Thieme

Cholangioscopy-based convoluted neuronal network vs. confocal laser endomicroscopy in identification of neoplastic biliary strictures

Authors
Carlos Robles-Medranda[1], Jorge Baquerizo-Burgos[1], Miguel Puga-Tejada[1], Domenica Cunto[1], Maria Egas-Izquierdo[1], Juan Carlos Mendez[2], Martha Arevalo-Mora[1], Juan Alcivar Vasquez[1], Hannah Lukashok[1], Daniela Tabacelia[3,4]

Institutions
1 Gastroenterology, Instituto Ecuatoriano de Enfermedades Digestivas – IECED, Guayaquil, Ecuador
2 Research and Development, mdconsgroup, Guayaquil, Ecuador
3 Gastroenterology, Elias Emergency University Hospital, Bucuresti, Romania
4 Universitatea de Medicină și Farmacie Carol Davila din București, Bucuresti, Romania

Key words
Pancreatobiliary (ERCP/PTCD), Cholangioscopy, Tissue diagnosis, Strictures

received 19.1.2024
accepted after revision 24.7.2024

Bibliography
Endosc Int Open 2024; 12: E1118–E1126
DOI 10.1055/a-2404-5699
ISSN 2364-3722

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Corresponding author
Dr. Carlos Robles-Medranda, Instituto Ecuatoriano de Enfermedades Digestivas – IECED, Gastroenterology, Av. Abel Romeo Castillo y Av. Juan Tanca Marengo , 090505 Guayaquil, Ecuador
carlosoakm@yahoo.es

ABSTRACT

Background and study aims Artificial intelligence (AI) models have demonstrated high diagnostic performance identifying neoplasia during digital single-operator cholangioscopy (DSOC). To date, there are no studies directly comparing AI vs. DSOC-guided probe-base confocal laser endomicroscopy (DSOC-pCLE). Thus, we aimed to compare the diagnostic accuracy of a DSOC-based AI model with DSOC-pCLE for identifying neoplasia in patients with indeterminate biliary strictures.

Patients and methods This retrospective cohort-based diagnostic accuracy study included patients ≥ 18 years old who underwent DSOC and DSOC-pCLE (June 2014 to May 2022). Four methods were used to diagnose each patient's biliary structure, including DSOC direct visualization, DSOC-pCLE, an offline DSOC-based AI model analysis performed in DSOC recordings, and DSOC/pCLE-guided biopsies. The reference standard for neoplasia was a diagnosis based on further clinical evolution, imaging, or surgical specimen findings during a 12-month follow-up period.

Results A total of 90 patients were included in the study. Eighty-six of 90 (95.5%) had neoplastic lesions including cholangiocarcinoma (98.8%) and tubulopapillary adenoma (1.2%). Four cases were inflammatory including two cases with chronic inflammation and two cases of primary sclerosing cholangitis. Compared with DSOC-AI, which obtained an area under the receiver operator curve (AUC) of 0.79, DSOC direct visualization had an AUC of 0.74 (P = 0.763), DSOC-pCLE had an AUC of 0.72 (P = 0.634), and DSOC- and pCLE-guided biopsy had an AUC of 0.83 (P = 0.809).

Conclusions The DSOC-AI model demonstrated an offline diagnostic performance similar to that of DSOC-pCLE, DSOC alone, and DSOC/pCLE-guided biopsies. Larger multicenter, prospective, head-to-head trials with a proportional sample among neoplastic and nonneoplastic cases are advisable to confirm the obtained results.

Introduction

Early and accurate diagnosis of neoplastic biliary strictures increases patient survival; nevertheless, the most commonly used technique, endoscopic retrograde cholangiopancreatography (ERCP) with brush cytology and biopsy forceps, has low diagnostic yield [1].

To overcome this limitation, new techniques, such as digital single-operator cholangioscopy (DSOC) and probe-based confocal laser endomicroscopy (pCLE), have been developed [1, 2]. DSOC enables high-resolution visualization of the bile duct system, tissue sampling, and interventional therapies [3, 4]. This technique, involving the visual interpretation of biliary malignancies, is superior to ERCP in terms of diagnosis of indeterminate biliary lesions, with an overall sensitivity of 94% and a specificity of 95% [5]. However, DSOC is an advanced endoscopic technique limited to specialized training and lack of training facilities and formal training guidelines [6].

On the other hand, pCLE enables in vivo histological evaluation of the mucosa through administration of a contrast agent such as fluorescein, which is administered intravenously and distributed within the epithelial extracellular matrix and lamina propria, allowing in vivo evaluation [7, 8]. This technique achieved adequate sensitivity and specificity for identifying neoplasia, ranging from 75% to 87% and 76% to 100%, respectively [1, 2]. However, when pCLE is combined with conventional tissue sampling through a DSOC-guided pCLE (DSOC-pCLE), the sensitivity and specificity are 93% and 82%, respectively [2]. Considering the cost, difficulty obtaining adequate interobserver agreement, lack of availability, and minimal improvement in diagnostic accuracy, utilization of pCLE is quite limited [2].

Currently, DSOC-guided tissue biopsy is utilized where conventional methods do not achieve diagnosis and as an adjunct to pCLE (when available) [2]. Through direct observation of the lesions, endoscopists can discern areas suggestive of neoplasia and obtain directed biopsies based on macroscopic patterns, reducing the number of inconclusive or inadequate tissue sampling, compared with ERCP; thus, reducing reinterventions and performing early management [9, 10, 11, 12].

A recent study comparing several methods indicated that the combination of pCLE and DSOC showed observed agreement superior to ERCP and DSOC alone (93.3% vs 70% and 90%, respectively) [13]. These results indicate that DSOC combined with pCLE can provide an accurate interpretation of biliary strictures. However, despite high diagnostic accuracy, these procedures are expensive, and the required equipment is not globally available. Therefore, there is a need to develop new technologies capable of providing additional data for detection of neoplasia and guide the best possible site for tissue acquisition.

In recent years, artificial intelligence (AI) tools have been applied in gastrointestinal endoscopy, mainly for cancer detection [14]. Notably, convolutional neural network models (CNNs) are capable of extracting information and classifying features from images and/or videos [14]. Currently, CNN models trained to identify areas suggestive of neoplasia in indeterminate biliary strictures have been proposed [15, 16, 17]. These models achieve high diagnostic accuracy when used to interpret images; nevertheless, until recently, these AI models could not be applied in real time during live procedures [15, 16]. However, a newly developed CNN model capable of real-time analysis (AI-Works – Cholangioscopy, mdconsgroup, Guayaquil, Ecuador) has shown an 80% observed agreement during clinical validation [18]. Currently, there is limited information comparing DSOC-based AI models and DSOC-pCLE for identification of neoplastic lesions. Thus, we aimed to compare the diagnostic accuracy of a DSOC-based AI model with DSOC-pCLE for identifying neoplasia in patients with indeterminate biliary strictures.

Patients and methods

Study design and ethics

This retrospective study was performed at the Instituto Ecuatoriano de Enfermedades Digestivas, a tertiary academic center in Guayaquil, Ecuador. The study protocol was approved by the institutional review board and designed in accordance with the Declaration of Helsinki and STARD 2015 guidelines. Patients or their legal guardians provided written informed consent for analysis and publication of the recorded videos before undergoing the procedures.

Population selection

Consecutive patients who underwent DSOC and DSOC-pCLE procedures from January 2014 to May 2022 were included. Data from patients aged ≥ 18 years old who underwent DSOC and DSOC-pCLE were included for analysis. Patients were excluded if: 1) any DSOC and/or DSOC-pCLE videos and/or images were unavailable; 2) they had no histological confirmation, based on biopsy or surgical resection; and/or 3) < 12 months of follow-up data were available.

Endoscopic procedures

DSOC procedure

All patients were placed in a supine position under general anesthesia and received antibiotic prophylaxis. Then they were assessed using a standard duodenoscope (Pentax ED 3670TK; Pentax Medical, Hoya Corp., Tokyo, Japan), Pentax EPK-I and EPK-i5010 video processors, and a second-generation SpyGlass DS Digital System (Boston Scientific, Marlborough, Massachusetts, United States) or a 9F eyeMAX cholangioscope (Microtech, Nanjing, China). The SpyScope DS II catheter (SpyGlass DS Digital System) or the eyeMAX cholangioscope were passed proximally into the bile duct, suction was used to remove bile, sterile saline solution was infused to optimize imaging, and the cholangioscope was slowly withdrawn. Ultimately, systematic inspection of the ductal mucosa was performed. A minimum of four biopsy samples were taken from areas suggestive of neoplasia and visual impressions of the endoscopist was classified according to the Carlos Robles-Medranda classification [11].

pCLE procedure

Consecutively after DSOC, patients underwent a DSOC-pCLE procedure. Patients were intravenously injected with 5 mL 10% fluorescein (BioGlo, Sofar Productos, Bogota, Colombia) and then pCLE was performed by passing a 1.0-mm Cellvizio CholangioFlex probe (Mauna Kea Technologies, Paris, France) through the cholangioscope working channel. The probe was gently placed in contact with the mucosa to avoid trauma. For DSOC-pCLE, the Miami malignancy criteria and the Paris inflammatory criteria were used. Miami malignancy criteria constituted thick white and dark bands, dark clumps, or epithelium; on the other hand, the Paris classification inflammatory criteria included vascular congestion, roughness aspect, increased interglandular space, and thickened reticular stricture [19, 20].

Offline video analysis by expert endoscopists

The pCLE videos were reviewed by two expert endoscopists (C. R.M. & J.A.V.) who were blinded to any clinical or ERCP information at the moment of offline video analysis. They indicated which descriptive criteria were present or absent in the videos.

pCLE Miami and Paris disaggregated criteria (pCLE, probe-based confocal laser endomicroscopy.):

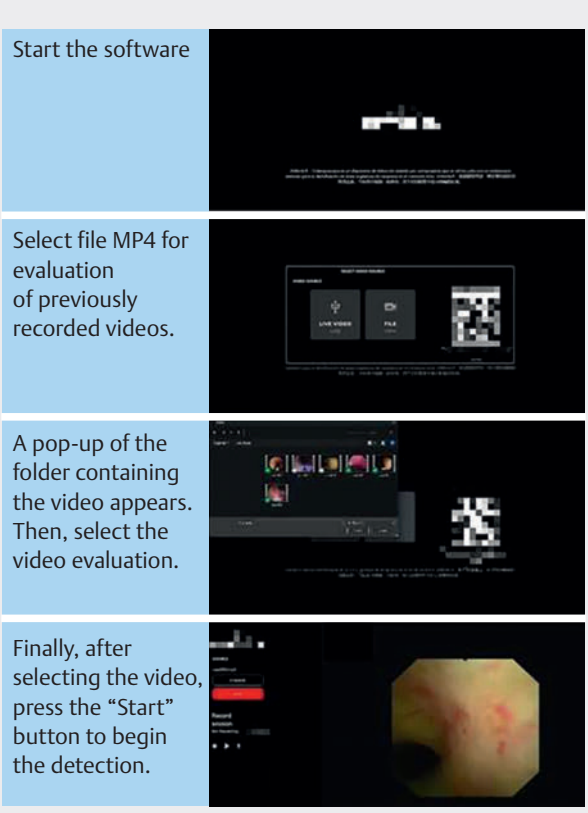
- Presence or absence of dark bands
- Presence or absence of white bands
- Presence or absence of dark clumps
- Presence or absence of epithelium
- Presence or absence of fluorescein leak
- Presence or absence of vascular congestion
- Presence or absence of rough appearance
- Presence or absence of reticular thickening
- Presence or absence of increased intraglandular space

The expert endoscopists evaluated all cases using both classifications (Miami and Paris) and more Miami classification malignancy criteria, observed by the endoscopists, than Paris classification inflammatory criteria constituted neoplasia. For DSOC- and pCLE-guided biopsy, corresponding histological findings constituted neoplasia.

Offline video analysis with AIWorks-Cholangioscopy software

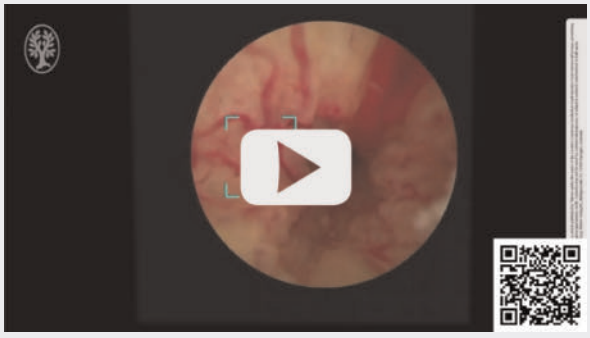
All recorded DSOC procedures were analyzed offline with the AIWorks-Cholangioscopy software (mdconsgroup, Guayaquil, Ecuador). AIWorks-Cholangioscopy is a CNN trained to identify bile duct mucosal abnormalities, commonly observed features of neoplastic lesions that can be recognized by various currently available classification systems (CRM, Mendoza) [11, 21], when identifying the following criteria: neovascularity, surface irregularities, polyploids, ulcerations or friability. This model was developed using YOLOv5 (You Only Look Once version 5, Washington, United States) and can delimit potential neoplastic lesions with a bounding box with high recall and precision (▶ **Fig. 1**) [18].

For each patient, model output highlighted any areas that were deemed suggestive of neoplasia with a bounding box, and a screenshot that could be accessed within the software was automatically saved (▶ **Video 1**). For the present study, an observer (J.B-B.) blinded to any clinical data or ERCP information recorded the judgement output by the AI and marked detection or absence of detection of areas suggestive of neoplasia provided by the AIWorks-Cholangioscopy software for each DSOC video. The information was registered in a database for comparison with the pathology, 12-month follow-up, DSOC-pCLE, and DSOC visual impression data.



▶ **Fig. 1** Step-by-step visual interpretation of application of the AIWorks-Cholangioscopy (mdconsgroup, Guayaquil, Ecuador) software for video analysis.

VIDEO



▶ **Video 1** utomatic detection of areas suggestive of neoplasia using the AIWorks-Cholangioscopy software in a patient with biopsy-confirmed cholangiocarcinoma. The detected areas are highlighted with green bounding boxes.

Statistical analysis

Technical considerations

Statistical analysis was performed using R v4.1.2 (R Foundation for Statistical Computing; Vienna, Austria) by our institutional biostatistician (M.P-T.). $P < 0.05$ was considered to indicate statistical significance.

Sample size

Based on the observed agreement of 93.3% and 80% of DSOC-pCLE [13] and DSOC-based AI model [18], respectively, a size effect of $h = 0.4036$ was calculated. Considering a 5% alfa and 20% beta errors, a sample size of 76 pairs was necessary to demonstrate DSOC-pCLE observed agreement was higher than that reached by the DSOC-based AI model, with 80% power statistic. A pair was defined as a patient who was assessed through DSOC-pCLE and DSOC-based AI model.

Descriptive analysis

Numerical variables are presented as the mean (standard deviation) or median (interquartile range) depending on the normality of their statistical distribution, which was assessed with the Kolmogorov–Smirnov test. Corresponding categorical variables are described as frequencies (%) with 95% confidence intervals.

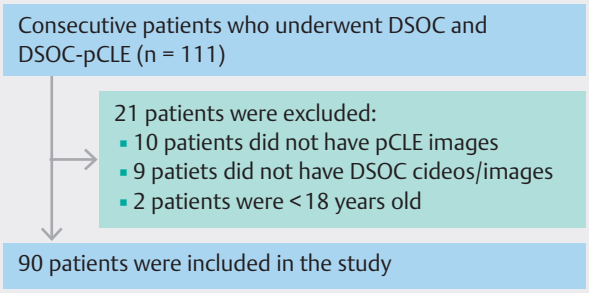
Diagnostic accuracy

Sensitivity, specificity, positive and negative predictive values, and observed agreement of DSOC visual impression, DSOC-pCLE, DSOC and pCLE-guided biopsy, and DSOC-AI model were calculated to evaluate diagnostic accuracy. Receiver operator characteristic analysis was performed, and to compare the areas under the receiver operator curve (AUCs) of the diagnostic methods, we performed DeLong's test. The reference standard for neoplasia was based on further clinical evaluation, imaging, or surgical specimen findings during a 12-month follow-up period. A subanalysis by distal vs proximal lesion location was performed.

Results

Baseline characteristics

A total of 90 patients were included in the study and 21 patients were excluded (▶ **Fig. 2**). Baseline characteristics of the patients were as follows (▶ **Table 1**). Mean age was 66.4 years ± 13.7 and 56.7% were female. Tumor suspicion was the most common indication (55.6%). Strictures were most commonly proximal and middle lesions from the common bile duct (CBD) (42.2%), with a median size of 22.0 mm (14.3–25.0 mm). Twenty-five of 90 patients (27.8%) had previously undergone ERCP at least once. A total of 26.1% of cases had previous plastic stent placement, with a median duration of 60 days (18.0–96.0). The most common neoplastic lesion based on histopathological results was cholangiocarcinoma (94.4%) (▶ **Fig. 3**). Of the patients, 15.5% underwent surgery.



▶ **Fig. 2** Flowchart of patient recruitment procedures.

Diagnostic accuracy

For evaluating diagnostic accuracy, the four diagnostic methods were compared with respect to several performance metrics (▶ **Table 2** [total]). Diagnoses based on DSOC visual impressions achieved 95.6% observed agreement, with 96.5% sensitivity and 75% specificity. pCLE achieved a 94.4%, 94.2%, and 100% observed agreement, sensitivity, and specificity, respectively. DSOC and pCLE-guided biopsies achieved 97.7%, 100% and 97.9% sensitivity, specificity, and observed agreement, respectively. Finally, the DSOC-based AI model detection software achieved 97.7% sensitivity, 75% specificity and 96.7% observed agreement. The AUC for DSOC-AI was 0.790, DSOC direct visualization had an AUC of 0.740, pCLE had an AUC of 0.720, and DSOC and pCLE-guided biopsies had an AUC of 0.830. When comparing with DSOC-AI using the DeLong's test, there were no statistically significant differences among diagnostics accuracy of these methods (DSOC direct visualization, $P = 0.763$; pCLE, $P = 0.634$; and pCLE-guided biopsies, $P = 0.809$).

In evaluating distal lesions, specifically from the distal or middle CBD, DSOC direct visualization, as well as DSOC and pCLE-guided biopsy, demonstrated superior performance compared with the DSOC-based AI model. When comparing with the AI, both methods achieved the same higher sensitivity (97.22% vs 94.44%), negative predictive value (NPV) (66.67% vs 50%), and observed agreement (97.37% vs 94.74%), although the differences were not statistically significant ($P = 0.707$). Conversely, in distal lesions, DSOC-guided pCLE showed the lowest sensitivity (88.89%), NPV (33.33%), and observed agreement (89.47%), with the differences again not statistically significant ($P = 0.643$) (▶ **Table 2** [lesions from distal and middle CBD]).

Although DSOC direct visualization performed well in assessing lesions in the distal or middle CBD, a contrasting situation was observed when evaluating lesions in the proximal CBD, common hepatic duct, hilum, and intrahepatic duct (52; 57.8%). Here, the DSOC-based AI model achieved the highest sensitivity (100%) and NPV (100%), along with the same highest observed agreement (98.08%) as DSOC-guided pCLE and DSOC and pCLE-guided biopsy. Although the DSOC-based AI model did not significantly outperform DSOC-guided pCLE and DSOC and pCLE-guided biopsy, it did achieve a significantly higher

▶ **Table 1** Baseline characteristics.

	N = 90
Age (years), mean ± SD	66.4 ± 13.7
▪ 18–39 years old, n (%)	3 (3.3)
▪ 40–64 years old	34 (37.8)
▪ ≥ 65 years old	53 (58.9)
Sex (female), n (%)	51 (56.7)
Main indication, n (%)	
▪ Tumor suspicion	50 (55.6)
▪ Bile duct obstruction	31 (34.4)
▪ Indeterminate bile stricture	9 (10.0)
Lesion location, n (%)	
▪ Distal lesions	38 (42.2)
– Distal CBD	12/48
– Middle CBD	26/48
▪ Proximal lesions	52 (57.8)
– Proximal CBD	30/52
– Common hepatic duct	6/52
– Hepatic hilum	11/52
– Intrahepatic duct	5/52
Lesion size (mm), median (IQR)	20 (14.3–25.0)
Previously performed ERCP, n (%)	25 (27.8)
Previous stent placement, n (%)	18 (20.0)
Stent placement duration (days), median (IQR)	60.0 (18.0–96.0)
Performed biopsy, n (%)	90 (100.0)
No. of biopsy samples taken, median (IQR)	4 (1–6)
Diagnosis after histopathological confirmation, n (%)	
▪ Cholangiocarcinoma	85 (94.4)
▪ Tubulopapillary adenoma	1 (1.1)
▪ Chronic inflammation	2 (2.2)
▪ Primary sclerosing cholangitis	2 (2.2)

SD, standard deviation; CBD, common bile duct; IQR, interquartile range; ERCP, endoscopic retrograde cholangiopancreatography.

AUC compared with DSOC direct visualization (0.990 vs 0.656; $P < 0.001$) (▶ **Table 2** [lesions from proximal CBD, common hepatic duct, hilum, and intrahepatic duct]).

Discussion

The present study aimed to compare the diagnostic accuracy of a DSOC-based AI model with DSOC-pCLE for identifying neoplasia in patients with indeterminate biliary strictures. We found that the diagnostic accuracy for neoplastic lesion identification of the AI model was similar to that for DSOC-pCLE, DSOC- and pCLE-guided biopsies and diagnoses based on DSOC visual impressions, indicating that AI tools can accurately aid endoscopists in identification of neoplasia during DSOC.

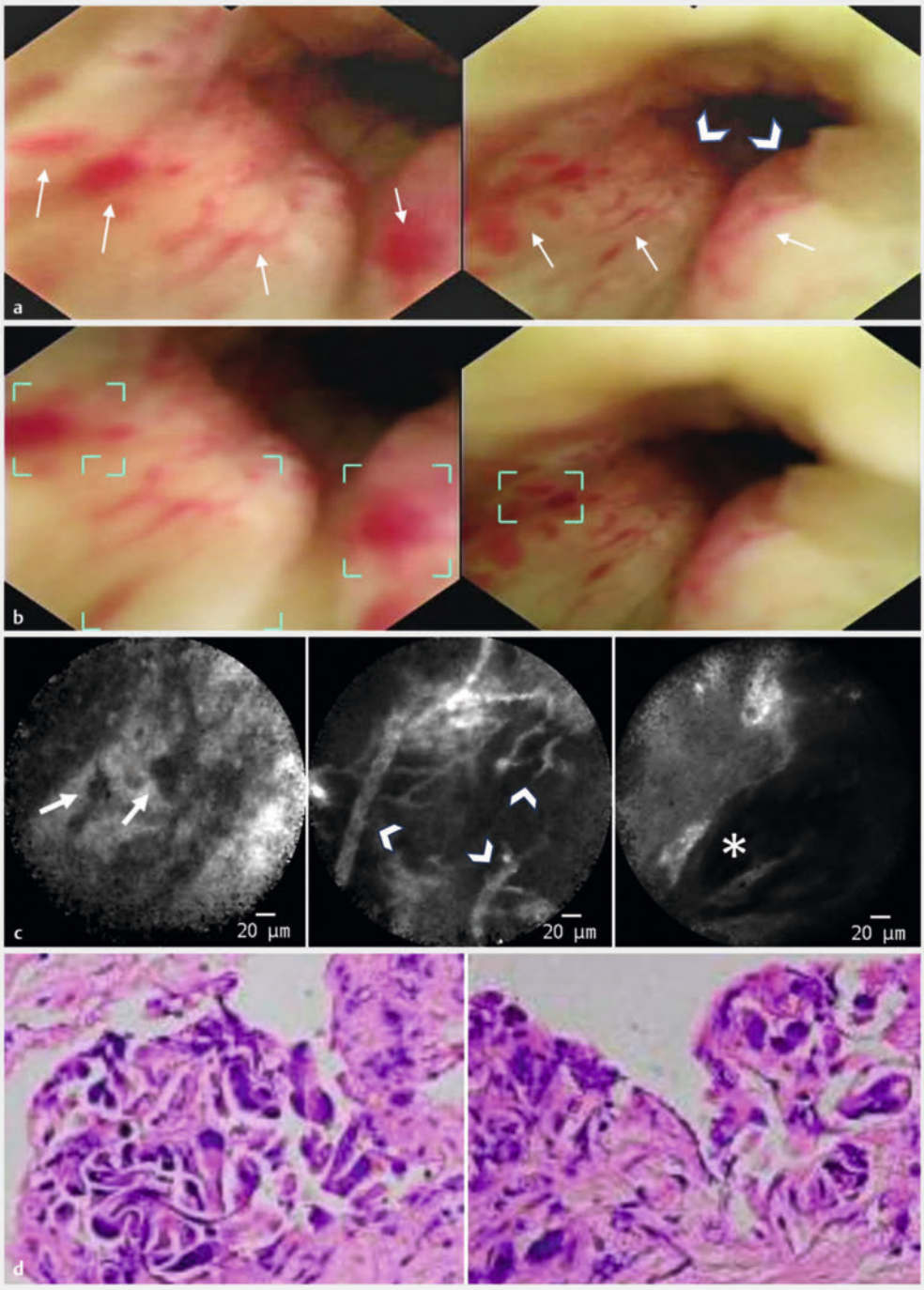
AI models can be classified as computer-assisted detection (CADe) or diagnostic devices (CADx). This classification depends on whether the model detects mucosal lesions or patterns (CADe) or characterizes lesions, for example, as benign or malignant lesions (CADx) [22]. Currently, there are still no CADx tools for DSOC capable of accurately classifying indeterminate biliary strictures, and thus, diagnosis is dependent on biopsy sampling. On the other hand, several CADe tools have been proposed, but their widespread use is limited to application in still images and it has not been able to perform adequately during real-time procedures [15, 16]. In addition, the diagnostic accuracy of CADe tools previously has been evaluated with still images only, and thus, the high diagnostic accuracy obtain in these cases cannot be extrapolated to reflect real clinical applications.

Robles-Medranda et al. performed a multicenter study initially evaluating application of the AIWorks-Cholangioscopy software (mdconsgroup, Guayaquil, Ecuador) to still images and obtained values similar to those obtained by Saraiva et al and Pereira et al., with an observed agreement > 90.0% [15, 16, 18]. However, when recorded DSOC procedures of patients with neoplasia were input into the model, its diagnostic accuracy changed, showing a sensitivity of 90.5%, a specificity of 68.2%, a PPV of 74.0%, an NPV of 87.8%, and an observed agreement of 80.0%.

In the present study, we performed offline assessment of prerecorded DSOC videos of patients with a confirmed diagnosis of neoplastic lesions. The AI model obtained the following parameters: 97.7% sensitivity, 75.0% specificity, 98.8% PPV, 60.0% NPV, and 96.7% observed agreement. These values were similar to those obtained for diagnosis based on endoscopists' visual impressions (96.5%, 75.0%, 98.8%, 50.0%, 95.6%, respectively). In addition, Robles-Medranda et al. compared the AI software with diagnosis based on visual impressions of a group of expert advanced endoscopists and a group of nonexperts; the model obtained higher diagnostic accuracy than both groups. Similarly, here, the AI model achieved a higher AUC (0.790) than diagnosis based on visual impressions (0.740, $P = 0.763$); however, the difference between the two methods was not significant.

A recent systematic review evaluating application of CADe tools during cholangioscopy included five studies. The models analyzed in these studies obtained high diagnostic accuracy [23]. The CNN model applied for lesion detection in still images achieved 98.6%, 98.0%, and 98.0% sensitivity, specificity, and observed agreement, respectively [18], with a 60 frames-per-second reading rate. Thus, CNN models appear to be promising AI tools for diagnosis of potentially malignant biliary strictures.

pCLE technology allows microscopy detection in vivo through application of fluorescein [24, 25, 26]. Two classifications have been proposed for biliary stricture analysis: the Miami classification, based on the malignancy criteria, and the



▶ **Fig. 3** The case of an 88-year-old male with a lesion extending from the middle third portion of the common bile duct to the common hepatic duct is presented as representative example. **a** Using the 9F eyeMAX cholangioscope, a raised intraductal lesion (arrowhead) with increased vascularity (arrows) was observed. **b** AIWorks-Cholangioscopy software detected areas suggestive of neoplasia (green bounding box) in the same locations identified with the cholangioscope. **c** pCLE findings revealed dark bands (*), white bands (arrowhead) and dark clumps (arrows). **d** Hematoxylin and eosin biopsy slide (100x) revealed malignant glandular tissue covered with stratified cylindric cells and atypical hypertrophic hyperchromatic nuclei compatible with cholangiocarcinoma.

Paris criteria, based on inflammatory criteria [26]. The Miami classification has a high sensitivity (98.0%) but low specificity (67.0%) due to increased false-positive rates caused by lack of criteria ensuring unique identification of malignant lesions [27]. Compared with the Miami classification, the sensitivity and specificity achieved by the Paris criteria for inflammatory lesions were 81.0% and 83.0%, respectively [28].

▶ **Table 2** Diagnostic accuracy of visual impressions during digital single-operator cholangioscopy (DSOC), DSOC-guided probed-based confocal endomicroscopy (pCLE), DSOC and pCLE-guided biopsy, and a DSOC-based artificial intelligence (AI) model [n/T; % (95% CI)].

	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Observed agreement	AUC (*P* value*)
Total (N = 90)						
DSOC-based AI model	97.7% (91.9–99.7)	75% (19.4–99.4)	98.8% (93.6–99.9)	60% (14.7–94.7)	96.7% (90.6–99.3)	0.790 (reference)
DSOC direct visualization	96.5% (90.1–99.3)	75.0% (19.4–99.4)	98.8% (93.5–99.9)	50.0% (11.8–88.2)	95.6% (89.0–98.8)	0.740 (*P* = 0.763)
DSOC-guided pCLE	94.2% (86.9–98.1)	100% (39.8–100)	100% (95.6–100)	44.4% (13.7–78.8)	94.4% (87.5–98.2)	0.720 (*P* = 0.634)
DSOC and pCLE-guided biopsy	97.7% (91.9–99.7)	100% (39.8–100)	100% (95.7–100)	66.7% (22.3–95.7)	97.8% (92.2–99.7)	0.830 (P = 0.809)
Lesions from distal and middle CBD (*n* = 38)						
DSOC-based AI model	94.44 (81.34–99.32)	100 (15.81–100)	100 (89.72–100)	50 (6.76–93.24)	94.74 (82.25–99.36)	0.750 (reference)
DSOC direct visualization	97.22 (85.47–99.93)	100 (15.81–100)	100 (90–100)	66.67 (9.43–99.16)	97.37 (86.19–99.93)	0.833 (*P* = 0.707)
DSOC-guided pCLE	88.89 (73.94–96.89)	100 (15.81–100)	100 (89.11–100)	33.33 (4.33–77.72)	89.47 (75.2–97.06)	0.667 (*P* = 0.643)
DSOC and pCLE-guided biopsy	97.22 (85.47–99.93)	100 (15.81–100)	100 (90–100)	66.67 (9.43–99.16)	97.37 (86.19–99.93)	0.833 (*P* = 0.707)
Lesions from proximal CBD, common hepatic duct, hilum, and intrahepatic duct (*n* = 52)						
DSOC-based AI model	100 (92.89–100)	50 (1.26–98.74)	98.04 (89.55–99.95)	100 (2.5–100)	98.08 (89.74–99.95)	0.990 (reference)
DSOC direct visualization	96 (86.29–99.51)	50 (1.26–98.74)	97.96 (89.15–99.95)	33.33 (0.84–90.57)	94.23 (84.05–98.79)	0.656 (*P* < 0.001)
DSOC-guided pCLE	98 (89.35–99.95)	100 (15.81–100)	100 (92.75–100)	66.67 (9.43–99.16)	98.08 (89.74–99.95)	0.833 (*P* = 0.108)
DSOC and pCLE-guided biopsy	98 (89.35–99.95)	100 (15.81–100)	100 (92.75–100)	66.67 (9.43–99.16)	98.08 (89.74–99.95)	0.833 (*P* = 0.108)

CI, confidence interval; AUC, area under the curve; pCLE, probed-based confocal laser endomicroscopy
*DeLong's test.

Using both classifications simultaneously to establish an objective measurement, such as image scoring systems, could improve pCLE diagnostic performance [29]. In the present study, two observers analyzed the pCLE data and marked the presence or absence of observed features from both the Miami and Paris classifications in a disaggregated manner. Then, applying both classifications though Boolean operator, we determined if a lesion was malignant or inflammatory based on presence of most Miami criteria detected or most Paris criteria, respectively. Using this analysis, DSOC-pCLE achieved 94.2% and 100% sensitivity and specificity, respectively.

Both modalities (DSOC-AI and pCLE) had similar AUCs (0.760 vs. 0.720, *P* = 0.634) and high observed agreement (96.7% and 94.4%, respectively). However, both techniques obtained nominally lower diagnostic accuracy than biopsies. Despite the lack of a significant difference between the DSOC-AI, DSOC-pCLE, and DSOC/pCLE-guided biopsies, biopsies are still necessary for diagnosis confirmation; however, AI tools will be able to provide real-time assistance in terms of AI-guided biopsy sampling during DSOC procedures, improving the quality of samples obtained and reducing the rate of inadequate biopsies. In addition, AI models could be used as second opinion tools aiding in the decision-making process in complicated cases in which visual impression is uncertain [30]. As computational resources improve, faster and more accurate AI models could be developed to improve diagnostic accuracy and provide higher detection and recall, increasing the quality of samples [31].

In addition to their diagnostic performance, the high cost of DSOC and pCLE must be considered [13]. Few studies have evaluated the costs of single-use pCLE probes, and none have investigated the cost of AI-assisted DSOC. Notably, Tanisaka et al. reported that the costs for pCLE with DSOC surpass those for ERCP [13]. The potential reason for the procedure's high cost includes the single-use pCLE probe and the elevated costs of

the most commonly available cholangioscope. The AIWorks-Cholangioscopy model used here was trained with different brands of cholangioscopes (eyeMAX, Micro-Tech, Nianjing, China; Spyglass, Boston Scientific, Marlborough, Massachusetts, United States) and is cloud-based software; therefore, the software can be paired with different brands of cholangioscopes with potentially lower cost. This tool may allow AI-guided DSOC biopsy sampling to be performed as an alternative to DSOC-pCLE biopsy, requiring the same amount of equipment as a traditional DSOC, and therefore, less equipment compared with DSOC-pCLE. In addition, by performing biopsies guided by AI, the number of samples required, and reinterventions could be reduced [18].

The limitations of the present study included its retrospective analysis and single-center design, and the low number of benign cases. Considering the low NPV obtained by the different studied methods, it is advisable to continue using pCLE in cases in in which no neoplastic lesions were recognized during an AI-guided DSOC. Even so, this study provided information about the clinical application of several DSOC-related methods. Real-world studies are required to compare the AI model and pCLE to confirm the obtained results. In addition, studies evaluating the cost-effectiveness of both procedures are needed.

Conclusions

In conclusion, the DSOC-AI model demonstrated offline diagnostic accuracy similar to that of DSOC-pCLE, DSOC alone, and DSOC- and pCLE-guided biopsies, providing insight about novel technologies that could aid endoscopists in accurate identification of neoplasia.

Conflict of Interest

Carlos Robles-Medranda is a key opinion leader and consultant for Pentax Medical, Steris, Medtronic, Motus, Micro-tech, G-Tech Medical Supply, CREO Medical, EndoSound, and mdconsgroup. The other authors declare no conflicts of interest.

References

[1] Koda H, Hara K, Nozomi O et al. High-resolution probe-based confocal laser endomicroscopy for diagnosing biliary diseases. Clin Endosc 2021; 54: 924–929 doi:10.5946/ce.2020.191

[2] Pilonis ND, Januszewicz W, di Pietro M. Confocal laser endomicroscopy in gastro-intestinal endoscopy: technical aspects and clinical applications. Transl Gastroenterol Hepatol 2022; 7: 1–20 doi:10.21037/tgh.2020.04.02

[3] Machicado JD, Raijman I, Shah RJ. Future of cholangioscopy. Gastrointest Endosc Clin N Am 2022; 32: 583–596 doi:10.1016/j.giec.2022.03.002

[4] Wen LJ, Chen JH, Xu HJ et al. Efficacy and safety of digital single-operator cholangioscopy in the diagnosis of indeterminate biliary strictures by targeted biopsies: a systematic review and meta-analysis. Diagnostics (Basel) 2020; 10: 666 doi:10.3390/diagnostics10090666

[5] de Oliveira PVAG, de Moura DTH, Ribeiro IB et al. Efficacy of digital single-operator cholangioscopy in the visual interpretation of indeterminate biliary strictures: a systematic review and meta-analysis. Surg Endosc 2020; 34: 3321–3329

[6] Sethi A, Shah RJ. Cholangioscopy and Pancreatoscopy. In: Cohen Jonathan (eds.) Successful Training in Gastroenterology Endoscopy.. New Jersey: Wiley; 2022: 133–142

[7] ASGE Technology Committee. Confocal Laser Endomicroscopy. Gastrointest Endosc 2014; 80: 928–938 doi:10.1016/j.gie.2014.06.021

[8] Wallace MB, Meining A, Canto MI et al. The safety of intravenous fluorescein for confocal laser endomicroscopy in the gastrointestinal tract. Aliment Pharmacol Ther 2010; 31: 548–552 doi:10.1111/j.1365-2036.2009.04207.x

[9] Kahaleh M, Raijman I, Gaidhane M et al. Digital cholangioscopic interpretation: when north meets the south. Dig Dis Sci 2021; 67: 1345–1351

[10] Robles-Medranda C, Oleas R, Sánchez-Carriel M et al. Vascularity can distinguish neoplastic from non-neoplastic bile duct lesions during digital single-operator cholangioscopy. Gastrointest Endosc 2021; 93: 935–941

[11] Robles-Medranda C, Valero M, Soria-Alcivar M et al. Reliability and accuracy of a novel classification system using peroral cholangioscopy for the diagnosis of bile duct lesions. Endoscopy 2018; 50: 1059–1070 doi:10.1055/a-0607-2534

[12] Sethi A, Tyberg A, Slivka A et al. Digital single-operator cholangioscopy (DSOC) improves interobserver agreement (IOA) and accuracy for evaluation of indeterminate biliary strictures: the Monaco classification. J Clin Gastroenterol 2022; 56: 94–97 doi:10.1097/MCG.0000000000001321

[13] Tanisaka Y, Ryozawa S, Nonaka K et al. Diagnosis of biliary strictures using probe-based confocal laser endomicroscopy under the direct view of peroral cholangioscopy: results of a prospective study (with video). Gastroenterol Res Pract 2020; 2020: 1–9

[14] Min JK, Kwak MS, Cha JM. Overview of deep learning in gastrointestinal endoscopy. Gut Liver 2019; 13: 388–393 doi:10.5009/gnl18384

[15] Pereira P, Mascarenhas M, Ribeiro T et al. Automatic detection of tumor vessels in indeterminate biliary strictures in digital single-operator cholangioscopy. Endosc Int Open 2022; 10: 262–268 doi:10.1055/a-1723-3369

[16] Saraiva MM, Ribeiro T, Ferreira JPS et al. Artificial intelligence for automatic diagnosis of biliary stricture malignancy status in single-operator cholangioscopy: a pilot study. Gastrointest Endosc 2022; 95: 339–348 doi:10.1016/j.gie.2021.08.027

[17] Marya NB, Powers PD, Petersen BT et al. Identification of patients with malignant biliary strictures using a cholangioscopy-based deep learning artificial intelligence (with video). Gastrointest Endosc 2023; 97: 268–278

[18] Robles-Medranda C, Baquerizo-Burgos J, Alcívar-Vásquez J et al. Artificial intelligence for diagnosing neoplasia on digital cholangioscopy: development and multicentric validation of a convolutional neural network model. Endoscopy 2023; 55: 719–727

[19] Taunk P, Singh S, Lichtenstein D et al. Improved classification of indeterminate biliary strictures by probe-based confocal laser endomicroscopy using the Paris criteria following biliary stenting. J Gastroenterol Hepatol 2017; 32: 1778–1783 doi:10.1111/jgh.13782

[20] Wallace M, Lauwers GY, Chen Y et al. Miami classification for probe-based confocal laser endomicroscopy. Endoscopy 2011; 43: 882–891 doi:10.1055/s-0030-1256632

[21] Kahaleh M, Gaidhane M, Shahid HM et al. Digital single-operator cholangioscopy interobserver study using a new classification: the Mendoza classification (with video). Gastrointest Endosc 2022; 95: 319–326

[22] Firmino M, Angelo G, Morais H et al. Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy. Biomed Eng Online 2016; 15: 2

[23] Njei B, McCarty TR, Mohan BP et al. Artificial intelligence in endoscopic imaging for detection of malignant biliary strictures and cholangiocarcinoma: a systematic review. Ann Gastroenterol 2023; 36: 223–230 doi:10.20524/aog.2023.0779

[24] Abramov I, Dru AB, Belykh E et al. Redosing of fluorescein sodium improves image interpretation during intraoperative ex vivo confocal laser endomicroscopy of brain tumors. Front Oncol 2021; 11: 1–14

[25] Becker V, von Delius S, Bajbouj M et al. Intravenous application of fluorescein for confocal laser scanning microscopy: evaluation of contrast dynamics and image quality with increasing injection-to-imaging time. Gastrointest Endosc 2008; 68: 319–323

[26] Wallace MB, Meining A, Canto MI et al. The safety of intravenous fluorescein for confocal laser endomicroscopy in the gastrointestinal tract. Aliment Pharmacol Ther 2010; 31: 548–552 doi:10.1111/j.1365-2036.2009.04207.x

[27] Taunk P, Singh S, Lichtenstein D et al. Improved classification of indeterminate biliary strictures by probe-based confocal laser endomicroscopy using the Paris criteria following biliary stenting. J Gastroenterol Hepatol 2017; 32: 1778–1783 doi:10.1111/jgh.13782

[28] Kahaleh M, Giovannini M, Jamidar P et al. Probe-based confocal laser endomicroscopy for indeterminate biliary strictures: refinement of the image interpretation classification. Gastroenterol Res Pract 2015; 2015: 1–5 doi:10.1155/2015/675210

[29] Sievert M, Mantsopoulos K, Mueller SK et al. Systematic interpretation of confocal laser endomicroscopy: larynx and pharynx confocal imaging score. Acta Otorhinolaryngol Ital 2022; 42: 26–33 doi:10.14639/0392-100X-N1643

[30] Ricaurte-Ciro J, Baquerizo-Burgos J, Carvajal-Gutierrez J et al. Usefulness of artificial intelligence-assisted digital single-operator cholangioscopy as a second-opinion consultation tool during interhospital assessment of an indeterminate biliary stricture: a case report. VideoGIE 2023; 8: 364–366

[31] Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng 2018; 2: 719–731 doi:10.1038/s41551-018-0305-z