcare: a systematic review on clinical applications and technical architecture. *Cell Rep Med.* Feb 20, 2024;5(2):101419. [doi: [10.1016/j.xcrm.2024.101419](https://doi.org/10.1016/j.xcrm.2024.101419)] [Medline: [38340728](https://pubmed.ncbi.nlm.nih.gov/38340728/)]

Abbreviations

AI: artificial intelligence

IRCAD: Institut de Recherche contre les Cancers de l'Appareil Digestif

STANDING: Standards for Data Diversity, Inclusivity, and Generalisability

STARD-AI: Standards for Reporting Diagnostic Accuracy–Artificial Intelligence

TRIPOD-AI: Transparent Reporting of a Multivariable Model for Individual Prognosis or Diagnosis–Artificial Intelligence

Edited by Amy Schwartz; peer-reviewed by Shuang Zhao, Simon Laplante; submitted 30.May.2024; final revised version received 05.Jul.2025; accepted 12.Aug.2025; published 06.Mar.2026

Please cite as:

*Sampaio Soares A, Bano S, Castro LT, Pascoal M, Rocha R, Alves P, Mira P, Costa J, Chand M, Stoyanov D, Barata C
AI Models to Reduce Surgical Complications Through Intraoperative Video Analysis: Protocol for a Prospective Cohort Study*

JMIR Res Protoc 2026;15:e62734

URL: <https://www.researchprotocols.org/2026/1/e62734>

doi: [10.2196/62734](https://doi.org/10.2196/62734)

© António Sampaio Soares, Sophia Bano, Laura T Castro, Margarida Pascoal, Ricardo Rocha, Paulo Alves, Paulo Mira, Joao Costa, Manish Chand, Danail Stoyanov, Catarina Barata. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 06.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.