



Evaluation of surgical complexity by automated surgical process recognition in robotic distal gastrectomy using artificial intelligence

Masashi Takeuchi¹ · Hirofumi Kawakubo¹ · Takayuki Tsuji¹ · Yusuke Maeda¹ · Satoru Matsuda¹ · Kazumasa Fukuda¹ · Rieko Nakamura¹ · Yuko Kitagawa¹

Received: 25 May 2022 / Accepted: 28 January 2023 / Published online: 23 February 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Background Although radical gastrectomy with lymph node dissection is the standard treatment for gastric cancer, the complication rate remains high. Thus, estimation of surgical complexity is required for safety. We aim to investigate the association between the surgical process and complexity, such as a risk of complications in robotic distal gastrectomy (RDG), to establish an artificial intelligence (AI)-based automated surgical phase recognition by analyzing robotic surgical videos, and to investigate the predictability of surgical complexity by AI.

Method This study assessed clinical data and robotic surgical videos for 56 patients who underwent RDG for gastric cancer. We investigated (1) the relationship between surgical complexity and perioperative factors (patient characteristics, surgical process); (2) AI training for automated phase recognition and model performance was assessed by comparing predictions to the surgeon-annotated reference; (3) AI model predictability for surgical complexity was calculated by the area under the curve.

Result Surgical complexity score comprised extended total surgical duration, bleeding, and complications and was strongly associated with the intraoperative surgical process, especially in the beginning phases (area under the curve 0.913). We established an AI model that can recognize surgical phases from video with 87% accuracy; AI can determine intraoperative surgical complexity by calculating the duration of beginning phases from phases 1–3 (area under the curve 0.859).

Conclusion Surgical complexity, as a surrogate of short-term outcomes, can be predicted by the surgical process, especially in the extended duration of beginning phases. Surgical complexity can also be evaluated with automation using our artificial intelligence-based model.

Gastric cancer is the most common solid malignancy and a leading cause of mortality worldwide [1, 2]. Although radical gastrectomy with lymph node dissection (LND) is the standard treatment for gastric cancer, the complication rate remains high. Among gastrectomy cases in Japan, postoperative data show a complication rate of 6.8% and mortality of 1.2% [3]. Therefore, to safely perform surgery, the ability to estimate surgical complexity and determine the risk for complications is important. To date, there are several factors associated with surgical complexity for gastrectomy, including body mass index, visceral fat area, and tumor characteristics [4–6]. Another important factor for evaluating surgical complexity and predicting complications is

the surgeon's skill and experience because outcomes after gastrectomy have varied between certified and non-certified institutions or surgeons [7, 8]. Straightforward cases for surgeons experienced with gastrectomy are often complicated for inexperienced surgeons. Surgeons may discover intraoperative complexity because of easy bleeding, complex anatomic landmarks, degree of adhesions, tissue fragility, and abdominal cavity narrowing—all of which are difficult to evaluate preoperatively with precision [9, 10]. Stulberg et al. noted most efforts to date to improve outcomes have focused on systems of care that surround the surgical episode rather than actual processes for the surgery itself. In fact, as an intraoperative factor, the surgeon's technical performance is strongly associated with outcomes [11]. This finding suggests surgical outcomes can be attributed to intraoperative factors versus preoperative findings. Thus, it may be necessary to use intraoperative factors to objectively and accurately evaluate surgical complexity and predict

✉ Hirofumi Kawakubo
hkawakubo@z3.keio.jp

¹ Department of Surgery, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-Ku, Tokyo 160-8582, Japan

complications for each patient. However, some intraoperative factors, such as easy bleeding or anatomic landmark complexity, are difficult to evaluate because these are not objective. Therefore, this study focuses on the intraoperative surgical process.

Artificial intelligence (AI), particularly computer vision (CV), has significantly impacted imaging and video analysis. CV enables computers to understand meaningful information from images or videos. It has a considerable ability to support clinical decision-making in various medical fields, including automatic diagnosis of intestinal tumors during endoscopy and automated detection of pulmonary lesion on computed tomography (CT) [12, 13]. We previously reported that using deep learning AI could detect esophageal cancer on CT images with substantial accuracy, better than that of radiologists [14]. This technology is beginning to be used in surgical fields, particularly laparoscopic cholecystectomy. Several studies suggested possible automated recognition of surgical phases on laparoscopic cholecystectomy, showing surgical progress with an accuracy of 74.5%–97.3% [15, 16]. This tool informs us about what surgeons do during surgery in real-time, resulting in the actual sharing of surgical processes among surgical teams. Automated recognition of laparoscopic cholecystectomy phases can also help create a video summary for use as a postoperative educational tool. In addition to laparoscopic cholecystectomy, a similar tool for laparoscopic sleeve gastrectomy and colorectal cancer surgery has been established with considerable accuracy. An automated surgical process using phase recognition can intraoperatively evaluate surgical complexity. However, no AI-based surgical phase recognition systems have been reported for robotic procedures. Moreover, the clinical relevance of automatic surgical phase recognition has not yet been indicated.

We hypothesized that the surgical process reflects the surgical complexity during robotic distal gastrectomy (RDG). The aims of the current study were, therefore: 1) to investigate the association between surgical process and surgical complexity, such as the risk of complications in RDG; 2) to establish an AI-based automated surgical phase recognition for RDG by analyzing robotic surgical videos, and 3) to investigate the predictability of surgical complexity for RDG by AI. We believe that AI automation can help surgeons perform optimal decision-making during surgery, such as changing the operator to the expert or predicting complications. This research also could help create educational tools for understanding and using robotics, especially for young surgeons.

Methods

Data sets

This study retrospectively assessed consecutive 56 patients who underwent RDG with D1 + (10 cases) or D2 LND (46 cases) for gastric cancer at Keio University Hospital, Tokyo, Japan, between 2018 and 2021. Exclusion criteria were conversion to total gastrectomy or undergoing combined cholecystectomy during the same surgery. Although all datasets from 56 patients were used to establish the AI model, data for 46 patients who underwent D2 LND were investigated to evaluate surgical complexity. Patient clinical characteristics, including age, sex, clinical findings, and short-term outcomes, were retrospectively extracted from hospital electric records. Institutional Review Board (IRB) approval was obtained prior to the start of the study and we obtained patients' informed Consent.

Surgical procedure

Surgical indications and extent of the LND were determined by Japanese Gastric Cancer Treatment Guidelines [17]. With the patient in the supine position, RDG was performed using the da Vinci Xi system (Intuitive Surgical, Sunnyvale, California, USA). Three board-certified experts performed the surgeries. Accordingly, 4 ports for the da Vinci Xi and 1 port for the assistant were inserted.

After abdominal cavity entry, the omentum was incised 3 cm from the stomach wall toward the spleen. The incision continued until the left gastroepiploic vessels, which were then divided. Omentum dissection was continued to the right and down to the transverse colon. The right gastroepiploic vein was divided just above bifurcation of the anterior superior pancreaticoduodenal vein and right gastroepiploic vein. After right gastroepiploic artery division, pre-pancreatic soft tissues were removed.

After supraduodenal LND, the duodenum was transected using a 60-mm stapler. Suprapancreatic LND was performed with common hepatic LND, celiac LND, and left gastric LND. For cases in which D1 + LND was performed, proximal splenic and hepatoduodenal LND were omitted. After completely removing the lymph node of the lesser curvature side of the gastric wall, the stomach was transected using two or three 60-mm staplers. Although Billroth-I or Roux-en-Y reconstruction was performed by the surgeon's preference, Roux-en-Y tended to be chosen if the remnant stomach was small. For study analysis of the surgeon's learning curve, 46 patients were divided into 2 groups based on the timing of the surgeries at the institution during the study period: the early period group, when

the first 20 RDG surgeries were performed in our institution (before February 2020), and the late-period group, when the most recent surgeries were performed (after March 2020) [18].

Surgical complexity

To evaluate surgical complexity, the surgeon examined the association between 3 factors that are surrogates of surgical level—estimated total operative time, bleeding, and complications [4, 10, 19, 20]—and perioperative factors, such as age, sex, clinical stage, surgical process. In addition, a surgical complexity score based on these three factors was established, and we investigated the factors related to this score (Table 1). A score less than 2 was defined as *low complexity*; a score of 2 or more represented *high complexity*. In this analysis, we used the estimated total surgical time, which excludes the duration of reconstruction from total surgical time because we wanted to

Table 1 Definition of surgical complexity score

Variable	Value	Score
Extended total surgical time	Extended	1
	No-extended	0
Bleeding	Presence	1
	Absence	0
Complications	Presence	1
	Absence	0
Total difficulty score 0–3		

Total score less than 2 low complexity; total score 2 or more high complexity

use all data for patients who underwent Roux-en-Y or Billroth-I reconstruction. We divided the estimated total surgical time into two groups on the 25th percentile. According to the anesthesiologist's chart, bleeding was defined as the presence of blood loss during surgery, whereas no bleeding was defined as absence of blood loss. Any grade 1 or higher complication in the Clavien-Dindo classification was considered a postoperative complication.

Annotation for surgical phase recognition

The RDG phases were divided into 10 surgical phases (phases 1–10): (1) preparation, (2) left gastroepiploic LND, (3) infrapyloric LND, (4) supraduodenal LND, (5) duodenal resection, (6) suprapancreatic LND, (7) lesser curvature LND, (8) gastric resection, (9-a) Roux-en-Y reconstruction, (9-b) Billroth-I reconstruction, and (10) after dissection to completion of surgery (Fig. 1). We determined each phase's starting and ending points based on the surgical procedure and anatomic characteristics (Table 2). The phase for the suprapancreatic LND included common hepatic LND, celiac LND, left gastric LND, and proximal splenic and hepatoduodenal LND. The “no-step phase” indicated video sequences during camera removal from the abdominal cavity was also defined to clarify the time for cleaning the camera, port insertion, or instrument exchange. Two board-certified gastrointestinal surgeons performed video annotations manually and independently (MT and YM). Discrepancies in the annotation were addressed by discussion between these two surgeons.

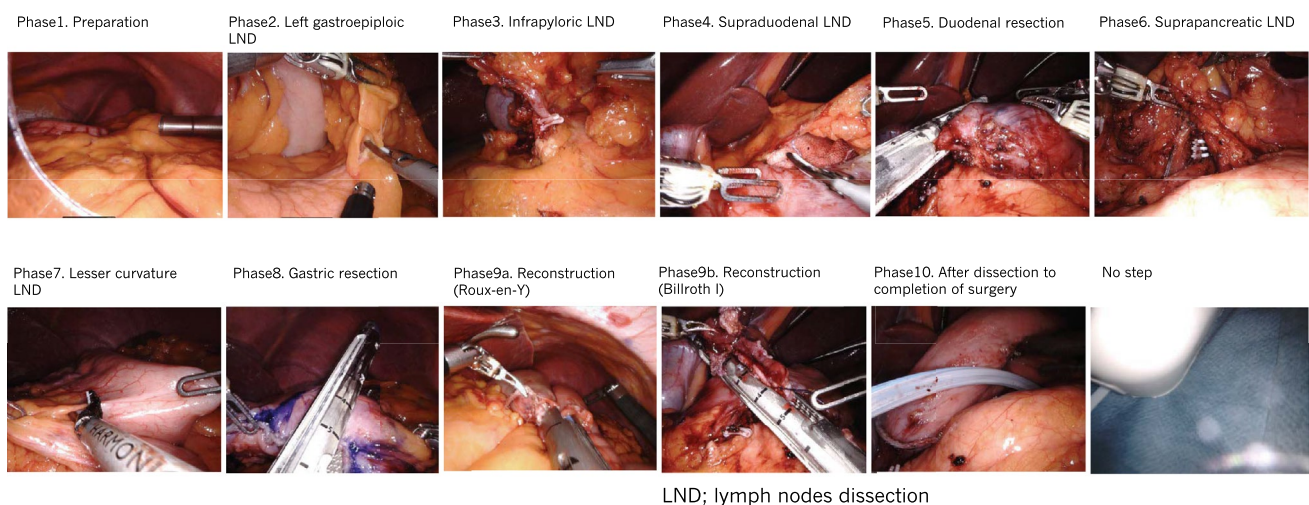


Fig. 1 Representative image for each of the ten surgical phases (phases 1–10); (1) preparation, (2) left gastroepiploic LND, (3) infrapyloric LND, (4) supraduodenal LND, (5) duodenal resection,

(6) suprapancreatic LND, (7) lesser curvature LND (8) gastric resection, (9-a) Roux-en-Y reconstruction, (9-b) Billroth-I reconstruction, and (10) after dissection to completion of surgery

Table 2 Definition of each phase of robotic gastrectomy

No	Step	start	end
1	Preparation	Camera is inserted into abdominal cavity	Start next phase
2	Left gastroepiploic LND	Start omental incision 3 cm from stomach wall toward spleen	End dissection of left gastroepiploic LND
3	Infrapyloric LND	Start omental incision 3 cm from stomach wall toward right gastroepiploic vein	End dissection of infrapyloric LND
4	Supraduodenal LND	Start lesser omentum incision between supraduodenal vessel and right gastric vessel	End dissection of supraduodenal LND
5	Duodenal resection	Start using stapler to divide duodenum	Finish dividing duodenum
6	Suprapancreatic LND	Start lesser omentum incision toward cardia, or start dissection for suprapancreatic LND	End dissection of suprapancreatic LND
7	Lesser curvature LND	Start dissection for soft tissue attached to lesser curvature side of gastric wall	End dissection of lesser curvature LND
8	Gastric resection	Start using stapler to divide stomach	Finish dividing stomach
9-a	Roux-en-Y reconstruction	Start treating small bowel or stomach for reconstruction	Finish reconstruction
9-b	Billroth-I reconstruction	Start treating duodenum or stomach for reconstruction	Finish reconstruction
10	After dissection to completion of surgery	End of all dissection	Finish all procedures
11	No step	Camera removed from abdominal cavity	Camera is inserted into abdominal cavity

LND lymph node dissection

AI models for computer vision

We used TeCNO [21], which uses a multi-stage temporal convolutional network for the hierarchical prediction of surgical phases. This approach can model the temporal context of phases over relatively long periods during the procedure. We applied these parameters to train the models: 12 output features, input image height of 256 and width of 256, and a sampling rate of 1. We used four-fold cross validation to train and assess the AI model; a random set comprising 75% of the videos (42 videos) was used as the training set, and the remaining 25% (14 videos) was used as the test set. This process was repeated 4 times to use all datasets for the testing. The model was implemented in Python 3.6. Frames were extracted from each video at a rate of one frame per second (fps), with an average of $14,590 \pm 5135$ frames per video. The AI models' performance was assessed by comparing predictions to the reference annotated by the surgeon. We assessed the performance by normalized confusion matrices (NCM), precision, recall, F-value, and accuracy. Rows in the NCM corresponded to the annotated phases (ground truth), whereas columns corresponded to the predicted phases. Values in the diagonal elements of the NCM represented, for each phase, the proportion of time points where the prediction was correct (true positive rate). These measurements were investigated as follows: *accuracy* [(true positive + true negative)/(true positive + false positive + false negative + true negative)]; *precision* [true positive/(true positive + false positive)]; *recall* [true positive/(true positive + false negative)]; and *F-value* [$2 \times (\text{recall} \times \text{precision}) / (\text{recall} + \text{precision})$].

Statistical analysis

All statistical analyses were calculated using Stata/IC 16 for Mac (StataCorp, Texas, USA), with a *p*-value of <0.05 indicating statistical significance. We calculated between-group differences using the chi-square test for categorical variables and the Mann–Whitney *U* test for continuous variables. Finally, accuracy for the prediction of surgical complexity was confirmed using the area under the curve (AUC), determined by analysis of the receiver operator characteristics curve.

Results

Patient characteristics

Of all 56 videos, 46 patients (35 male, 11 female) underwent RDG with D2 LND and were used for analysis to evaluate the relationship between surgical process and surgical complexity; the other 10 patients underwent D1 + LND. Clinical staging was stage I for 38 patients (83%) and stage II or higher for 8 patients. In 1 patient each, 5 complications were observed: (1) postoperative bleeding, (2) pancreatic fistula, (3) anastomotic leakage, (4) pulmonary embolism, and (5) delayed gastric emptying.

Relationship between surgical complexity and perioperative factors

The relationship between surgical complexity and several perioperative factors were investigated (Supplemental Table 1). We determined the cutoff value of the estimated

total surgical time as 18,315 s, the point of 25% percentile. For the estimated total operation duration, higher clinical stage ($p = 0.005$); higher clinical T stage ($p = 0.045$); early surgical period at the institution ($p = 0.003$); and surgical duration for all phases, except phase 10 (after dissection to surgery completion); and the no-step phase were significant-associated factors for the extended total surgical duration. For bleeding, significant-associated factors were extended total surgical duration ($p < 0.001$); Roux-en-Y reconstruction ($p = 0.027$); and extended duration of these phases: preparation ($p = 0.004$), left gastroepiploic LND ($p = 0.024$), infrapyloric LND ($p < 0.001$), duodenal resection ($p = 0.024$), suprapancreatic LND ($p < 0.001$), and the no-step phase ($p = 0.002$). For complications, significant-associated factors were extended total surgical duration ($p = 0.039$), Roux-en-Y reconstruction ($p = 0.047$), and extended duration of infrapyloric LND ($p < 0.001$), and of suprapancreatic LND ($p = 0.023$).

We also focused on the duration of early phases, which showed a similar tendency: preparation to left gastroepiploic LND (phases 1–2) and preparation to infrapyloric LND (phases 1–3), in which significant differences in bleeding and estimated surgical duration were observed.

We compared surgical complexity scores between high (score 2 or more) and low (score 1 or less) complexity. Significantly associated factors of high complexity were: clinical stage II or more ($p = 0.001$); reconstruction method (Roux-en-Y, $p = 0.044$); extended duration for these phases: preparation ($p = 0.002$), left gastroepiploic LND ($p < 0.001$), infrapyloric LND ($p < 0.001$), supraduodenal LND ($p = 0.003$), duodenal resection ($p = 0.003$), suprapancreatic LND ($p < 0.001$), lesser curvature LND ($p = 0.002$), and gastric resection ($p = 0.009$); duration from phase 1 to phase 2 ($p < 0.001$); and duration from phase 1 to phase 3 ($p < 0.001$). Furthermore, the early stage, which shows the first 20 cases in our institution, was also a significant factor ($p < 0.001$). These AUC values were compared (Supplemental Fig. 1) to evaluate the predictability for high complexity. Duration from phase 1 to phase 3 had the highest AUC values; 0.913.

Establishment of automated surgical phase recognition

We established the AI model to recognize the surgical processes using 56 videos with an overall accuracy of 87%. Supplemental Fig. 2 shows the NCM phases, which indicate true positive rates (diagonal items) ranging from 62% (Billroth-I reconstruction phase) to 96% (no-step and duodenal resection phases). Supplemental Table 2 shows other accuracy statistics (F-value, precision, and recall).

To visualize the predictive accuracy of our model, Fig. 2 shows timelines for two representative cases. The upper and

lower timelines show the annotated and predicted phases, respectively. These cases achieved nearly complete agreement between ground truth and predicted phases.

Relationship between surgical complexity and surgical processes predicted by AI

Several AUC values for phase duration were compared (Fig. 3) to evaluate the AI model's predictability of high complexity. The AUC value of predicted duration from phase 1 to phase 2 and duration from phase 1 to phase 3 were 0.865 and 0.860, respectively, which is a higher value than preoperative factors.

Discussion

This study showed that surgical complexity as a surrogate of short-term outcomes could be predicted by surgical process, especially with extended duration of the beginning phases, based on extended total surgical duration, intraoperative bleeding, and postoperative complications. We established the AI-based system to recognize the surgical phase automatically with high accuracy. Surgical complexity can also be evaluated automatically by our present system. This approach enables intraoperative decision-making, such as optimal timing of changing the surgeon to expert, to predict intraoperative bleeding and complications. To the best of our knowledge, this report is the first to use AI to determine surgical complexity.

Evaluating surgical complexity is important to avoid complications and propose optimal decision-making. To date, many factors such as obesity, tumor size, and advanced stage were reported as risk factors for surgical complexity in gastrointestinal surgery [4, 6]. We can preoperatively use these patient factors to determine surgical complexity; however, short-term outcomes are associated with patient factors evaluated preoperatively and intraoperative factors such as surgeon proficiency. Therefore, the importance of evaluating surgical skills is increasing [8]. Indeed, the preoperative prediction of surgical complexity has several advantages, such as additional preparation before surgery. However, intraoperative findings can provide more information than preoperative ones. The connection of pre- and intraoperative findings can help predict surgical skill as a next step.

To assess surgical skill, objective indications are needed. Stulberg et al. showed that higher technical skill scores for colectomy may be associated with lower complication rates; however, surgeons scored skills based on performance such as instrumental handling or operative flow [11]. Han et al. showed the Kappa value among surgeons was low for assessing surgical skills, suggesting consistent assessment for surgical skill is challenging [7]. Thus, the intraoperative

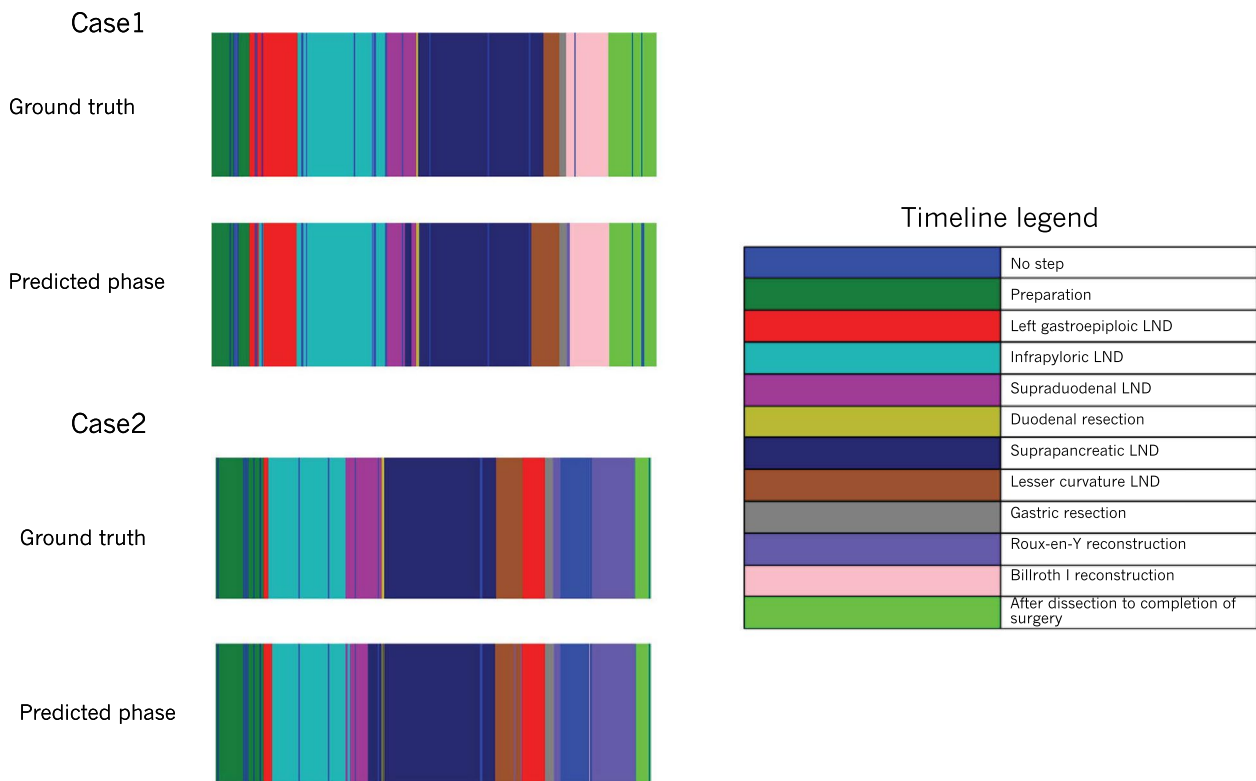
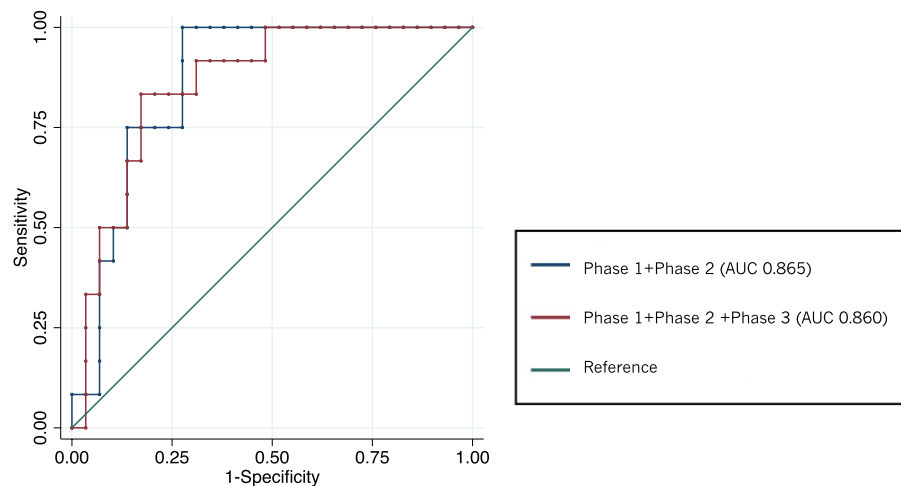


Fig. 2 Example of timeline visualization of the phases from the videos for TeCNO prediction. Upper row, ground truth; lower row, TeCNO predictions

Fig. 3 Receiver operating characteristic curve analysis for the beginning of surgical process predicted by artificial intelligence



process may be a good objective indicator for evaluating surgical skills. Moreover, it may reflect the patient factors, such as anatomic complexity and obesity, because of related intraoperative delays. This system would be useful for predicting complications with high accuracy and to support decision-making, such as changing the operator to expert during surgery and modifying surgical procedures.

In our study, the beginning phases, from the start to end of left gastroepiploic LND or infrapyloric LND, had high

predictability for surgical complexity. These phases included many technical procedures (e.g., adhesiotomy, instrument handling, retraction, mobilization, vessel resection), and thus durations were strongly associated with surgical skill and complexity.

In addition to evaluating surgical complexity, automated phase recognition can be useful in clinical practice settings, such as surgical summaries, operating room assistance, and education. Several web-based educational

platforms are widely used for minimally invasive surgery [22, 23]. Video platforms now have a major role in surgical education due to the coronavirus disease 2019 pandemic. We can easily and automatically index actual surgical steps in RDG on these platforms. Furthermore, this approach may also improve operating room efficiency by helping surgical staff prepare the next equipment or patients by showing the remaining duration and phase transition timing. To date, laparoscopic cholecystectomy has been most common surgical procedure in which automated phase recognition was attempted, given its lower complexity and robust standardization in general surgery. This study is the first to apply automated phase duration to RDG, the most common upper gastrointestinal surgery, which reinforces the importance of our research.

Our NCM results showed that AI had relatively low accuracy for recognizing reconstruction compared with other phases. First, we conducted two types of reconstructions based on the surgeon's preference—Roux-en-Y and Billroth-I reconstruction. A sufficient number of datasets may result in better outcomes; more datasets for each reconstruction are needed to increase accuracy for these phases in the future.

This study has several limitations, primarily its single-center design and inclusion of patients who underwent RDG by only several surgeons in one department, which have led to overfitting. Second, phase duration may not consistently directly show surgical skill because processes may be often delayed unpredictably for reasons such as machine failure. Ideally, automated recognition of processes would be an essential task, such as recognizing whether an organ was injured. However, procedures cannot be evaluated consistently for gastrectomy [7], which challenges automated recognition by AI [24]. Thus, it is reasonable to first focus on objective indicators such as phase duration. In addition to phase recognition, automated procedure recognition is needed in the future. Third, although the AUC value for surgical processes was higher than that for the surgical period, the learning curve of our procedure for RDG may also be associated with extended total duration and bleeding, as indicated by the surgical complexity score. To determine the actual cutoff value for the duration of beginning phases, only cases for which the surgeon has reached the learning curve should be used. Fourth, we had adopted three factors; extended total surgical time, bleeding and complication; as binary variables to establish a simple scoring system, which can enable easy interpretation. However, diminishing the resolution of data are a limitation of our surgical complexity scoring system. More detailed subdivision of each variable is necessary to reflect surgical complexities precisely. However, a detailed scoring system could not be established because of the small amount of data. The collection of more video data may enable us to establish a detailed scoring system.

In conclusion, our study shows that surgical complexity as a surrogate of short-term outcomes can be predicted by surgical processed, especially in the extended duration of beginning phases. Surgical complexity can also be evaluated automatically by our present AI-based model, and it enables intraoperative decision-making and prediction of intraoperative bleeding and complications.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00464-023-09924-9>.

Acknowledgements We wish to thank Kumiko Motooka, a staff member at the Department of Surgery in Keio University School of Medicine, for her help with the preparation of this manuscript.

Declarations

Disclosure Masashi Takeuchi, Hirofumi Kawakubo, Takayuki Tsuji, Yusuke Maeda, Satoru Matsuda, Kazumasa Fukuda, and Rieko Nakamura have no conflicts of interest or financial ties to disclose. Yuko Kitagawa received lecture fees from CHUGAI PHARMACEUTICAL CO., LTD., TAIHO PHARMACEUTICAL CO., LTD, ASAHI KASEI PHARMA CORPORATION, Otsuka Pharmaceutical Factory Inc., SHIONOGI & CO., LTD., Nippon Covidien Inc., ONO PHARMACEUTICAL CO., LTD., Bristol-Myers Squibb K.K. Author Y.K was supported by grants from CHUGAI PHARMACEUTICAL CO., LTD., TAIHO PHARMACEUTICAL CO., LTD, Yakult Honsha Co. Ltd., AsahiKASEI Co., Ltd., Otsuka Pharmaceutical Co., Ltd., Takeda Pharmaceutical Co., Ltd., ONO PHARMACEUTICAL CO., LTD., TSUMURA & CO., Kyouwa Hakkou Kirin Co., Ltd., DAINIPPON SUMITOMO PHARMA Co., Ltd., EA Pharma Co., Ltd., Astellas Pharma Inc., Toyama Chemical Co., Ltd., MEDICON INC., KAKEN PHARMACEUTICAL CO. LTD., Eisai Co., Ltd., Otsuka Pharmaceutical Factory Inc., TEIJIN PHARMA LIMITED., NIHON PHARMACEUTICAL CO., LTD., and Nippon Covidien Inc. Author Y.K held an endowed chair provided by CHUGAI PHARMACEUTICAL CO., LTD. and TAIHO PHARMACEUTICAL CO., LTD, outside the submitted work.

References

1. Joshi SS, Badgwell BD (2021) Current treatment and recent progress in gastric cancer. *CA Cancer J Clin* 71:264–279
2. Kamangar F, Dores GM, Anderson WF (2006) Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol* 24:2137–2150
3. Kakeji Y, Takahashi A, Hasegawa H et al (2020) Surgical outcomes in gastroenterological surgery in Japan: report of the National Clinical Database 2011–2018. *Ann Gastroenterol Surg* 4:250–274
4. Roh CK, Lee S, Son S-Y, Hur H, Han S-U (2021) Risk factors for the severity of complications in minimally invasive total gastrectomy for gastric cancer: a retrospective cohort study. *J Gastric Cancer* 21:352
5. Kunisaki C, Makino H, Takagawa R et al (2009) Predictive factors for surgical complications of laparoscopy-assisted distal gastrectomy for gastric cancer. *Surg Endosc* 23:2085–2093
6. Takeuchi M, Ishii K, Seki H et al (2016) Excessive visceral fat area as a risk factor for early postoperative complications of total

- gastrectomy for gastric cancer: a retrospective cohort study. *BMC Surg* 16:10–16
7. Han SU, Hur H, Lee HJ et al (2021) Surgeon quality control and standardization of D2 lymphadenectomy for gastric cancer: a prospective multicenter observational study (KLASS-02-QC). *Ann Surg* 273:315–324
 8. Suda K, Yamamoto H, Nishigori T et al (2021) Safe implementation of robotic gastrectomy for gastric cancer under the requirements for universal health insurance coverage: a retrospective cohort study using a nationwide registry database in Japan. *Gastric Cancer*. <https://doi.org/10.1007/s10120-021-01257-7>
 9. Li P, Huang CM, Lin JX et al (2016) A preoperatively predictive complexity scoring system for laparoscopic spleen-preserving splenic hilar lymph node dissection for gastric cancer: experience from a large-scale single center. *Surg Endosc* 30:4092–40101
 10. Serra-Aracil X, Rebasas-Cladera P, Mora-Lopez L, Pallisera-Lloveras A, Serra-Pla S, Navarro-Soto S (2020) A scoring system to predict complex transanal endoscopic surgery. *Surg Endosc* 34:4828–4836
 11. Stulberg JJ, Stulberg JJ, Huang R et al (2020) Association between surgeon technical skills and patient outcomes. *JAMA Surg* 155:960–968
 12. Esteva A, Kuprel B, Novoa RA et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542:115–118
 13. Zhao W, Yang J, Sun Y et al (2018) 3D deep learning from CT scans predicts tumor invasiveness of subcentimeter pulmonary adenocarcinomas. *Cancer Res* 78:6881–6889
 14. Takeuchi M, Seto T, Hashimoto M et al (2021) Performance of a deep learning-based identification system for esophageal cancer from CT images. *Esophagus* 18:612–620
 15. Garrow CR, Kowalewski KF, Li L et al (2021) Machine learning for surgical phase recognition: a systematic review. *Ann Surg* 273:684–693
 16. Guédon ACP, Meij SEP, Osman KN et al (2020) Deep learning for surgical phase recognition using endoscopic videos. *Surg Endosc*. <https://doi.org/10.1007/s00464-020-08110-5>
 17. Shiva T (2021) Japanese gastric cancer treatment guidelines 2018. *Gastric Cancer* 24(1):1–21
 18. Kang BH, Xuan Y, Hur H, Ahn CW, Cho YK, Han SU (2012) Comparison of surgical outcomes between robotic and laparoscopic gastrectomy for gastric cancer: the learning curve of robotic surgery. *J Gastric Cancer* 12:156–163
 19. Wang X, Yao Y, Qian H, Li H, Zhu X (2019) Longer operating time during gastrectomy has adverse effects on short-term surgical outcomes. *J Surg Res* 243:151–159
 20. Escal L, Nougaret S, Guiu B et al (2018) MRI-based score to predict surgical complexity in patients with rectal cancer. *Br J Surg* 105:140–146
 21. Czempiel T, Paschali M, Keicher M et al (2020) TeCNO: Surgical phase recognition with multi-stage temporal convolutional networks. *Lect Notes Comput Sci* 12263:343–352
 22. Yuksel C, Culcu S (2022) New learning area in laparoscopic gastrectomy for gastric cancer: YouTube or WebSurg? *J Minim Access Surg* 18:129–135
 23. Mutter D, Vix M, Dallemagne B, Perretta S, Leroy JM (2011) WeBSurg: an innovative educational Web site in minimally invasive surgery—principles and results. *Surg Innov* 18:8–14
 24. Mascagni P, Vardazaryan A, Alapatt D et al (2020) Artificial intelligence for surgical safety. *Ann Surg* 275(5):955–961

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.