

Original Article

Open Access



Operative difficulty in laparoscopic cholecystectomy: considering the role of machine learning platforms in clinical practice

Isaac Tranter-Entwistle¹, Tim Eglinton^{1,2}, Saxon Connor², Thomas J. Hugh^{3,4}

¹Department of Surgery, The University of Otago Medical School, Dunedin 9016, New Zealand.

²Department of General Surgery, Christchurch Hospital, Christchurch 8011, New Zealand.

³Upper Gastrointestinal Surgical Unit, Royal North Shore Hospital and North Shore Private Hospital, St Leonards NSW 2065, Australia.

⁴Northern Clinical School, University of Sydney, Camperdown NSW 2006, Australia.

Correspondence to: Dr. Isaac Tranter-Entwistle, Department of Surgery, The University of Otago Medical School, 362 Leith Street, Dunedin 9016, New Zealand. E-mail: Isaac.Tranter-Entwistle@cdhb.health.nz

How to cite this article: Tranter-Entwistle I, Eglinton T, Connor S, Hugh TJ. Operative difficulty in laparoscopic cholecystectomy: considering the role of machine learning platforms in clinical practice. *Art Int Surg* 2022;2:46-56. <https://dx.doi.org/10.20517/ais.2022.01>

Received: 13 Jan 2022 **First Decision:** 8 Feb 2022 **Revised:** 24 Feb 2022 **Accepted:** 21 Mar 2022 **Published:** 25 Mar 2022

Academic Editor: Andrew A. Gumbs, Marialuisa Lugaresi **Copy Editor:** Jia-Xin Zhang **Production Editor:** Jia-Xin Zhang

Abstract

Aim: Computer vision is a subset of machine learning (ML) technology that allows automated analysis of large operative video datasets. The aim of this study was to use a commercially available ML-driven platform to evaluate a subjective grading of operative difficulty in laparoscopic cholecystectomy (LC).

Methods: Patients undergoing LC prospectively consented, and their operations were recorded. The intra-operative findings were prospectively graded (1-4) based on intraoperative gallbladder appearance assessments. Deidentified videos were uploaded to Touch Surgery™ and run through the platform's algorithm, providing automated analytics including the total operative length and operative phase length. The rate of critical view of safety (CVS) achievement was also included in the analysis.

Results: 206 LC were included. 27 LC were excluded due to incomplete video recording and were therefore not amenable to the final data analysis. Grade 1 and 2 patients had significantly shorter operative time than grade 3 and 4 patients [17min and 53s (IQR 15min and 24s- 21min and 38s) vs. 25 min and 49s (IQR 20min and 12s-38min and 38s) ($P < 0.010$)]. The operative phases for each step were significantly longer in patients with gallbladders graded



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



3 or 4 compared to those patients graded 1 or 2 ($P < 0.043$). The CVS was achieved in 94% of grade 1 patients, 88% of grade 2 patients, 85% of grade 3 patients and 73% of grade 4 patients ($P = 0.177$).

Conclusion: Increased operative time and decreased ability to achieve the CVS with more difficult intraoperative findings supports the utility of the proposed grading system. ML in surgery is a nascent field, but this study demonstrates the potential of commercially available platforms for use in operative analytics, documentation, audit and training of future surgeons.

Keywords: Laparoscopic cholecystectomy, machine learning, artificial intelligence, difficulty grading

INTRODUCTION

Computer vision is a subset of machine learning (ML) that allows automated analysis of large operative video datasets. Laparoscopic cholecystectomy is a high-volume procedure with consistent steps suitable for the application of ML techniques. Recent advances have included automated identification of operative steps and anatomical structures, but the impact of these technologies has been confined to research studies^[1,2]. Their use in clinical practice has been limited due to a lack of surgeon awareness of the potential applications, concerns regarding the black box nature of algorithms, and limited high-quality surgical video data sets. Given the significant barriers to entry in developing these systems, including computer science expertise and data requirements, it is possible the commercial versions of these tools will become increasingly widespread. In this context, surgeon-led consideration of how these tools add value in clinical practice is needed.

Traditionally, clinicians have used pre-operative variables to predict the degree of gallbladder inflammation and thus surgical difficulty^[3]. Increasingly intraoperative grading scores have been shown to be associated with operative outcomes and technical difficulty^[4-7]. Given outcomes are often related to actions taken intraoperatively, quantification of technical difficulty allows for operative benchmarking, prediction of post-operative outcomes, and development of research standards^[8]. We hypothesize that an artificial intelligence platform can confirm the impact of a “difficult” cholecystectomy by assessing a subjective intra-operative cholecystectomy grading system. The aim of this study was to use a commercially available ML-powered surgical video management and analytics platform (Touch Surgery™) to evaluate subjective intraoperative grading of operative difficulty during laparoscopic cholecystectomy using a stepwise workflow approach and thereby consider the implications for clinical practice.

METHODS

Study Design

Patients undergoing elective laparoscopic cholecystectomy and routine operative cholangiogram (IOC) by a single specialist hepatobiliary surgeon (North Shore Private Hospital, Sydney, Australia) were consented preoperatively to undergo video recording of their operation. This study was approved by the Ramsay Health Care research ethics committee (approval no. RG2020.153). Video footage from camera insertion to the removal of the specimen was captured as part of routine patient care with an intraoperative photo of the critical view of safety (CVS) taken in every operation and measured operative time excluded time setting up the equipment, establishing the pneumoperitoneum and closing the wounds. Laparoscopic Cholecystectomy procedures were recorded, saved, de-identified, and then uploaded to Touch Surgery™ (<https://www.touchsurgery.com/professional>), a web-based platform for surgical video storage and surgical analytics, powered by ML. Upon upload, all videos were run through the Touch Surgery RedactOR™ algorithm to ensure any remaining patient identifiable information was removed. RedactOR™ detects portions of the video where the camera is outside of the patient and pixelates the video stream in real-time

on upload to prevent the recording of any potentially identifiable information. Operations are automatically broken down into phases and steps to provide insights into surgical performance, variation, and standardization, which provides opportunities for pre-operative rehearsal and post-operative review. The underpinning ML is based on Convolutional Neural Networks architectures for classifying and extracting frames into their feature representation (step one). A single frame, however, is normally not sufficient to correctly identify the operative phase, as it may depict anatomical landmarks that appear throughout the operation. To overcome this limitation and process the temporal information together with the spatial information, these features are then fed into a Recurrent Neural Network (step two) to improve temporal consistency and representation^[9,10]. Touch Surgery™ phase identification is based on previous works including DeepPhase, EndoNet, and a phase recognition model with an F1 (a composite score used to assess ML accuracy generated by taking the mean of the positive predictive value and sensitivity) score of 91.1% in predicting phase of total knee joint replacement^[9-11]. The network used to annotate Laparoscopic Cholecystectomy videos in this paper was developed by Digital Surgery Ltd. (UK) using a large dataset of combined videos from surgeons of different countries and hospitals. It achieves 95% accuracy in detecting phase transitions in laparoscopic cholecystectomy. When tested on the video data included in this study, the model also achieved 95% accuracy. Qualified annotators, trained on surgically-validated guidelines, quality-assured the model outputs.

Operative Phases

In the present platform, Touch Surgery™ defined the surgical workflow phases for the automated analysis by liaising with key opinion leaders and consulting the literature^[12-15]. Based on this, the laparoscopic cholecystectomy videos were divided into the following five operative phases for the purposes of automated analysis:

- 1 Port insertion and gallbladder exposure.
- 2 Dissection of Calot's triangle.
- 3 Ligation and division of the cystic duct & artery.
- 4 Gallbladder dissection.
- 5 Specimen removal and closure.

CVS

Presence of the CVS was manually documented as part of the Touch Surgery™ digital analytics service by trained annotators in accordance with the SAGES safe cholecystectomy program^[16]. This approach has previously shown validity, with Deal *et al.*^[17] demonstrating a statistically significant correlation between expert and crowd workers' ratings of CVS achievement.

Grading of Operative Difficulty

The North Shore system uses a 4-point "operative difficulty" grading score which has been recorded prospectively in the operation record for every patient since 1998. This was modified from an earlier grading system first described by Hugh *et al.* in 1992 in an unselected consecutive series of 100 patients undergoing laparoscopic cholecystectomy^[5,18]. Assessment of the intraoperative findings was performed and documented at the commencement of the procedure by the attending surgeon in keeping with the scale as described by O'Neill *et al.* [Figure 1]^[5,18].

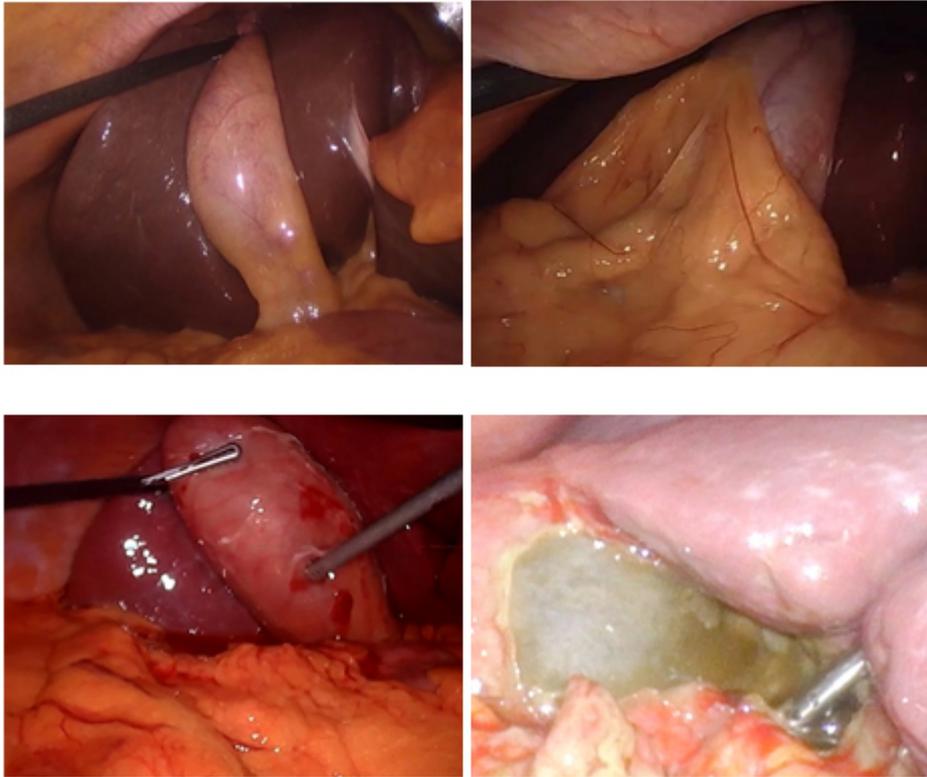


Figure 1. Intraoperative grades: Grade 1: Thin wall, normal-appearing gallbladder, no adhesions (Top left). Grade 2: Mildly abnormal-appearing GB (slightly thick-walled or distended) And/or thin-filmy GB adhesions (Top right). Grade 3: Moderately abnormal appearing GB (thick-walled, oedematous, with mucocele, or large distended gallbladder. And/or overlying moderate adhesions (Bottom left). Grade 4: Severely inflamed or grossly abnormal-appearing GB (e.g., necrotic or perforated). And/or extensive dense adhesions (Bottom right).

Inclusion and exclusion criteria

The present cohort includes both elective and acute patients presenting to a single surgeon HPB surgeon at the Royal North Shore Hospital and North Shore Private Hospital, St Leonards, NSW, Australia. To be eligible, captured videos had to have all phases including port insertion, dissection of Calot's triangle, ligation and division of the cystic duct and artery, gallbladder dissection, and specimen removal. Videos that did not have all five phases due to late recording or early stopping were excluded from the analysis.

Statistical analysis

Statistical analysis was performed using SciPy and Pingouin^[19,20]. D'Agostino-Pearson's test of normality was performed; where there was a normal distribution, a Levene test of variance was performed, or if non-parametric, Bartlett's test. Mann-Whitney U tests were performed for non-parametric samples with equal variance and Brunner-Munzel for those with unequal variance. For parametric samples with equal variance, a t-test was performed, or Welch's test for those with unequal variance.

RESULTS

During the study period, 233 patients consented to the video recording of their procedures, and from this group, 206 (88%) videos met the inclusion criteria. 27 LC were excluded due to incomplete video recording and were therefore not amenable to analysis. The videos analyzed included a consecutive series of patients operated on by a single surgeon over a 3-year period. Most operations were done electively, and in all cases,

a standardized operative approach including routine intra-operative cholangiography was undertaken. Demographic and peri-operative details of the cohort are seen in [Table 1](#).

The median operative time was 19min and 53s (IQR 15min and 53s-26min and 16s). In total, 143 (69%) patients were classified as either grade 1 or 2, with a median operative time of 17min and 53s (IQR 15min and 24s-21min and 38s). In comparison, 63 (31%) patients were classified as either grade 3 or 4 with a median operative time of 25 min and 49s (IQR 20min and 12s-38min and 38s). Operative time was significantly shorter for grade 1 and 2 than for the patients graded 3 or 4 ($P < 0.01$) [[Figure 2](#)]. The variation in operative length was greatest in patients who were assigned a grade of 3 or 4. The time differences and P -values between phases are documented in [Table 2](#).

There were 33 (16%) grade 1 patients, with a median operative time of 15min and 49s (IQR 13min and 14s-18min and 15s), and 110 (54%) grade 2 patients, with a median operative time of 18 min and 25s (IQR 15min and 45s-21min and 51s). Fifty-two (25%) grade 3 patients were analyzed with a median operative time of 23min and 48s (IQR 19min and 56s-33min and 34s), and 11 (5%) grade 4 patients' videos were analyzed, with a median operative time of 56min and 4s (IQR 41min and 18s-71min and 11s).

When the operations were analyzed according to the five predetermined operative steps, all phases took significantly longer to complete in grade 3 and 4 patients compared with grade 1 and 2 patients [[Table 2](#)] [[Figure 3](#)].

The rate of achievement of the CVS for each operative grade is shown in [Table 3](#). The rate of achievement of the CVS when comparing grade 1-2 and grade 3-4 was not significantly different ($P = 0.177$)

DISCUSSION

The ML-powered system allowed automated analysis of a large video dataset, confirming that the total operative time and individual operative phases were correlated with an intraoperative difficulty rating. Operative time is a consistent marker of technical ability and operative difficulty across the surgical literature, and grading of laparoscopic cholecystectomy difficulty has been shown to have validity in predicting outcome^[4-6,8,21-24]. This study provides an example of the emerging clinical utility of computer vision technology in providing automated operative analytics in clinical practice.

Accurate identification of the operative phase is important in allowing workflow planning and the development of intraoperative decision support systems. However, to have utility, operative phases need to be clinically relevant. While previous publications have considered the accuracy of automated phase identification, there is currently no universal standard in laparoscopic cholecystectomy^[25]. The present study investigated the clinical utility of automated phase identification by considering the impact of a subjective gradings score on operative phase times. A significant difference in phases times was seen across all phase times when comparing grade 1 and 2 gallbladders with grade 3 and 4 gallbladders. The major time difference between grades was seen in the time taken in initial exposure and the time to dissect Calot's triangle, which is arguably the most critical step in avoiding a bile duct injury. The image findings of the IOC were not captured as part of the laparoscopic recording, which meant this could not be included as a discrete phase in this study; however, routine performance ensured there was no biasing effect between groups. While further work is needed to create a unified standard of phase identification, the data presented here suggest clinical utility of the chosen phases.

Table 1. Patient and operative demographics

Preoperative variables	(n = 206)
Median Age (IQR)	52 (14-93)
Gender, (%)	
Female	68
Median Charlson Comorbidity Index (IQR)	2 (0-7)
Pancreatitis, (%)	3.9
Jaundice, (%)	2.5
Urgent/Semi-urgent operation, (%)	12
Operative variables	
IOC attempted, (%)	97
IOC successful, (%)	95
Postoperative outcomes	
Median Hospital stay in days (IQR)	1 (0-21)
Complications	
Minor (\leq Grade 2)	6.2%
Major ($>$ Grade 2)	1%

Table 2. Operative time by Grade

Phase	Grade 1 and 2 median time in minutes (minutes: seconds)	IQR (minutes: seconds)	Grade 3 and 4 median time in minutes (minutes: seconds)	IQR (minutes: seconds)	P-value
Port insertion and gallbladder exposure	02:02	01:28-02:54	04:35	02:55-07:55	$P = 0.00$
Dissection of Calots triangle	05:15	03:55-07:00	07:00	04:39- 10:29	$P < 0.01$
Ligation and division of cystic duct & artery	02:53	01:16-05:10	03:53	02:47-07:04	$P < 0.01$
Gallbladder Dissection	02:20	01:03-04:27	03:34	01:30-05:44	$P < 0.05$
Specimen removal and closure	03:24	02:31-05:00	04:06	03:02-07:23	$P < 0.01$

Table 3. Achievement of CVS by grade

Grade	Anterior CVS	Posterior CVS	No CVS	Total
Grade 1	91% (n = 30)	3% (n = 1)	6% (n = 2)	100% (n = 33)
Grade 2	86% (n = 95)	2% (n = 2)	12% (n = 13)	100% (n = 110)
Grade 3	83% (n = 43)	2% (n = 1)	15 % (n = 8)	100% (n = 52)
Grade 4	73% (n = 8)	0% (n = 0)	27% (n = 3)	100% (n = 11)

CVS: Critical view of safety.

Achievement of the CVS is an established requirement in safe cholecystectomy^[16,26]. Rates of CVS achievement are often overstated, with one study finding CVS was only achieved in 10.8% of patients despite a documented achievement rate of 80%^[27]. Intraoperative photo documentation of the CVS has been suggested as a quality control measure; however, this is surgeon dependent and necessitates subsequent external audit to ensure consistency^[28]. In contrast, routine intraoperative video recording removes barriers to capture and may ensure consistency of achievement^[29]. The high rate of CVS achievement in the current study (88%) is in keeping with operations being performed in the elective setting by a sub-specialist hepatobiliary surgeon. The inverse relationship between patient grade and CVS achievement demonstrated

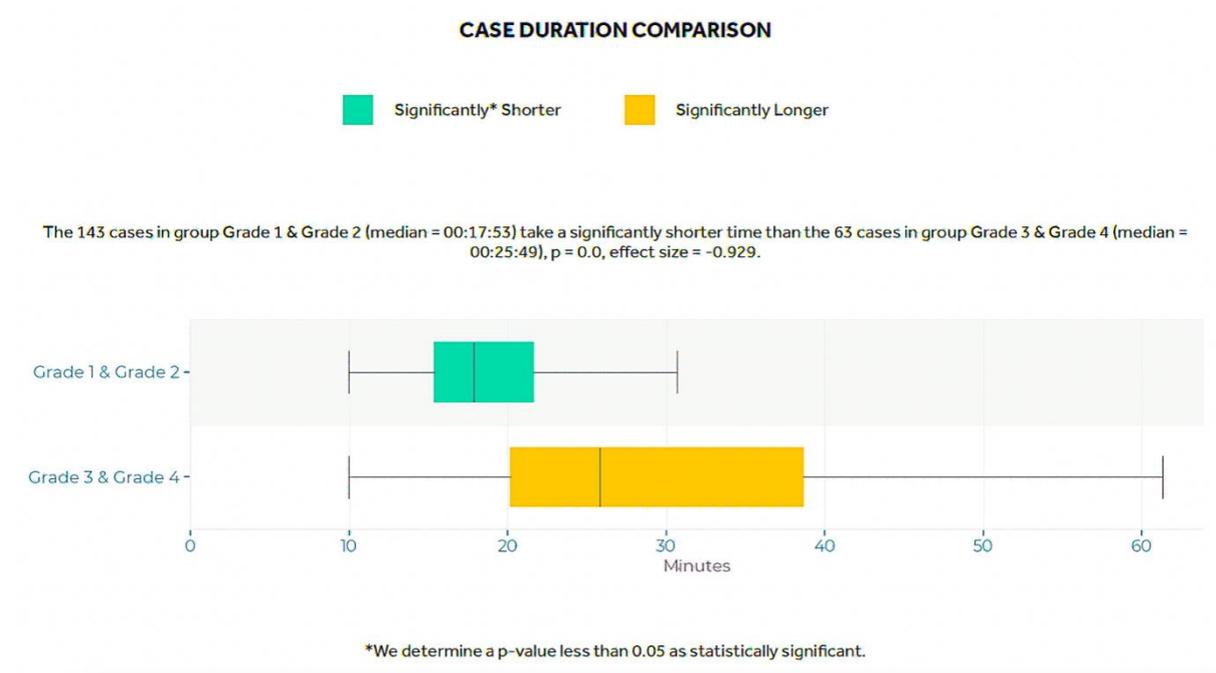


Figure 2. Median operative times by grade. Copyright. All rights reserved. Digital Surgery Ltd. 2021.

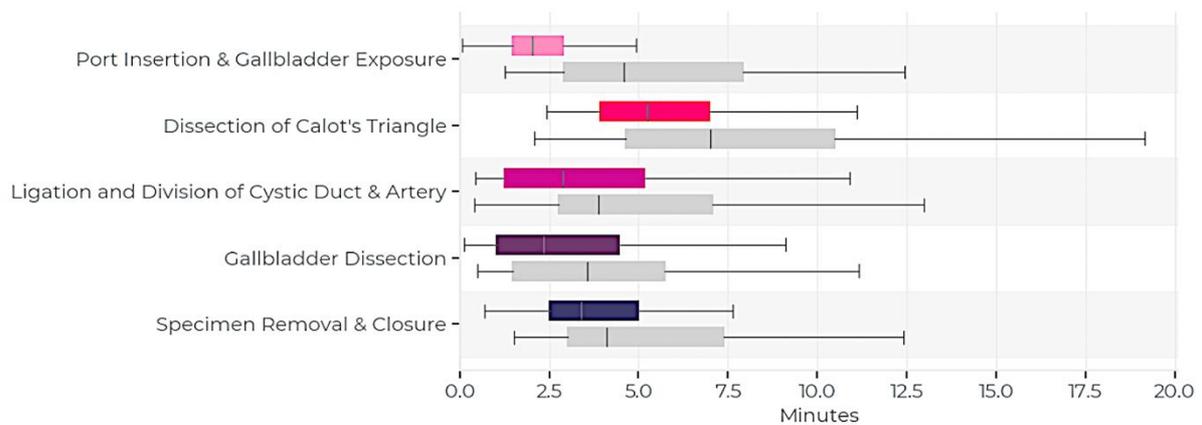


Figure 3. Median Operative Time by Phase. Phase duration comparisons between Grade 1 and 2 (colored boxes) and Grade 3 and 4 (grey boxes). All phases took significantly longer to complete in grade 3 and 4 patients compared with grade 1 and 2 patients. Copyright. All rights reserved. Digital Surgery Ltd. 2021.

is concordant with an accurate grading score. Broader validation could allow for a benchmark rate of CVS achievement, prompting audit and review if rates persistently drop below this. While in the future, a prospective analysis could provide intraoperative prompts with manual override to ensure the CVS is achieved.

Surgical curricula are increasingly relying on competency-based models as a means of capturing progress^[30-32]. This approach reflects the operative learning curve, in which trainees perform different segments of each operation under supervision before progressing to perform the entirety of the operation.

By creating agreed phases or steps of each operation as part of a training curriculum, these competencies can be captured, and accurate feedback provided. Capture and automated assessment of these phases with ML techniques is a logical step in this pathway. While manual review of large volumes of video is not feasible, employing AI allows automated analysis and segmentation of phases. This study provides timeframes for each stage of the operation that represents a technical gold standard as the operations were performed by an experienced laparoscopic hepatobiliary surgeon. Although further data is needed for each level of trainee and each grade of gallbladder difficulty, this forms the first part of establishing competency-based standards for a surgical procedure. In the future, failure to meet expected time requirements might trigger a manual review of the technique with surgeon mentors. Prospective capture with automated grading and analysis could allow for focused video review between surgeon and trainee. Routine operative difficulty grading would quantify the operative technical difficulty of the procedures trainees are undertaking. Given the operative technical skill and the operative difficulty grade are predictive of patient outcomes, both need to be taken into account when considering trainee progress^[4,5,8]. Understanding the degree of difficulty of the operations the trainee is undertaking and what phases of these are challenging would more accurately quantify the trainees' progression through their learning curve.

Given the documented utility of the classification system for quantifying the difficulty of laparoscopic cholecystectomy in both classical and ML evaluations, validation of clinical usefulness needs to be confirmed in a large cohort of surgeons at different operative levels. This would allow for the generation of normal curves for expected operating time for each phase of the identified operation. The novel test set from this study could potentially be used to develop automated identification of the intraoperative difficulty grade.

The present study focused on overall and phase timing as measures of operative difficulty as a means of considering the clinical utility of the computer vision platform. Time is only one aspect of operative performance that can be assessed using ML techniques. In particular, automated assessment of CVS attainment would represent a significant advancement. Other factors that could be captured automatically include the rate of gallstone spillage, the number of instrument changes, and the economy of instrument movement. Incorporating these and other factors in automated analysis could produce a more comprehensive assessment of operative techniques for both audit and training purposes.

AI models are able to segment and automatically identify critical operative steps¹. However, in most cases, this has involved retrospective capture and analysis of video in relatively small sample sizes, and this approach is limited by the physical time cost required for surgeon video labeling^[17]. Through pooled data sets, increased surgeon interest, and possibly unsupervised ML, these issues are slowly being addressed. It is even possible to envisage that soon the operative video will be stored as part of the patient notes and with an automated operative note generation. As these difficulties are overcome, and AI tools become readily available in the workplace, clinician involvement with decision-making regarding utility, utilization, and value will be needed. Engagement ensures the tools developed will be driven by clinical applicability and provide value in patient care rather than an externally imposed quality indicator adding to the already burgeoning paperwork load.

Computer vision tools lack easy explainability due to the opaque nature of the internal logic of their underpinning neural networks algorithms, limiting clinicians' ability to understand and explain how these tools reach their conclusions. This concern has been particularly pronounced when these tools are used to guide treatment decisions. Where the inability to explain fully how a decision is reached precludes a clinician's ability to undertake informed consent with their patients^[33]. However, the recent federal drug

administration approval of the GI Genius system for automated polyp identification following clinical trial data showing increased adenoma detection rate signifies the increasing acceptability of these systems where they are clinically explainable and improve outcomes^[34,35]. The current retrospective nature of surgical video analysis platforms means that they do not directly impact decision-making around patient treatment and therefore do not violate the principles of informed consent due to a lack of algorithmic explainability. While this lessens the ethical barrier to uptake, it is still imperative for clinicians to consider how they should be used in clinical practice and if outputs are consistent with clinical intuitions. Clinician input is therefore needed to link these systems to clinical practice and consider if their results have clinical explainability. In particular, while phase identification algorithms in laparoscopic cholecystectomy have shown reasonable accuracy, their consistency with real-life clinical intuition needs to be considered. In this context, the association seen between increasing operative time and increasing operative difficulty, particularly in the dissection of Calot's triangle, is consistent with clinical intuition and clinically explainable.

The study presents a single specialist surgeon cohort of prospectively captured laparoscopic cholecystectomy operations. While the universality of laparoscopic cholecystectomy means that from a technical perspective, this study is generalizable, this may not be true for the ML analysis. This is because these systems can be brittle with significant changes in analysis quality due to seemingly irrelevant changes in operative approaches or equipment^[36]. It should also be noted that the operative times cannot be extrapolated due to the procedures being undertaken by a single expert HPB surgeon. Further validation of intraoperative grading is needed in external data sets encompassing a broader number of centers. ML in surgery is a nascent field, but this study and others like it demonstrate the potential in operative analytics, documentation, audit and training of future surgeons.

DECLARATIONS

Acknowledgements

With thanks to Touch Surgery™ for their access to the platform, advice and provision of output data throughout this study.

Authors Contributions

Made substantial contributions to the conception, acquisition, analysis of the data, drafting and revision of this work: Tranter-Entwistle I, Eglinton T, Connor S, Hugh TJ.

Availability of Data and Research Materials

Data could be provided on reasonable request.

Financial support and sponsorship

None.

Conflicts of interest

Isaac Tranter-Entwistle has received funding from Medtronic (Touch Surgery™ is a subsidiary of Medtronic) to undertake a PhD through the University of Otago from February 2021. This study was conducted in 2020.

Thomas Hugh has undertaken consultancy for Touch Surgery™ separate to the present study and was not involved in the data collection or analysis of results in this study.

Saxon Connor has undertaken pro bono consultancy for Medtronic/Touch Surgery™ developing a freely available educational application around laparoscopic cholecystectomy, as well as video annotation as part of an unrelated study.

Tim Eglinton has no conflicts of interest to declare.

Ethical approval and consent to participate

This study was approved by the Ramsay Health Care research ethics committee (approval no. RG2020.153). Patients undergoing elective laparoscopic cholecystectomy and routine operative cholangiogram (IOC) by a single specialist hepatobiliary surgeon (North Shore Private Hospital, Sydney, Australia) were consented preoperatively to undergo video recording of their operation. All study participants provided informed consent.

Consent for publication

All videos were run through the Touch Surgery RedactORTM algorithm to ensure any remaining patient identifiable information was removed.

Copyright

© The Author(s) 2022.

REFERENCES

1. Hashimoto DA, Rosman G, Witkowski ER, et al. Computer vision analysis of intraoperative video: automated recognition of operative steps in laparoscopic sleeve gastrectomy. *Ann Surg* 2019;270:414-21. [DOI](#) [PubMed](#) [PMC](#)
2. Volkov M, Hashimoto DA, Rosman G, Meireles OR, Rus D. Machine learning and coresets for automated real-time video segmentation of laparoscopic and robot-assisted surgery. *Proc - IEEE Int Conf Robot Autom* :2017. p. 754-59. [DOI](#)
3. Gupta N, Ranjan G, Arora MP, et al. Validation of a scoring system to predict difficult laparoscopic cholecystectomy. *Int J Surg* 2013;11:1002-6. [DOI](#) [PubMed](#)
4. Madni TD, Leshikar DE, Minshall CT, et al. The Parkland grading scale for cholecystitis. *The American Journal of Surgery* 2018;215:625-30. [DOI](#) [PubMed](#)
5. Wennmacker SZ, Bhimani N, van Dijk AH, Hugh TJ, de Reuver PR. Predicting operative difficulty of laparoscopic cholecystectomy in patients with acute biliary presentations. *ANZ J Surg* 2019;89:1451-6. [DOI](#) [PubMed](#) [PMC](#)
6. Griffiths EA, Hodson J, Vohra RS, et al; West Midlands Research Collaborative. Utilisation of an operative difficulty grading scale for laparoscopic cholecystectomy. *Surg Endosc* 2019;33:110-21. [DOI](#) [PubMed](#) [PMC](#)
7. Hugh TB, Chen FC, Hugh TJ, Li B. Laparoscopic cholecystectomy. A prospective study of outcome in 100 unselected patients. *Med J Aust* 1992;156:318-20. [DOI](#) [PubMed](#)
8. Birkmeyer JD, Finks JF, O'Reilly A, et al; Michigan bariatric surgery collaborative. surgical skill and complication rates after bariatric surgery. *N Engl J Med* 2013;369:1434-42. [DOI](#) [PubMed](#)
9. Zisimopoulos O, Flouty E, Luengo I, et al. DeepPhase: surgical phase recognition in CATARACTS videos. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-lópez C, Fichtinger G, editors. *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018*. Cham: Springer International Publishing; 2018. p. 265-72. [DOI](#)
10. Kadkhodamohammadi A, Sivanesan Utharaj N, Giataganas P, et al. Towards video-based surgical workflow understanding in open orthopaedic surgery. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 2021;9:286-93. [DOI](#)
11. 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* 2017;36:86-97. [DOI](#) [PubMed](#)
12. Fried G, Neville A. *Mastery of endoscopic and laparoscopic surgery*. 4th ed. (Lee L. Swanström NJS, ed.). Lippincott Williams and Wilkins; 2014.
13. Pignata G, Bracale U, Lazzara F. *Laparoscopic surgery: key points, operating room setup and equipment*. Springer; 2016. [DOI](#)
14. Ellison CE, Zollinger RM. *Zollinger's Atlas of surgical operations*.
15. Bonjer JH. *Surgical principals of minimally invasive procedures: manual of the European association of endoscopic surgery*. Available from: <https://link.springer.com/book/10.1007/978-3-319-43196-> [Last accessed on 23 Mar 2022].
16. Pucher PH, Brunt LM, Fanelli RD, Asbun HJ, Aggarwal R. SAGES expert Delphi consensus: critical factors for safe surgical practice in laparoscopic cholecystectomy. *Surg Endosc* 2015;29:3074-85. [DOI](#) [PubMed](#)
17. Deal SB, Stefanidis D, Telem D, et al. Evaluation of crowd-sourced assessment of the critical view of safety in laparoscopic cholecystectomy. *Surg Endosc* 2017;31:5094-100. [DOI](#) [PubMed](#)
18. O'Neill RS, Wennmacker SZ, Bhimani N, van Dijk AH, de Reuver P, Hugh TJ. Unsuspected choledocholithiasis found by routine intra-operative cholangiography during laparoscopic cholecystectomy. *ANZ J Surg* 2020;90:2279-84. [DOI](#) [PubMed](#)
19. Virtanen P, Gommers R, Oliphant TE, et al; SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020;17:261-72. [DOI](#) [PubMed](#) [PMC](#)
20. Vallat R. Pingouin: statistics in Python. *J Open Source Softw* 2018;3:1026. [DOI](#)
21. Schneider DF, Mazeh H, Oltmann SC, Chen H, Sippel RS. Novel thyroidectomy difficulty scale correlates with operative times. *World J Surg* 2014;38:1984-9. [DOI](#) [PubMed](#) [PMC](#)

22. Tseng JF, Pisters PW, Lee JE, et al. The learning curve in pancreatic surgery. *Surgery* 2007;141:456-63. [DOI PubMed](#)
23. Bourgouin S, Mancini J, Monchal T, Calvary R, Bordes J, Balandraud P. How to predict difficult laparoscopic cholecystectomy? *Am J Surg* 2016;212:873-81. [DOI PubMed](#)
24. Cheng K, You J, Wu S, et al. Artificial intelligence-based automated laparoscopic cholecystectomy surgical phase recognition and analysis. *Surg Endosc* 2021. [DOI PubMed](#)
25. Garrow CR, Kowalewski KF, Li L, et al. Machine learning for surgical phase recognition: a systematic review. *Ann Surg* 2021;273:684-93. [DOI PubMed](#)
26. Wakabayashi G, Iwashita Y, Hibi T, et al. Tokyo guidelines 2018: surgical management of acute cholecystitis: safe steps in laparoscopic cholecystectomy for acute cholecystitis (with videos). *J Hepatobiliary Pancreat Sci* 2018;25:73-86. [DOI PubMed](#)
27. Nijssen MA, Schreinemakers JM, Meyer Z, van der Schelling GP, Crolla RM, Rijken AM. Complications after laparoscopic cholecystectomy: a video evaluation study of whether the critical view of safety was reached. *World J Surg* 2015;39:1798-803. [DOI PubMed](#)
28. Sebastian M, Sebastian A, Rudnicki J. Recommendation for photographic documentation of safe laparoscopic cholecystectomy. *World J Surg* 2021;45:81-7. [DOI PubMed PMC](#)
29. Mascagni P, Fiorillo C, Urade T, et al. Formalizing video documentation of the critical view of safety in laparoscopic cholecystectomy: a step towards artificial intelligence assistance to improve surgical safety. *Surg Endosc* 2020;34:2709-14. [DOI PubMed](#)
30. Greenberg JA, Minter RM. Entrustable professional activities: the future of competency-based education in surgery may already be here. *Ann Surg* 2019;269:407-8. [DOI PubMed](#)
31. Knox ADC, Gilardino MS, Kasten SJ, Warren RJ, Anastakis DJ. Competency-based medical education for plastic surgery: where do we begin? *Plast Reconstr Surg* 2014;133:702e-10e. [DOI PubMed](#)
32. Nousiainen MT, Mironova P, Hynes M, et al; CBC Planning Committee. Eight-year outcomes of a competency-based residency training program in orthopedic surgery. *Med Teach* 2018;40:1042-54. [DOI PubMed](#)
33. Kundu S. AI in medicine must be explainable. *Nat Med* 2021;27:1328. [DOI PubMed](#)
34. Repici A, Badalamenti M, Maselli R, et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. *Gastroenterology* 2020;159:512-520.e7. [DOI PubMed](#)
35. Repici A, Spadaccini M, Antonelli G, et al. Artificial intelligence and colonoscopy experience: lessons from two randomised trials. *Gut* 2022;71:757-65. [DOI PubMed](#)
36. Madani A, Namazi B, Altieri MS, et al. Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. *Ann Surg* 2020. [DOI PubMed PMC](#)