



Systematic review on the use of artificial intelligence to identify anatomical structures during laparoscopic cholecystectomy: a tool towards the future

Diletta Corallino^{1,2} · Andrea Balla^{3,4} · Diego Coletta⁵ · Daniela Pacella⁶ · Mauro Podda⁷ · Annamaria Pronio⁸ · Monica Ortenzi⁹ · Francesca Ratti¹⁰ · Salvador Morales-Conde^{3,4} · Pierpaolo Sileri¹¹ · Luca Aldrighetti¹⁰

Received: 23 July 2024 / Accepted: 12 February 2025
© The Author(s) 2025

Abstract

Purpose Bile duct injury (BDI) during laparoscopic cholecystectomy (LC) is a dreaded complication. Artificial intelligence (AI) has recently been introduced in surgery. This systematic review aims to investigate whether AI can guide surgeons in identifying anatomical structures to facilitate safer dissection during LC.

Methods Following PROSPERO registration CRD-42023478754, a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)-compliant systematic search of MEDLINE (via PubMed), EMBASE, and Web of Science databases was conducted.

Results Out of 2304 articles identified, twenty-five were included in the analysis. The mean average precision for biliary structures detection reported in the included studies reaches 98%. The mean intersection over union ranges from 0.5 to 0.7, and the mean Dice/F1 spatial correlation index was greater than 0.7/1. AI system provided a change in the annotations in 27% of the cases, and 70% of these shifts were considered safer changes. The contribution to preventing BDI was reported at 3.65/4.

Conclusions Although studies on the use of AI during LC are few and very heterogeneous, AI has the potential to identify anatomical structures, thereby guiding surgeons towards safer LC procedures.

Keywords Artificial intelligence · Laparoscopic cholecystectomy · Minimally invasive surgery · Bile duct injuries

Introduction

Laparoscopic cholecystectomy (LC) is the gold standard treatment for symptomatic cholelithiasis and one of the most performed surgical procedures worldwide, both in elective and emergency settings and by senior and trainee surgeons [1–4]. The most dreaded complication of LC is bile duct injuries (BDIs) and a primary cause of BDI is a misunderstanding of the anatomical structures [1, 5]. The iatrogenic BDI rate after LC is higher than open cholecystectomy and ranges between 0.4 and 1.5% [5]. BDIs are responsible

for further surgical, endoscopic, or radiological procedures, increasing cost management of patients [5, 6].

Several optimal strategies for prevention of the BDIs are reported in the literature, which include the systematic use of the Critical View of Safety (CVS), the bailout procedures, such as sub-total cholecystectomy and many others [7]. The use of indocyanine green (ICG) fluorescence and the possibility to perform a fluorescent cholangiography (FC) during LC allows the visualization of extrahepatic biliary structures and facilitates intraoperatively dissection, aiming to reduce the BDI risk [8–19]. However, in the case of obese patients or acute cholecystitis, misinterpretation of the biliary anatomy can still occur, with a high risk of BDI [20–23].

Artificial intelligence (AI) has been recently introduced in MIS [24–28]. Computer vision is a field of AI that focuses on identifying and accurately analysing digital images, videos and other input, simulating and optimizing human recognition ability [25–28]. Although the use of deep learning

Diletta Corallino and Andrea Balla shared first authorship.

Pierpaolo Sileri and Luca Aldrighetti shared last authorship.

Extended author information available on the last page of the article

and computer vision in diagnostics, especially endoscopic, is now well established, its real-time application during surgery is still in the earliest stages [24–28].

The present systematic review aims to investigate whether AI can guide surgeons in identifying anatomical structures to facilitate safer dissections during LC.

Materials and methods

Institutional review board approval and informed consent from participants were unnecessary for the present study. This study was registered in the International Prospective Register of Systematic Reviews (PROSPERO; CRD-42023478754).

Search strategy

A systematic review of published articles according to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) Statement 2020 [29] and according to the Cochrane Handbook for systematic reviews of interventions was conducted [30]. The aim of the present study was to explore the potential of AI to improve the safety of LC by aiding surgeons in the identification of anatomical structures.

The PICO question was generated from a discussion within the authors (D.C., D.C., D.P., A.B.).

The following PICO question was adopted:

P(opulation). Patients undergoing LC.

I(ntervention). Identification of the anatomical landmarks during LC by AI.

C(omparison). Identification of the anatomical landmarks during LC without AI.

O(utcomes). Feasibility of anatomical structures identification during LC using AI.

Study identification

A computerized search was performed in MEDLINE (via PubMed), EMBASE and Web of Science databases for articles published up to 30/09/2024 [31], without language restrictions (Supplementary Material-Table 1). Reviews, systematic reviews, meta-analysis, comments, case reports, congress abstracts, correspondence and letters to editor, editorials, technical surgical notes, imaging studies and studies with animals' involvement were excluded. For a more objective evaluation, studies in which AI was used to identify adverse effects, timing, operators, and surgical tools, were excluded from this systematic review.

Eligibility criteria, screening process, and data extraction

All articles in which AI was used to anatomical structures identification, groups of anatomical structures that configure safe/unsafe dissection zones, or CVS assessment during LC were eligible for inclusion in the present study. Two independent reviewers (D.C. and D.C.) conducted the screening process and data extraction in a double-masked fashion. Discrepancies were resolved with a discussion with a third reviewer (A.B.). Data collected from each study included the following predefined items: (1) Study identifier (first author, year of publication); (2) Study design; (3) Population; (4) Type of dataset; (5) Inclusion and exclusion criteria; (6) Anatomical structures identified; (7) AI model; (8) "Ground truth" establishment; (9) Validation and application of the model; (10) Outcome measure; (11) Main results; (12) Conclusion; (13) Risk of bias/quality assessment. Data were stored in the Microsoft Excel program (Microsoft Corporation, Redmond, Washington, USA).

Outcomes of interest

According to the PICO criteria, the aim of this systematic review is to report the current evidence of the use of AI for anatomical structures identification during LC.

Structure identification concerns the hepatocystic anatomy: gallbladder, cystic duct (CD), cystic artery (CA), cystic plate (CP), common bile duct (CBD), hepatocystic triangle (HCT), lower edge of the left medial liver segment (LEoLMLS), Rouviere's sulcus (RS) [27, 32–55]. Some authors distinguish the anatomical identification into safe and dangerous zone for dissection, called "Go" and "No-Go" zone, respectively [36, 37, 45, 54]. The "Go" zone was defined as the area located within the HCT (closer to the inferior edge of the gallbladder) that is deemed safe to proceed with dissection with a low probability of causing a major BDI. The "No-Go" zone was defined as the deeper region within the HCT, where further dissection was deemed unnecessary and dangerous with an unacceptable probability of causing a major BDI. The "No-Go" zone also includes the hepatoduodenal ligament, liver hilum, and all structures inferiorly [27, 35]. Non-segmented regions of dissection are deemed as "in-between" (i.e., neither Go, or No-Go) zone [45].

As CVS is based on the proper identification of anatomical structures, studies involving CVS assessment were included in this systematic review [33, 27, 35, 36, 39, 41, 44, 45, 47, 48, 50–52]. The achievement of CVS is based on three criteria: the 2-structure criterion (CA and CD

entering the gallbladder), the HCT criterion (a carefully dissected HCT), and the CP criterion (the division of the lower third of the gallbladder from the CP) [27].

Anatomical identification and CVS assessment by AI could be achieved in several ways [32–34, 27, 35–54]. Semantic segmentation [36, 37, 45, 54], computer vision model for workflow analysis [27, 35, 36], Deep Convolutional Neural Network (DCNN) and Graph neural networks (GNNs) [33, 41, 44, 45], bounding boxes [32, 34, 39, 41, 42, 48] and MIL models [38, 51] are some examples.

The AI model performance was compared to surgical annotations [32, 27, 33–55]. It was evaluated in terms of mean average precision (mAP), mean or balanced accuracy (b-ACC), Area Under the Curve (AUC), sensitivity or recall, specificity, negative predictive value (NPV), positive predictive value (PPV), detection rate (DR) and positive ratios with PMLA (partial matches between landmark detection and annotation) [32–34, 27, 35–55]. mAP evaluates the overall performance of the model, defined as the average precision across all recall values between 0 and 1, while bACC is the average of sensitivity and specificity for the best model selected using the validation set [36]. AUC is an effective metric to summarize the overall accuracy of the test. It takes values from 0 to 1, where a value of 0 indicates a perfectly inaccurate test and a value of 1 reflects a perfectly accurate test [56]. In general, an AUC from 0.7 to 0.8 is considered acceptable, 0.8 to 0.9 is considered excellent, and more than 0.9 is considered outstanding [51]. The performance of the segmentation network was measured with mean intersection over union (IOU) and Dice/F1 score, which are two commonly used metrics in computer vision to quantify the percentage of overlap between the segmentation output and the “ground truth” [37]. The IOU value will be closer to 1 as much as the segmentation is correct (and therefore the greater the overlap area), while the F1 score is calculated as the harmonic mean of the precision and recall scores. It ranges from 0–100%, and a higher F1 score denotes a better-quality classifier. [35, 36, 40, 43]. In case of CVS achievement, the validity of the AI annotations of CVS component (AI-surgeon inter-rater agreement) was assessed with Cohen’s kappa, percentage of agreement and Gwet’s AC2, a weighted measure that adjusts for scales with high probability of random agreement and represented by a collapsed dichotomous scoring rubric of “poor—unsafe” vs. “adequate/excellent—safe”. Gwet’s AC2 metric ranges from 0 to 100%: excellent (81% to 100%), very good (61% to 80%), moderate (41% to 60%), and poor (<40%). [33, 27, 47, 50, 57].

Risk of bias assessment in the included studies

Risk of bias of the included studies was assessed three authors (D.C., D.C., A.B.) using the revised Cochrane

risk-of-bias tool for randomized trials (RoB 2) and the risk of bias in non-randomized studies for interventions (ROB-INS-I) tool [58, 59]

Grading the quality of evidence

According to the Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) approach, two authors (D.C., A.B.) independently evaluated the quality of evidence for imprecision, inconsistency, indirectness, and publication bias [60]. Moreover, the quality of evidence was classified as very low, low, moderate, or high [60]. Subsequently, a summary table was created using the GRADE profiler software (version 3.6.1) (available at: <http://www.gradeworkinggroup.org/>) [60].

Results

The search revealed 2304 articles, of these, 1164 were eliminated because they were duplicates. Of the remaining 1139 articles, 935 were excluded after screening title and abstract because they did not meet the inclusion criteria. Fifty-seven articles were thoroughly analysed, and 33 further articles were excluded. Finally, twenty-five articles published between June 2020 and September 2024 were included [27, 32–43], as shown in the PRISMA flow diagram (Fig. 1) [29].

Table 1 reports studies characteristics and results, while the risk of bias based on ROBINS-I and RoB 2 of the included articles is reported in Tables 2 and 3. The assessment of evidence according to the GRADE method of the included articles is reported in Supplementary Material-Table 2.

Discussion

The present study aims to report current evidence about using AI to identify anatomical structures during LC.

Even if the evaluation was heterogeneous among the included studies, AI models can be used to identify anatomical structures in the surgical field during LC, increasing surgical team awareness and playing a potential role in preventing BDI. [27, 32–55].

As reported and as expected, the AI appears more effective in identifying areas of dissection, rather than individual structures (lower mAP value for CBD, CD, RS, and LEO-MLS, compared to Go and No-Go zone and CVS) [32, 37, 44]. Conversely, bACC values were also reported to be high for individual structures, reaching 3.8/4 (95%) in CBD recognition. Likewise, AI’s performance is reported grater for No-Go zone compared to Go zones [37, 38, 55].

These results can be explained because identifying dissection areas is certainly easier than a single anatomical

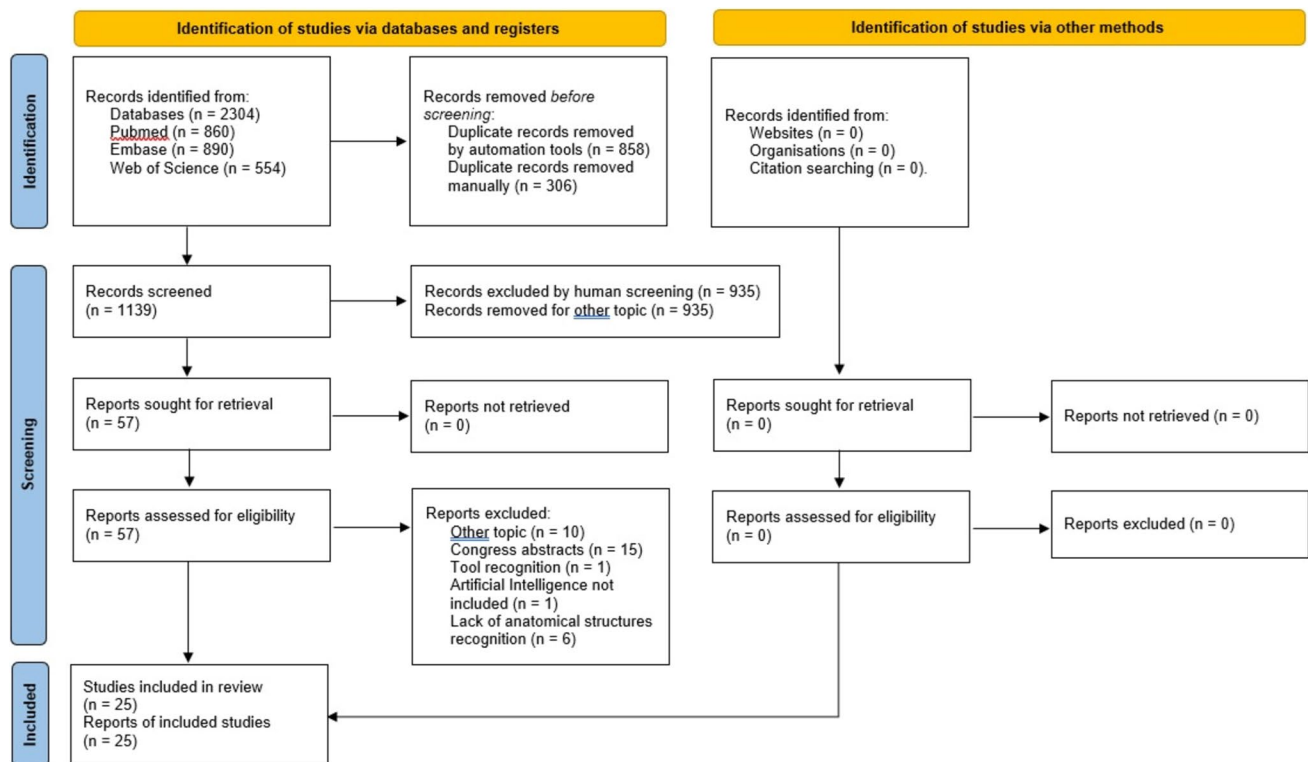


Fig. 1 Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) flow diagram [29]

structure. Equally, the volume of a No-Go zone within a frame is greater and therefore more easily identifiable compared to a Go zone [32–34, 27, 35–55].

It is reported in the literature that almost all BDIs are caused by errors in human visual perception and that technology is probably the most important weapon against these complications [61, 62]. AI models aim to create an accurate mental model that reflects the proper surgical anatomy, thanks to the segmentation of images collected from surgical videos [37–66].

While the agreement between AI and surgeons in identifying CVS reported in the included papers is high overall [33, 27, 35, 36, 50], it decreases significantly (to 56%) in cases of low-severity disease. Moreover, Korndorffer et al. reported considerably higher agreement ($p < 0.001$) in more complex cases compared to simpler ones. This finding suggests that the nuances of the surgeon's decision-making process, especially in individual cases, may be difficult to capture within a single AI model [33].

The results reported in this systematic review demonstrate that the use of AI can lead to a shift in surgical strategy, resulting in safer dissections [43]. Furthermore, Leifman et al. show that intraoperative AI use results in an increased rate of achieving CVS (85% versus 19% in their standard LC and versus the 7–52% range reported in the literature) [41]. This result could be partially attributed to the well-known

"Hawthorne effect", where the awareness of being monitored by AI potentially incentivized surgeons to optimize their performance [51].

It is crucial to differentiate between papers that use AI for post-hoc analysis of collected videos [33, 27, 35–37, 39, 41, 42, 45–48] and those that, departing from academic use, apply AI in clinical practice [32, 34, 38, 40, 43, 44, 49–55] and it is noteworthy that recent studies employ more complex technologies built upon extensive databases [49–55].

We strongly believe in the potential educational role of using AI during LC. As we know, and as reported in the literature the advice from a member of the surgical team other than the operator is one of the most adopted ways to prevent BDI and the risk of BDI appears lower in hospital with a surgical residency program and in high-volume centres, underlining the importance of the expert surgeon [67, 68]. We do not believe that AI can replace the opinion of an expert surgeon during a complex LC, but we assess that it could represent a valid tool especially for younger surgeons during MIS, since there is a lack of tactile sensitivity and visual information becomes predominant.

All the literature included in this review was published after 2020, highlighting the actuality and novelty of this paper, furthermore, although other systematic reviews on the role of AI in surgery exist in the literature [69–78], to our knowledge, this is the first to address its use in identifying

Table 1 Summary of the 25 included studies

Author (year), [reference]	Aim	Dataset (n of videos)	Exclusion criteria	Anatomical structures identified	AI model	Ground truth establishment/annotations	Outcome
Tokuyasu et al. [32]	identification of individual anatomical structures	99	high inflammation and fibrosis, bleeding, or less-visible landmarks	CD, CBD, LEoLMs, RS	YOLOv3	two expert surgeons (> 200 LC)	mAP for CBD: 32%; CD: 7%; LEoLMs: 3%; RS: 10%
Korndorffer et al. [33]	CVS assessment	385	None	HCT, CP, CD, CD + CA	ConvNet; 3D DCNN	Eight board-certified surgeons	% of agreement: HCT 81%, cystic plate 84%, CD + CA 81%, CVS 93% Kappa: HCT 0.6, cystic plate 0.6, CD + CA 0.4, CVS 0.6
Inomata et al. [34]	identification of individual anatomical structures	99	high inflammation and fibrosis, bleeding, or less-visible landmarks	CD, CBD, LEoLMs, RS	YOLOv3	two expert surgeons (> 200 LC)	High DR of landmarks in case of mild to moderate inflammation
Mascagni et al. [27]	CVS assessment	155	Incomplete recordings, bailout procedures, IOC	HCT, CP, CD, CD + CA	EndoDigest	two independent HPB surgeons	Average error for cystic duct division identification: 1.9%, CVS assessment: 91% Inter-rater agreement for CVS 85%, AC + CD 91%, HCT 82%, CP 84%
Mascagni et al. [35]	CVS assessment	144	Incomplete recordings, bailout procedures, IOC	HCT, CP, CD, CD + CA	EndoDigest	two independent HPB surgeons	75% of agreement for CVS achievement
Mascagni et al. [36]	CVS assessment	201	abnormal hepatocystic anatomy and absence of anterior view of CVS	Gallbladder, CD, CA, CP, HCT	DeepCVS	one surgeon in training and one HPB surgeon, in case of disagreement a senior HPB surgeon mediated	Inter-rater agreement for CVS 82%; CA + CD 93%; HCT 68%; CP 74% CVS mAP: 72%; CVS bACC: 71%; Se: 70%; Sp: 72%; Mean IOU: 0.7

Table 1 (continued)

Author (year), [reference]	Aim	Dataset (n of videos)	Exclusion criteria	Anatomical structures identified	AI model	Ground truth establishment/annotations	Outcome
Madani et al. [37]	identification of individual anatomical structures and safe/dangerous zones of dissection	308	subtotal and top-down approach LC	Go zone, No-Go zone and target anatomy (gallbladder, liver, HCT)	ResNet50 (GoNoGoNet) CholeNet	three acute care and MIS surgeon and one high-volume HPB surgeon	Go zone: ACC 94%; Se 69%; Sp 94%; NPV 96%; PPV 74%; IOU 0.5; F1 score 70% No-Go zone: ACC 95%; Se 80%; Sp 98%; NPV 97%; PPV 86%; IOU 0.7; F1 score 83% HCT: ACC 93%; Se 80%; Sp 95%; NPV 96%; PPV 78%; IOU 0.7; F1 score 79%
Laplante et al. [38]	identification of safe/dangerous zones of dissection in comparison to expert surgeons	25	top-down approach LC, and poor visualization of the hepatocystic triangle	Go zone, No-Go zone	GoNoGoNet	six high-volume expert surgeons from the SAGES Safe Cholecystectomy Task Force	Go zone: ACC 92%; Se 52%; Sp 97%; NPV 70%; PPV 94%; IOU > 0.4; F1 score 58% No-Go zone: ACC 92%; Se 80%; Sp 95%; NPV 95%; PPV 84%; IOU > 0.7; F1 score 80%
Colbeci et al. [39]	CVS assessment	556	None	HCT, CP, CD, CD + CA	MIL model	A group of surgeons	ACC: HCT 83%, cystic plate 83%, CA + CD 82%, mean ACC 83% CVS: ACC 84%, Sp 87%, Se 72%, F1 score 66% F1 score for CBD 72%; CD 49%; S4 46%; RS 66%; bACC 4.2/5 bACC: CVS 67%, CA 55%, CD 55%, liver 60%, CP 56%, CA + CD 54%
Nakanuma et al. [40]	identification of individual anatomical structures	10	high inflammation and fibrosis, bleeding, or less-visible landmarks	CD, CBD, LEoLMLS, RS	YOLOv3	three expert surgeons (not belonging to the research group)	
Ban et al. (2023) [41]	CVS assessment	100	None	HCT, CP, CD, liver, CD + CA	ConceptNet-Resnet50; ConceptNet-ViT	clinical CVS experts	

Table 1 (continued)

Author (year), [reference]	Aim	Dataset (n of videos)	Exclusion criteria	Anatomical structures identified	AI model	Ground truth establishment/annotations	Outcome
Fujinaga et al. [42]	identification of individual anatomical structures and prevention of BDIs	30	conversion to open surgery	CD, CBD, LEoLMs, RS	Cross-AI (YOLOv3 + EfficientNet-B5 and a SAM optimizer)	two expert surgeons; IEC: three surgeons involved in the study; EEC three certified surgeons from multiple institutions	F1 score for CBD 45%, CD 20%, S4 27%, RS 30%; PMLA for CBD 84%, CD 77%, S4 87%, RS 70%; ACC for CBD and CD 3.8/4 and 3.7/4, respectively. The contribution to preventing BDI was at 3.7/4
Endo et al. [43]	evaluation of the influence of the AI system on the identification of individual anatomical structures	95	severe inflammation and abnormal biliary anatomy	RS, LEoLMs, CD and EHBD	YOLOv3	four beginners and four expert surgeons	27% of changed annotations (70% safer changes)
Kawamura et al. [44]	CVS assessment	72	severe cholecystitis	CD, CA, CP, HT	EfficientNet-B5 and a SAM optimizer	a board-certified surgeon in gastroenterology A JSES board certified surgeon performed the final confirmation of the annotation data	mAP:97%; Se:74%; F1 score:83%; Sp:97%; ACC:83%
Alkhamaisch et al. [45]	CVS assessment	200	poor image quality (bleeding, less visible landmarks, anatomical differences) and no-standard LC	HCT, CP, CD, CD + CA	Auto-Encoder; U-Net	two expert surgeons and three 5th-year resident	ACC 92%; Mean IOU: 0.75
Khalid et al. [46]	identification of safe/dangerous zones of dissection in LC with BDIs	22	poor visual quality	Go zone, No-Go zone, or neither (i.e., in between the two)	U-Net	two independent high-volume surgeons (members of the SAGES Safe Cholecystectomy Task Force)	LC videos with BDIs had a significantly greater proportion of dissection occurring outside of the Go zone compared to LC videos in the Control group ($p < 0.01$)
Adrales et al. [47]	CVS assessment	36	None	HCT, CP, CA, CD	LC-CVS OPSA	12 expert LC surgeon	Inter-rater reliability: Gwet's AC2: HCT 41%; CP 50%; CA 54%; CD 72%. Total agreement: HCT 76%; CP 78%; CA 80%; CD 86%

Table 1 (continued)

Author (year), [reference]	Aim	Dataset (n of videos)	Exclusion criteria	Anatomical structures identified	AI model	Ground truth establishment/annotations	Outcome
Hedge et al. [48]	identification of individual anatomical structures and CVS assessment	160	None	HCT, CP, CD, CD + CA	3D-CNN	three expert surgeons	CVS: precision 83%; Se 85%; F1 score:84% CD + CA: precision 77%; Se 65%; F1 score:71% LEoLMLS + RS: pr 92%; Se:87%, mAP 91% IOU 0.5. Mean surgeons' acceptances of the guided line dissection: 96%
Smithmaitrie et al. [49]	identification of individual anatomical structures and visualization of guided dissection line	40	None	LEoLMLS, RS	YOLOv7	three independent surgeons	100% achievement for CVS achievement
Petracchi et al. [50]	CVS assessment	40	Non-elective LC performed for reasons different from symptomatic gallstones and Tokyo 1 cholecystitis	CD, CA and CVS	PyCharm, Google Colab Pro and YOLOv8	three blinded surgeons	100% achievement for CVS achievement
Leifman et al. [51]	CVS assessment	40	None	CVS: HCT, CP, CD, CA, liver	SpineNet-49, Resnet50	three independent surgeons	Sp 100%; Se 97%; CVS AUC 0.9 (CD and CA: 0.89, HCT: 0.96, CP: 0.88, liver: 0.9)
Fried et al. [52]	CVS assessment	279	None	CVS: HCT, CP, CD, CA	MIL model	two expert surgeons	ACC 95%; CA and CD visibility rise from 42 to 75% ($p < 0.001$) thanks to AI-CVS achievement
Tashiro et al. [53]	LCT with AI and anatomical structures with ICG	1	LC without ICG	LCT, CD, gallbladder, liver	Eureka	-	LCT identification: IOU 0.6; F1 score:60%. The combination of AI and ICG may reduce BDI rate
Wu et al. [54]	CVS assessment and improving in surgical performance	90	Other concomitant procedures, patients with severe comorbidities, suspect of gallbladder malignancy and Parkland score ≥ 3	HCT, CP, CD, CD + CA	SmartCoach	hepatobiliary experts	Significant improvement in CVS achievement in AI-enhanced feedback group as compared to self-learning group (11% to 78%, $p = 0.02$)

Table 1 (continued)

Author (year), [reference]	Aim	Dataset (n of videos)	Exclusion criteria	Anatomical structures identified	AI model	Ground truth establishment/annotations	Outcome
Protserov et al. (2024) [55]	identification of safe/dangerous zones of dissection	314	subtotal and top-down approach LC	Go zone, No-Go zone	U-Net and SegFormer	A panel of expert high-volume surgeons for Go and No-Go zone, separately	U-Net: Go-zone: F1 score 57%; precision 48%; Se 82%; RAE + 92% No-Go-zone: F1 score 76%; precision 68%; Se 92%; RAE + 47% SegFormer: Go-zone: F1 score 60%; precision 53%; Se 75%; RAE + 48% No-Go-zone: F1 score 76%; precision 68%; Se 92%; RAE + 46%

AI Artificial intelligence; BDI Bile duct injury; CD Cystic duct; CA Cystic artery; CBD Common bile duct; LEO LMLS Lower edge of the left medial liver segment; RS Rouviere's sulcus; HCT hepatocystic triangle; YOLOv3 You only look once version 3; LC Laparoscopic cholecystectomy; mAP Mean average precision; DR Detection rate; CVS Critical view of Strasberg; CP Cystic plate; HPB Hepato-pancreato-biliary; DCNN Deep Convolutional Neural Network; IOU Intersection over union; bACC Balanced accuracy; ACC Accuracy; Se Sensitivity; Sp Specificity; MIS Minimally invasive surgery; NPV Negative predictive value; IOC Intraoperative Cholangiography; SAGES Society of American Gastrointestinal and Endoscopic Surgeons; MIL Multi Instance Learning; EEC External evaluation committee; EHBD Extrahepatic bile duct; S4 Segment 4; SAM Sharpness-Aware Minimization; IEC Internal evaluation committee; JSES Japanese Society for Endoscopic Surgery; PMLA Partial matches between landmark detection and annotation; AUC Area Under the Curve; CNN Convolutional Neural Network; LCT Loose connective tissue; ICG Indocyanine green; RAE Relative-Area-Error

Table 2 Assessment of risk of bias of the included articles based on Risk of Bias in Non-randomised Studies – of Interventions (ROBINS-I) [58]







Author, year, type of study	Bias due to confounding	Bias in selection participants	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of reported result	Overall
Tokuyasu et al. 2020, ambispective [32]	Moderate	Moderate	Low	Moderate	Moderate	Moderate	Moderate	Moderate
Korndorffer et al. 2020, retrospective [33]	Low	Low	Low	Low	Low	Moderate	Low	Low
Inomata et al. 2021, ambispective [34]	Low	Moderate	Low	Low	Moderate	Moderate	Moderate	Moderate
Mascagni et al. 2021, retrospective [27]	Low	Moderate	Low	Low	Low	Moderate	Low	Low
Mascagni et al. 2022, retrospective [35]	Low	Moderate	Low	Low	Low	Moderate	Low	Low
Mascagni et al. 2022, retrospective [36]	Low	Moderate	Moderate	Moderate	Low	Low	Low	Low
Madani et al. 2022, retrospective [37]	Low	Low	Low	Moderate	Low	Moderate	Moderate	Low
Laplante et al. 2022, prospective [38]	Low	Low	Low	Moderate	Low	Moderate	Moderate	Low
Colbeci et al. 2022, retrospective [39]	Low	Low	Low	Low	Low	Moderate	Low	Low
Nakanuma et al. 2022, prospective [40]	Low	Low	Low	Low	Low	Moderate	Low	Low
Ban et al. 2023, retrospective [41]	Low	Moderate	Moderate	Moderate	Low	Low	Low	Moderate
Endo et al. 2022, retrospective [42]	Low	Low	Low	Low	Low	Moderate	Low	Low
Fujinaga et al. 2022, prospective [43]	Low	Low	Low	Low	Low	Moderate	Low	Low
Kawamura et al. 2022, prospective [44]	Low	Low	Low	Low	Low	Moderate	Low	Low
Alkhamaiseh et al. 2023, retrospective [45]	Low	Low	Low	Low	Low	Moderate	Low	Low
Khalid et al. 2022, retrospective [46]	Low	Low	Low	Low	Low	Moderate	Low	Low
Adrales et al. 2023, retrospective [47]	Low	Moderate	Low	Low	Low	Moderate	Low	Low

Table 2 (continued)

Author, year, type of study	Bias due to confounding	Bias in selection participants	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of reported result	Overall
Hedge et al. 2023, retrospective [48]	Low	Low	Low	Low	Low	Moderate	Low	Low
Smithmaitrie et al. 2024, ambispective [49]	Low	Moderate	Low	Low	Low	Moderate	Low	Low
Petracchi et al. 2024, prospective [50]	Low	Low	Low	Low	Low	Moderate	Low	Low
Leifman et al. 2024, prospective [51]	Low	Moderate	Low	Low	Low	Moderate	Low	Low
Fried et al. 2024, ambispective [52]	Low	Moderate	Low	Low	Low	Moderate	Low	Low
Tashiro et al. 2024, ambispective [53]	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate
Protserov et al. 2024, ambispective [55]	Low	Moderate	Low	Low	Low	Moderate	Low	Low

Low: low risk of bias (the study is comparable to a randomised trial). Moderate: moderate risk of bias (the study provides sound evidence for a nonrandomised study but cannot be considered comparable to a randomised trial). Serious: serious risk of bias (the study has important problems)

Table 3 Assessment of risk of bias of the included articles based on Risk of Bias assessment-RoB-II [59]

Author, year, type of study	Bias due to randomisation process	Bias due to deviations from intended interventions	Bias due to missing outcome data	Bias in measurement of outcomes	Bias in selection of reported result	Overall
Wu <i>et al.</i> , 2024, prospective [55]						



anatomical structures during LC. This aspect represents an advantage and a limitation of this paper. Indeed, this is certainly a current topic, but there is active work in this field, with continuous publications, which has made it difficult to conduct a systematic review.

The main limitations of the present study are the overall poor quality and quantity of included studies and the presence of only one randomized control trials. The heterogeneity of the AI models employed, the heterogeneity of the inclusion criteria and the small number of cases analysed in

each study represent further limitations. Expert opinion is also very heterogeneous, potentially limiting the generalizability of such models and the data was extracted from the articles themselves without access to the raw data, and this did not allow for true statistical analysis. Another important limitation is that the risk of BDI increases in some conditions such as high inflammation and fibrosis, bleeding, anatomical abnormalities, which also represent the exclusion criteria in many of the included papers [32, 34, 36, 38, 40, 43, 44, 48, 50, 54].

Finally, a further non-negligible limitation is that all the included studies coming from very few research groups with small number and possible overlapping dataset.

Conclusions

Although current studies on using AI to identify anatomical structures during LC are few and with considerable heterogeneity, AI models seem to be valuable tools to anatomical identification, but we are still far from real-time guidance to minimize the risk of BDIs and others adverse events. The results seem promising, but new randomized trials and greater standardization are needed for these AI models to be reproducible.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00423-025-03651-6>.

Authors' contributions D.C, A.B, D.C: Study conception and design, Acquisition of data, Analysis and interpretation of data, Drafting of manuscript, Critical revision of manuscript. Final approval. D.P, F.R, S.M, P.S, L.A: Study conception and design, Critical revision of manuscript. Final approval. M.P, A.P, M.O: Critical revision of manuscript. Final approval.

Funding Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement. There are no grant support or financial relationship for the present study.

Data availability No datasets were generated or analysed during the current study.

Declarations

Ethical approval All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. For this type of study formal consent is not required.

Institutional review board approval and informed consent from participants were unnecessary for the present study. This study was registered in the International Prospective Register of Systematic Reviews (PROSPERO; CRD-42023478754).

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Fugazzola P, Cobianchi L, Di Martino M, Tomasoni M, Dal Mas F, Abu-Zidan FM, Agnoletti V, Ceresoli M, Coccolini F, Di Saverio S, Dominioni T, Farè CN, Frassini S, Gambini G, Leppäniemi A, Maestri M, Martín-Pérez E, Moore EE, Musella V, Peitzman AB, de la Hoz RÁ, Sargenti B, Sartelli M, Viganò J, Anderloni A, Biffl W, Catena F, Ansaloni L, S.P.Ri.M.A.C.C. Collaborative Group (2023) Prediction of morbidity and mortality after early cholecystectomy for acute calculous cholecystitis: results of the SPReMACC study. *World J Emerg Surg* 18(1):20. <https://doi.org/10.1186/s13017-023-00488-6>
2. Ortenzi M, Agresta F, Vettoretto N, Gerardi C, Allocati E, Botteri E, Montori G, Balla A, Arezzo A, Piatto G, Sartori A, Antoniou S, Podda M (2023) Use of High Energy Devices (HEDs) versus electrocautery for laparoscopic cholecystectomy: a systematic review and meta-analysis of randomised controlled trials. *Surg Endosc* 37(6):4249–4269. <https://doi.org/10.1007/s00464-023-10060-7>
3. Quaresima S, Balla A, Guerrieri M, Lezoche G, Campagnacci R, D'Ambrosio G, Lezoche E, Paganini AM (2023) Results of medium seventeen years' follow-up after laparoscopic choledochotomy for ductal stones. *Gastroenterol Res Pract*. <https://doi.org/10.1155/2016/9506406>
4. Quaresima S, Balla A, Guerrieri M, Campagnacci R, Lezoche E, Paganini AM (2017) A 23-year experience with laparoscopic common bile duct exploration. *HPB (Oxford)* 19(1):29–35. <https://doi.org/10.1016/j.hpb.2016.10.011>
5. Balla A, Quaresima S, Corona M, Lucatelli P, Fiocca F, Rossi M, Bezzi M, Catalano C, Salvatori FM, Fingerhut A, Paganini AM (2019) ATOM classification of bile duct injuries during laparoscopic cholecystectomy: analysis of a single institution experience. *J Laparoendosc Adv Surg Tech A* 29(2):206–212. <https://doi.org/10.1089/lap.2018.0413>
6. Guerra F, Coletta D, Gavioli M, Coco D, Patriti A (2020) Minimally invasive surgery for the management of major bile duct injury due to cholecystectomy. *J Hepatobiliary Pancreat Sci* 27(4):157–163. <https://doi.org/10.1002/jhbp.710>
7. de'Angelis N, Catena F, Memeo R, Coccolini F, Martínez-Pérez A, Romeo OM, De Simone B, Di Saverio S, Brustia R, Rhaïem R, Piardi T, Conticchio M, Marchegiani F, Beghdadi N, Abu-Zidan FM, Alikhanov R, Allard MA, Allievi N, Amaddeo G, Ansaloni L, Andersson R, Andolfi E, Azfar M, Bala M, Benkabbou A, Ben-Ishay O, Bianchi G, Biffl WL, Brunetti F, Carra MC, Casanova D, Celentano V, Ceresoli M, Chiara O, Cimbanassi S, Bini R, Coimbra R, de'Angelis LG, Decembrino F, De Palma A, de Reuver PR, Domingo C, Cotsoglou C, Ferrero A, Fraga GP, Gaiani F, Gheza F, Gurrado A, Harrison E, Henriquez A, Hofmeyr S, Iadarola R, Kashuk JL, Kianmanesh R, Kirkpatrick AW, Kluger Y, Landi F, Langella S, Lapointe R, Le Roy B, Luciani A, Machado F, Maggi U, Maier RV, Mefire AC, Hiramatsu K, Ordoñez C, Patrizi F, Planells M, Peitzman AB, Pekolj J, Perdigao F, Pereira BM, Pessaux P, Pisano M, Puyana JC, Rizoli S, Portigliotti L, Romito R, Sakakushev B, Sanei B, Scatton O, Serradilla-Martin M, Schneck AS, Sissoko ML, Sobhani I, Ten Broek RP, Testini M, Valinas R, Veloudis G, Vitali GC, Weber D, Zorcolo L, Giuliani F, Gavriliadis P, Fuks D, Sommacale D (2021) 2020 WSES guidelines for the detection and management of bile duct injury during cholecystectomy. *World J Emerg Surg* 16(1):30. <https://doi.org/10.1186/s13017-021-00369-w>
8. Morales-Conde S, Licardie E, Alarcón I, Balla A (2022) Indocyanine green (ICG) fluorescence guide for the use and indications in general surgery: recommendations based on the descriptive review of the literature and the analysis of experience. *Cir Esp (Engl Ed)* 100(9):534–554. <https://doi.org/10.1016/j.cireng.2022.06.023>

9. Agnus V, Pesce A, Boni L, Van Den Bos J, Morales-Conde S, Paganini AM, Quaresima S, Balla A, La Greca G, Plaudis H, Moretto G, Castagnola M, Santi C, Casali L, Tartamella L, Saadi A, Picchetto A, Arezzo A, Marescaux J, Diana M (2020) Fluorescence-based cholangiography: preliminary results from the IHU-IRCAD-EAES EURO-FIGS registry. *Surg Endosc* 34(9):3888–3896. <https://doi.org/10.1007/s00464-019-07157-3>
10. Morales-Conde S, Licardie E, Alarcón I, Balla A (2020) Indocyanine-green-guided, ureteric preserving, laparoscopic Hartmann's procedure for obstructing colonic adenocarcinoma with endometriosis - a video vignette. *Colorectal Dis* 22(11):1764–1765. <https://doi.org/10.1111/codi.15149>
11. Morales-Conde S, Alarcón I, Yang T, Licardie E, Camacho V, Aguilar Del Castillo F, Balla A (2020) Fluorescence angiography with indocyanine green (ICG) to evaluate anastomosis in colorectal surgery: where does it have more value? *Surg Endosc* 34(9):3897–3907. <https://doi.org/10.1007/s00464-019-07159-1>
12. Morales-Conde S, Alarcón I, Yang T, Licardie E, Balla A (2020) A decalogue to avoid routine ileostomy in selected patients with border line risk to develop anastomotic leakage after minimally invasive low-anterior resection: a pilot study. *Surg Innov* 27(1):44–53. <https://doi.org/10.1177/1553350619890720>
13. Morales-Conde S, Balla A, Alarcón I, Licardie E (2021) Management of postoperative complications after laparoscopic left hemicolectomy: an approach in modern times after incorporation of indocyanine green and full mobilization of the splenic flexure. *Minerva Surg* 76(4):303–309. <https://doi.org/10.23736/S2724-5691.21.08642-9>
14. Balla A, Sosa V, Licardie E, Alarcón I, Morales-Conde S (2019) Laparoscopic left hemicolectomy with indocyanine green fluorescence angiography for diverticular disease in a patient with intestinal malrotation - a video vignette. *Colorectal Dis* 21(8):978–979. <https://doi.org/10.1111/codi.14670>
15. Spota A, Al-Taher M, Felli E, Morales Conde S, Dal Dosso I, Moretto G, Spinoglio G, Baiocchi G, Vilallonga R, Impellizzeri H, Martin-Martin GP, Casali L, Franzini C, Silvestri M, de Manzini N, Castagnola M, Filauro M, Cosola D, Copaescu C, Garbarino GM, Pesce A, Calabrò M, de Nardi P, Anania G, Carus T, Boni L, Patané A, Santi C, Saadi A, Rollo A, Chautems R, Noguera J, Grosek J, D'Ambrosio G, Ferreira CM, Norcic G, Navarra G, Riva P, Quaresima S, Paganini A, Rosso N, De Paolis P, Balla A, Sauvain MO, Gialamas E, Bianchi G, La Greca G, Castoro C, Picchetto A, Franchello A, Tartamella L, Juvan R, Ioannidis O, Kosir JA, Bertani E, Stassen L, Marescaux J, Diana M (2021) Fluorescence-based bowel anastomosis perfusion evaluation: results from the IHU-IRCAD-EAES EURO-FIGS registry. *Surg Endosc* 35(12):7142–7153. <https://doi.org/10.1007/s00464-020-08234-8>
16. Guerra F, Coletta D, Greco PA, Eugeni E, Patriti A (2021) The use of indocyanine green fluorescence to define bowel microcirculation during laparoscopic surgery for acute small bowel obstruction. *Colorectal Dis* 23(8):2189–2194. <https://doi.org/10.1111/codi.15680>
17. Balla A, Corallino D, Quaresima S, Palmieri L, Meoli F, Cordova Herencia I, Paganini AM (2022) Indocyanine green fluorescence angiography during laparoscopic bariatric surgery: a pilot study. *Front Surg* 26(9):906133. <https://doi.org/10.3389/fsurg.2022.906133>
18. Picchetto A, Cinelli L, Bannone E, Baiocchi GL, Morales-Conde S, Casali L, Spinoglio G, Franzini C, Santi C, D'Ambrosio G, Copaescu C, Rollo A, Balla A, Lepiane P, Paganini AM, Detullio P, Quaresima S, Pesce A, Luciano T, Bianchi G, Marescaux J, Diana M (2023) Fluorescence-based sentinel lymph node mapping and lymphography evaluation: results from the IHU-IRCAD-EAES EURO-FIGS registry. *Surg Endosc* 37(7):5472–5481. <https://doi.org/10.1007/s00464-023-10043-8>
19. Aldrighetti L, Cipriani F, Fornoni G, Catena M, Ratti F (2023) Robo-Lap approach for anatomical resections of postero-superior liver segments by indocyanine green fluorescence: intraoperative navigation based on vascular landmarks. *Chirurgia (Bucur)* 118(2):170–179. <https://doi.org/10.21614/chirurgia.2816>
20. Hiwatashi K, Okumura H, Setoyama T, Ando K, Ogura Y, Aridome K, Maenohara S, Natsugoe S (2018) Evaluation of laparoscopic cholecystectomy using indocyanine green cholangiography including cholecystitis: A retrospective study. *Medicine (Baltimore)* 97(30):e11654. <https://doi.org/10.1097/MD.00000000000011654>
21. Kimura Y, Takada T, Kawarada Y, Nimura Y, Hirata K, Sekimoto M, Yoshida M, Mayumi T, Wada K, Miura F, Yasuda H, Yamashita Y, Nagino M, Hirota M, Tanaka A, Tsuyuguchi T, Strasberg SM, Gadacz TR (2007) Definitions, pathophysiology, and epidemiology of acute cholangitis and cholecystitis: Tokyo Guidelines. *J Hepatobiliary Pancreat Surg* 14(1):15–26. <https://doi.org/10.1007/s00534-006-1152-y>
22. Gallaher JR, Charles A (2022) Acute cholecystitis: a review. *JAMA* 327(10):965–975. <https://doi.org/10.1001/jama.2022.2350>
23. Pisano M, Allievi N, Gurusamy K, Borzellino G, Cimbanassi S, Boerna D, Coccolini F, Tufo A, Di Martino M, Leung J, Sartelli M, Ceresoli M, Maier RV, Poiasina E, De Angelis N, Magnone S, Fugazzola P, Paolillo C, Coimbra R, Di Saverio S, De Simone B, Weber DG, Sakakushev BE, Lucianetti A, Kirkpatrick AW, Fraga GP, Wani I, Biffl WL, Chiara O, Abu-Zidan F, Moore EE, Lepäniemi A, Kluger Y, Catena F, Ansaloni L (2020) World Society of Emergency Surgery updated guidelines for the diagnosis and treatment of acute calculus cholecystitis. *World J Emerg Surg* 15(1):61. <https://doi.org/10.1186/s13017-020-00336-x>
24. Yu KH, Beam AL, Kohane IS (2018) Artificial intelligence in healthcare. *Nat Biomed Eng* 2(10):719–731. <https://doi.org/10.1038/s41551-018-0305-z>
25. Cheng K, You J, Wu S, Chen Z, Zhou Z, Guan J, Peng B, Wang X (2022) Artificial intelligence-based automated laparoscopic cholecystectomy surgical phase recognition and analysis. *Surg Endosc* 36(5):3160–3168. <https://doi.org/10.1007/s00464-021-08619-3>
26. Ward TM, Hashimoto DA, Ban Y, Rosman G, Meireles OR (2022) Artificial intelligence prediction of cholecystectomy operative course from automated identification of gallbladder inflammation. *Surg Endosc* 36(9):6832–6840. <https://doi.org/10.1007/s00464-022-09009-z>
27. Mascagni P, Alapatt D, Urade T, Vardazaryan A, Mutter D, Marescaux J, Costamagna G, Dallemagne B, Padoy N (2021) A computer vision platform to automatically locate critical events in surgical videos: documenting safety in laparoscopic cholecystectomy. *Ann Surg* 274(1):e93–e95. <https://doi.org/10.1097/SLA.00000000000004736>
28. Balla A, Sartori A, Botteri E, Podda M, Ortenzi M, Silecchia G, Guerrieri M, Agresta F, ARMIS (Augmented Reality in Minimally Invasive Surgery) Collaborative Group (2023) Augmented reality (AR) in minimally invasive surgery (MIS) training: where are we now in Italy? The Italian Society of Endoscopic Surgery (SICE) ARMIS survey. *Updates Surg* 75(1):85–93. <https://doi.org/10.1007/s13304-022-01383-6>
29. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 29(372):n71. <https://doi.org/10.1136/bmj.n71>
30. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). Cochrane. Available from www.training.cochrane.org/handbook
31. Goossen K, Tenckhoff S, Probst P, Grummich K, Mihaljevic AL, Büchler MW, Diener MK (2018) Optimal literature search for systematic reviews in surgery. *Langenbecks Arch Surg* 403(1):119–129. <https://doi.org/10.1007/s00423-017-1646-x>

32. Tokuyasu T, Iwashita Y, Matsunobu Y, Kamiyama T, Ishikake M, Sakaguchi S, Ebe K, Tada K, Endo Y, Etoh T, Nakashima M, Inomata M (2021) Development of an artificial intelligence system using deep learning to indicate anatomical landmarks during laparoscopic cholecystectomy. *Surg Endosc* 35(4):1651–1658. <https://doi.org/10.1007/s00464-020-07548-x>
33. Korndorffer JR Jr, Hawn MT, Spain DA, Knowlton LM, Azagury DE, Nassar AK, Lau JN, Arnow KD, Trickey AW, Pugh CM (2020) Situating artificial intelligence in surgery: a focus on disease severity. *Ann Surg* 272(3):523–528. <https://doi.org/10.1097/SLA.00000000000004207>
34. Inomata M, Fujinaga A, Nakanuma H, Endo Y, Etoh T, Tokuyasu T, Kitano S (2021) Development of endoscopic surgery navigated by artificial intelligence. *Surg Technol Int* 26(39):99–102. <https://doi.org/10.52198/21.STI.39.CR1432>
35. Mascagni P, Alapatt D, Laracca GG, Guerriero L, Spota A, Fiorillo C, Vardazaryan A, Quero G, Alfieri S, Baldari L, Cassinotti E, Boni L, Cuccurullo D, Costamagna G, Dallemagne B, Padoy N (2022) Multicentric validation of EndoDigest: a computer vision platform for video documentation of the critical view of safety in laparoscopic cholecystectomy. *Surg Endosc* 36(11):8379–8386. <https://doi.org/10.1007/s00464-022-09112-1>
36. Mascagni P, Vardazaryan A, Alapatt D, Urade T, Emre T, Fiorillo C, Pessaux P, Mutter D, Marescaux J, Costamagna G, Dallemagne B, Padoy N (2022) Artificial intelligence for surgical safety: automatic assessment of the critical view of safety in laparoscopic cholecystectomy using deep learning. *Ann Surg* 275(5):955–961. <https://doi.org/10.1097/SLA.00000000000004351>
37. Madani A, Namazi B, Altieri MS, Hashimoto DA, Rivera AM, Pucher PH, Navarrete-Welton A, Sankaranarayanan G, Brunt LM, Okrainec A, Alseidi A (2022) Artificial intelligence for intraoperative guidance: using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. *Ann Surg* 276(2):363–369. <https://doi.org/10.1097/SLA.00000000000004594>
38. Laplante S, Namazi B, Kiani P, Hashimoto DA, Alseidi A, Pasten M, Brunt LM, Gill S, Davis B, Bloom M, Pernar L, Okrainec A, Madani A (2023) Validation of an artificial intelligence platform for the guidance of safe laparoscopic cholecystectomy. *Surg Endosc* 37(3):2260–2268. <https://doi.org/10.1007/s00464-022-09439-9>
39. Colbeci Y, Zohar M, Neimark D, Asselmann D, Bar O (2022) A Multi Instance Learning Approach for Critical View of Safety Detection in Laparoscopic Cholecystectomy. <i>Proceedings of the 7th Machine Learning for Healthcare Conference</i>, in <i>Proceedings of Machine Learning Research</i> 182:409–424 Available from <https://proceedings.mlr.press/v182/colbeci22a.html>
40. Nakanuma H, Endo Y, Fujinaga A, Kawamura M, Kawasaki T, Masuda T, Hirashita T, Etoh T, Shinozuka K, Matsunobu Y, Kamiyama T, Ishikake M, Ebe K, Tokuyasu T, Inomata M (2023) An intraoperative artificial intelligence system identifying anatomical landmarks for laparoscopic cholecystectomy: a prospective clinical feasibility trial (J-SUMMIT-C-01). *Surg Endosc* 37(3):1933–1942. <https://doi.org/10.1007/s00464-022-09678-w>
41. Ban Y, Eckhoff JA, Ward TM, Hashimoto DA, Meireles OR, Rus D, Rosman G (2024) Concept graph neural networks for surgical video understanding. *IEEE Trans Med Imaging* 43(1):264–274. <https://doi.org/10.1109/TMI.2023.3299518>
42. Fujinaga A, Endo Y, Etoh T, Kawamura M, Nakanuma H, Kawasaki T, Masuda T, Hirashita T, Kimura M, Matsunobu Y, Shinozuka K, Tanaka Y, Kamiyama T, Sugita T, Morishima K, Ebe K, Tokuyasu T, Inomata M (2023) Development of a cross-artificial intelligence system for identifying intraoperative anatomical landmarks and surgical phases during laparoscopic cholecystectomy. *Surg Endosc* 37(8):6118–6128. <https://doi.org/10.1007/s00464-023-10097-8>
43. Endo Y, Tokuyasu T, Mori Y, Asai K, Umezawa A, Kawamura M, Fujinaga A, Ejima A, Kimura M, Inomata M (2023) Impact of AI system on recognition for anatomical landmarks related to reducing bile duct injury during laparoscopic cholecystectomy. *Surg Endosc* 37(7):5752–5759. <https://doi.org/10.1007/s00464-023-10224-5>
44. Kawamura M, Endo Y, Fujinaga A, Orimoto H, Amano S, Kawasaki T, Kawano Y, Masuda T, Hirashita T, Kimura M, Ejima A, Matsunobu Y, Shinozuka K, Tokuyasu T, Inomata M (2023) Development of an artificial intelligence system for real-time intraoperative assessment of the critical view of safety in laparoscopic cholecystectomy. *Surg Endosc* 11. <https://doi.org/10.1007/s00464-023-10328-y>
45. Alkhamaiseh KN, Grantner JL, Shebrain S, Abdel-Qader I (2023) Towards reliable hepatocytic anatomy segmentation in laparoscopic cholecystectomy using U-Net with Auto-Encoder. *Surg Endosc* 37(9):7358–7369. <https://doi.org/10.1007/s00464-023-10306-4>
46. Khalid MU, Laplante S, Masino C, Alseidi A, Jayaraman S, Zhang H, Mashouri P, Protserov S, Hunter J, Brudno M, Madani A (2023) Use of artificial intelligence for decision-support to avoid high-risk behaviors during laparoscopic cholecystectomy. *Surg Endosc* 11. <https://doi.org/10.1007/s00464-023-10403-4>
47. Adrales G, Ardito F, Chowbey P, Morales-Conde S, Ferreres AR, Hensman C, Martin D, Matthaehi H, Ramshaw B, Roberts JK, Schrem H, Sharma A, Tabiri S, Vibert E, Woods MS (2024) Laparoscopic cholecystectomy critical view of safety (LC-CVS): a multi-national validation study of an objective, procedure-specific assessment using video-based assessment (VBA). *Surg Endosc* 38(2):922–930. <https://doi.org/10.1007/s00464-023-10479-y>
48. Hegde SR, Namazi B, Iyengar N, Cao S, Desir A, Marques C, Mahnken H, Dumas RP, Sankaranarayanan G (2024) Automated segmentation of phases, steps, and tasks in laparoscopic cholecystectomy using deep learning. *Surg Endosc* 38(1):158–170. <https://doi.org/10.1007/s00464-023-10482-3>
49. Smithmaitrie P, Khaonualsri M, Sae-Lim W, Wangkulangkul P, Jearanai S, Cheewatanakornkul S (2024) Development of deep learning framework for anatomical landmark detection and guided dissection line during laparoscopic cholecystectomy. *Heliyon* 10(3):e25210. <https://doi.org/10.1016/j.heliyon.2024.e25210>
50. Petracchi EJ, Olivieri SE, Varela J, Canullán CM, Zandalazini H, Ocampo C, Quesada BM (2024) Use of artificial intelligence in the detection of the critical view of safety during laparoscopic cholecystectomy. *J Gastrointestinal Surg : official journal of the Society for Surgery of the Alimentary Tract* 28(6):877–879. <https://doi.org/10.1016/j.gassur.2024.03.018>
51. Leifman G, Golany T, Rivlin E, Khoury W, Assalia A, Reissman P (2024) Real-time artificial intelligence validation of critical view of safety in laparoscopic cholecystectomy. *Intell-Based Med* 10. <https://doi.org/10.1016/j.ibmed.2024.100153>
52. Fried GM, Ortenzi M, Dayan D, Nizri E, Mirkin Y, Maril S, Asselmann D, Wolf T (2024) Surgical intelligence can lead to higher adoption of best practices in minimally invasive surgery. *Ann Surg* 280(3):525–534. <https://doi.org/10.1097/SLA.00000000000006377>
53. Tashiro Y, Aoki T, Kobayashi N, Tomioka K, Saito K, Matsuda K, Kusano T (2024) Novel navigation for laparoscopic cholecystectomy fusing artificial intelligence and indocyanine green fluorescent imaging. *J Hepato-Biliary-Pancreatic Sci* 31(5):305–307. <https://doi.org/10.1002/jhbp.1422>
54. Wu S, Tang M, Liu, J., Qin D, Wang Y, Zhai S, Bi E, Li Y, Wang C, Xiong Y, Li G, Gao F, Cai Y, Gao P, Wu Z, Cai H, Liu J, Chen Y, Fang C, Yao L, ... Wang X (2024) Impact of an AI-Based laparoscopic cholecystectomy coaching program on the surgical performance: a randomized controlled trial. *Int J Surg (London, England)* 110(12):7816–7823. <https://doi.org/10.1097/JS9.0000000000001798>
55. Protserov S, Hunter J, Zhang H, Mashouri P, Masino C, Brudno M, Madani A (2024) Development, deployment and scaling of operating room-ready artificial intelligence for real-time surgical

- decision support. *NPJ Digital Med* 7(1):231. <https://doi.org/10.1038/s41746-024-01225-2>
56. Mandrekar JN (2010) Receiver operating characteristic curve in diagnostic test assessment. *J Thoracic Oncol* : Official publication of the International Association for the Study of Lung Cancer 5(9):1315–1316. <https://doi.org/10.1097/JTO.0b013e3181ec173d>
 57. McHugh ML (2012) Interrater reliability: the kappa statistic. *Biochemica medica* 22(3):276–282
 58. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I, Carpenter JR, Chan AW, Churchill R, Deeks JJ, Hróbjartsson A, Kirkham J, Jüni P, Loke YK, Pigott TD, Ramsay CR, Regidor D, Rothstein HR, Sandhu L, Santaguida PL, Schünemann HJ, Shea B, Shrier I, Tugwell P, Turner L, Valentine JC, Waddington H, Waters E, Wells GA, Whiting PF, Higgins JP (2016) ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 12(355):i4919. <https://doi.org/10.1136/bmj.i4919>
 59. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan, A, ... Higgins JPT (2019) RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ (Clinical research ed.)* 366:4898. <https://doi.org/10.1136/bmj.4898>
 60. Schünemann H, Brozek J, Guyatt G, Oxman A (2023) GRADE handbook for grading quality of evidence and strength of recommendation Updated October 2013. [Available from: guidelinedevelopment.org/handbook. Last access on July 27th 2023]
 61. Brunt LM, Deziel DJ, Telem DA, Strasberg SM, Aggarwal R, Asbun H, Bonjer J, McDonald M, Alseidi A, Ujiki M, Riall TS, Hammill C, Moulton CA, Pucher PH, Parks RW, Ansari MT, Connor S, Dirks RC, Anderson B, Altieri MS, Tsamalaidze L, Stefanidis D (2020) Prevention of bile duct injury consensus work group. Safe cholecystectomy multi-society practice guideline and state-of-the-art consensus conference on prevention of bile duct injury during cholecystectomy. *Surg Endosc* 34(7):2827–2855. <https://doi.org/10.1007/s00464-020-07568-7>
 62. Restaino S, Scutiero G, Taliento C, Poli A, Bernardi G, Arcieri M, Santi E, Fanfani F, Chiantera V, Driul L, Scambia G, Greco P, Vizzielli G (2023) Three-dimensional vision versus two-dimensional vision on laparoscopic performance of trainee surgeons: a systematic review and meta-analysis. *Updates Surg* 75(3):455–470. <https://doi.org/10.1007/s13304-023-01465-z>
 63. Conrad C, Wakabayashi G, Asbun HJ, Dallemagne B, Demartines N, Diana M, Fuks D, Giménez ME, Goumarand C, Kaneko H, Memeo R, Resende A, Scatton O, Schneck AS, Soubrane O, Tanabe M, van den Bos J, Weiss H, Yamamoto M, Marescaux J, Pessaux P (2017) IRCAD recommendation on safe laparoscopic cholecystectomy. *J Hepatobiliary Pancreat Sci* 24(11):603–615. <https://doi.org/10.1002/jhbp.491>
 64. Way LW, Stewart L, Gantert W, Liu K, Lee CM, Whang K, Hunter JG (2003) Causes and prevention of laparoscopic bile duct injuries: analysis of 252 cases from a human factors and cognitive psychology perspective. *Ann Surg* 237(4):460–9. <https://doi.org/10.1097/01.SLA.0000060680.92690.E9>
 65. Liu H, Kuang J, Xu Y, Li T, Li P, Huang Z, Zhang S, Weng J, Lai Y, Wu Z, Lin F, Gu W, Huang Y (2023) Investigation of the optimal indocyanine green dose in real-time fluorescent cholangiography during laparoscopic cholecystectomy with an ultra-high-definition 4K fluorescent system: a randomized controlled trial. *Updates Surg* 75(7):1903–1910. <https://doi.org/10.1007/s13304-023-01557-w>
 66. Winkler-Schwartz A, Yilmaz R, Mirchi N, Bissonnette V, Ledwos N, Siyar S, Azarnoush H, Karlik B, Del Maestro R (2019) Machine learning identification of surgical and operative factors associated with surgical expertise in virtual reality simulation. *JAMA Netw Open* 2(8):e198363. <https://doi.org/10.1001/jamanetworkopen.2019.8363>
 67. Massarweh NN, Devlin A, Symons RG, Broeckel Elrod JA, Flum DR (2009) Risk tolerance and bile duct injury: surgeon characteristics, risk-taking preference, and common bile duct injuries. *J Am Coll Surg* 209(1):17–24. <https://doi.org/10.1016/j.jamcollsurg.2009.02.063>
 68. Kim MG, Kwon SJ (2014) Comparison of the outcomes for laparoscopic gastrectomy performed by the same surgeon between a low-volume hospital and a high-volume center. *Surg Endosc* 28:1563–1570
 69. Spinelli A, Carrano FM, Laino ME, Andreozzi M, Koleth G, Hassan C, Repici A, Chand M, Savevski V, Pellino G (2023) Artificial intelligence in colorectal surgery: an AI-powered systematic review. *Tech Coloproctol* 27(8):615–629. <https://doi.org/10.1007/s10151-023-02772-8>
 70. Eppler MB, Sayegh AS, Maas M, Venkat A, Hemal S, Desai MM, Hung AJ, Grantcharov T, Cacciamani GE, Goldenberg MG (2023) Automated capture of intraoperative adverse events using artificial intelligence: a systematic review and meta-analysis. *J Clin Med* 12(4):1687. <https://doi.org/10.3390/jcm12041687>
 71. Seetohul J, Shafiee M, Sirlantzis K (2023) Augmented Reality (AR) for surgical robotic and autonomous systems: state of the art, challenges, and solutions. *Sensors (Basel)* 23(13):6202. <https://doi.org/10.3390/s23136202>
 72. Kitaguchi D, Takeshita N, Matsuzaki H, Oda T, Watanabe M, Mori K, Kobayashi E, Ito M (2020) Automated laparoscopic colorectal surgery workflow recognition using artificial intelligence: Experimental research. *Int J Surg* 79:88–94. <https://doi.org/10.1016/j.ijsu.2020.05.015>
 73. Kitaguchi D, Takeshita N, Matsuzaki H, Takano H, Owada Y, Enomoto T, Oda T, Miura H, Yamanashi T, Watanabe M, Sato D, Sugomori Y, Hara S, Ito M (2020) Real-time automatic surgical phase recognition in laparoscopic sigmoidectomy using the convolutional neural network-based deep learning approach. *Surg Endosc* 34(11):4924–4931. <https://doi.org/10.1007/s00464-019-07281-0>
 74. Baltussen EJM, Snaebjornsson P, de Koning SGB, Sterenborg HJCM, Aalbers AGJ, Kok N, Beets GL, Hendriks BHW, Kuhlmann KFD, Ruers TJM (2017) Diffuse reflectance spectroscopy as a tool for real-time tissue assessment during colorectal cancer surgery. *J Biomed Opt* 22(10):1–6. <https://doi.org/10.1117/1.JBO.22.10.106014>
 75. Baltussen EJM, Brouwer de Koning SG, Sanders J, Aalbers AGJ, Kok NFM, Beets GL, Hendriks BHW, Sterenborg HJCM, Kuhlmann KFD, Ruers TJM (2020) Using diffuse reflectance spectroscopy to distinguish tumor tissue from fibrosis in rectal cancer patients as a guide to surgery. *Lasers Surg Med* 52(7):604–611. <https://doi.org/10.1002/lsm.23196>
 76. Shademan A, Decker RS, Opfermann JD, Leonard S, Krieger A, Kim PC (2016) Supervised autonomous robotic soft tissue surgery. *Sci Transl Med* 8(337):337ra64. <https://doi.org/10.1126/scitranslmed.aad9398>
 77. Garcia-Martinez A, Vicente-Samper JM, Sabater-Navarro JM (2017) Automatic detection of surgical haemorrhage using computer vision. *Artif Intell Med* 78:55–60. <https://doi.org/10.1016/j.artmed.2017.06.002>
 78. Kugener G, Zhu Y, Pangal DJ, Sinha A, Markarian N, Roshannai A, Chan J, Anandkumar A, Hung AJ, Wrobel BB, Zada G, Donoho DA (2022) Deep neural networks can accurately detect blood loss and hemorrhage control task success from video. *Neurosurgery* 90(6):823–829. <https://doi.org/10.1227/neu.0000000000001906>
 79. Liu R, Liu Z, Wang X (2022) Intraoperative time-out to promote the implementation of the critical view of safety in laparoscopic cholecystectomy: on the comprehension of criteria, disease severity, and Hawthorne effects. *J Am Coll Surg* 234(6):1261–1262. <https://doi.org/10.1097/XCS.0000000000000147>

Authors and Affiliations

Diletta Corallino^{1,2}  · **Andrea Balla**^{3,4}  · **Diego Coletta**⁵  · **Daniela Pacella**⁶  · **Mauro Podda**⁷  · **Annamaria Pronio**⁸ · **Monica Ortenzi**⁹  · **Francesca Ratti**¹⁰ · **Salvador Morales-Conde**^{3,4}  · **Pierpaolo Sileri**¹¹  · **Luca Aldrighetti**¹⁰ 

✉ Diletta Corallino
diletta.corallino1989@gmail.com

Andrea Balla
andrea.balla@gmail.com

Diego Coletta
diegocoletta1@gmail.com

Daniela Pacella
pacelladaniela@gmail.com

Mauro Podda
mauropodda@ymail.com

Annamaria Pronio
annamaria.pronio@uniroma1.it

Monica Ortenzi
monica.ortenzi@gmail.com

Francesca Ratti
ratti.francesca@hsr.it

Salvador Morales-Conde
smoralesc@gmail.com

Pierpaolo Sileri
piersileri@yahoo.com

Luca Aldrighetti
aldrighetti.luca@hsr.it

³ Department of General and Digestive Surgery, University Hospital Virgen Macarena, University of Sevilla, Seville, Spain

⁴ Unit of General and Digestive Surgery, Hospital Quirónsalud Sagrado Corazón, Seville, Spain

⁵ General and Hepatopancreatobiliary Surgery, IRCCS Regina Elena National Cancer Institute, Rome, Italy

⁶ Department of Public Health, University of Naples Federico II, Naples, Italy

⁷ Department of Surgical Science, University of Cagliari, Cagliari, Italy

⁸ Department of General Surgery and Surgical Specialties, Sapienza University of Rome, Viale del Policlinico 155, 00161 Rome, Italy

⁹ Department of General and Emergency Surgery, Polytechnic University of Marche, Ancona, Italy

¹⁰ Hepatobiliary Surgery Division, IRCCS San Raffaele Scientific Institute, Faculty of Medicine and Surgery, Vita-Salute San Raffaele University, Via Olgettina 60, 20132 Milan, Italy

¹¹ Coloproctology and Inflammatory Bowel Disease Surgery Unit, IRCCS San Raffaele Scientific Institute, Faculty of Medicine and Surgery, Vita-Salute University, Via Olgettina 60, 20132 Milan, Italy

¹ Hepatobiliary Surgery Division, IRCCS San Raffaele Scientific Institute, Via Olgettina 60, 20132 Milan, Italy

² Department of General Surgery and Surgical Specialties “Paride Stefanini”, Sapienza University of Rome, Viale del Policlinico 155, 00161 Rome, Italy