Application of artificial intelligence in pancreaticobiliary diseases

Hemant Goyal (iD)**, Rupinder Mann, Zainab Gandhi, Abhilash Perisetti, Zhongheng Zhang, Neil Sharma, Shreyas Saligram, Sumant Inamdar and Benjamin Tharian**

Abstract: The role of artificial intelligence and its applications has been increasing at a rapid pace in the field of gastroenterology. The application of artificial intelligence in gastroenterology ranges from colon cancer screening and characterization of dysplastic and neoplastic polyps to the endoscopic ultrasonographic evaluation of pancreatic diseases. Artificial intelligence has been found to be useful in the evaluation and enhancement of the quality measure for endoscopic retrograde cholangiopancreatography. Similarly, artificial intelligence techniques like artificial neural networks and faster region-based convolution network are showing promising results in early and accurate diagnosis of pancreatic cancer and its differentiation from chronic pancreatitis. Other artificial intelligence techniques like radiomics-based computer-aided diagnosis systems could help to differentiate between various types of cystic pancreatic lesions. Artificial intelligence and computer-aided systems also showing promising results in the diagnosis of cholangiocarcinoma and the prediction of choledocholithiasis. In this review, we discuss the role of artificial intelligence in establishing diagnosis, prognosis, predicting response to treatment, and guiding therapeutics in the pancreaticobiliary system.

Keywords: artificial intelligence, choledocholithiasis, computer-aided diagnosis, endoscopic ultrasound, pancreatic cancer

Received: 2 August 2020; revised manuscript accepted: 11 January 2021.

Correspondence to:

Hemant Goyal

The Wright Center for Graduate Medical Education, 501 S. Washington Avenue, Scranton, PA 18505, USA.
doc.hemant@yahoo.com

Rupinder Mann

Academic Hospitalist, Saint Agnes Medical Center, Fresno, CA, USA

Zainab Gandhi

Department of Medicine, Geisinger Community Medical Center, Scranton, PA, USA

Abhilash Perisetti

Department of Gastroenterology and Hepatology, The University of Arkansas for Medical Sciences, Little Rock, AR, USA

Zhongheng Zhang

Department of emergency medicine, Sir Run-Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China

Neil Sharma

Division of Interventional Oncology & Surgical Endoscopy (IOSE), Parkview Cancer Institute, Fort Wayne, IN, USA

Indiana University School of Medicine, Fort Wayne, IN, USA

Shreyas Saligram

Division of Advanced Endoscopy, Gastroenterology, Hepatology, and Nutrition, Department of Medicine, University of Texas Health, San Antonio, TX, USA

Sumant Inamdar & Benjamin Tharian

University of Arkansas for Medical Sciences, Little Rock, AR, USA

Introduction

Artificial intelligence (AI) is a generic term to address the capability of a computer to perform tasks as an intelligent being to imitate the "cognitive" functions of a human mind like the capability to "learn."[1] The concept of AI was present in as early as Aristotle's (384–322 BC) study of logic; however, Alan Turing built the first operational computer in 1940 from the modern world, which is known as Electromechanical Heath Robinson.[2] Intensive research focused on using AI applications in the medical field is underway, which could provide unprecedented opportunities to make an accurate diagnosis and improve healthcare quality in clinical practice.[3] The AI in medicine has two main branches, virtual (comprises of machine learning (ML) and deep learning (DL)), and physical.[4] ML and DL are frequently used interchangeably; however, each term has a specific meaning and significance. ML can be defined as the capability of a system to "learn" by giving data without any programming involved. In ML, the computer system understands the patterns or features of the data and makes predictions by recognition.[2] Support vector machines (SVMs), Bayesian inferences, decision trees, logistic regression, linear discriminants, and artificial neural networks (ANNs) are different models of ML.[5] The ML algorithms are used in the layers of non-linear processing in two steps for DL: (1) transformation refers to building an effective data model. (2) Feature extraction refers to computational automation to focus on a targeted featured aspect of the data to increase the power of prediction.[6] The physical branch is the second branch of AI in medicine, including medical devices and physical objects such as robots.[4]

The application of AI in the field of gastroenterology is increasing at a rapid pace, ranging from diagnosis and characterization of dysplastic and neoplastic polyps to cystic fluid analysis and accurate prediction models to determine the need for intervention with computer-aided diagnosis (CAD).[3,7,8] Furthermore, AI has shown to aid in diagnosing pancreaticobiliary diseases like detecting the various pancreaticobiliary tumors, differentiating cystic pancreatic lesions, and quality measures for endoscopic retrograde cholangiopancreatography (ERCP) (Table 1).[9] Moreover, endoscopic procedures also rely on inter-rater reliability mainly. With the use of AI, predictive models can be developed based on the area of interest to either increase the diagnostic accuracy or provide analysis on the probability of benefit from a therapeutic intervention. This AI implementation can bring personalized medicine to the commercial world and help provide better care to the patients.[3] In this review, we aim to discuss the role of the application of AI in benign and neoplastic diseases of the pancreaticobiliary system.

AI in endoscopic ultrasound

Endoscopic ultrasound (EUS) is one of the most common tests performed to diagnose pancreatic lesions. EUS-guided biopsy of pancreatic lesions is the commonly used modality for diagnosing pancreatic cancer (PC) whenever available. The diagnostic accuracy of EUS (85–90%) is much higher than computed tomography (CT) scan-based imaging methods (50%).[12] Early diagnosis of PC remains essential to improve the survival rates in these patients.[12] However, the reliability of EUS for PC diagnosis decreases in patients with chronic pancreatitis (CP) due to the presence of inflammation, scarring, and calcification. Besides, one of the limitations of EUS is the target area dependence and the accuracy of choosing the area to image. The subtle changes in tissue construction can be detected by digital image analysis. CAD can help differentiate various lesions and improve the diagnostic accuracy of EUS in these cases.[12,14,23] CAD uses both physicians (subjective aspect) and computers (objective aspect) equally to provide an output that can be used as a "second opinion" to mitigate the shortage of physicians and helps to improve the diagnosis.[12]

Zhu *et al.* studied EUS images to differentiate between CP and PC using the SVM. The study recruited 262 PC patients and 126 patients with CP. Overall, 16 features were identified and used SVM for training to provide the sensitivity 96.25 ± 0.45%, specificity 93.38 ± 0.21%, positive predictive value (PPV) 92.21 ± 0.42%, and negative predictive value (NPV) of 96.68 ± 0.15% for PC.[12] In another retrospective study, three groups of patients who underwent EUS, normal pancreas ($n = 22$), CP ($n = 12$), and PC ($n = 22$), were identified on review of an endoscopic data set. A total of 110, 99, and 110 regions of interest (ROI) in normal pancreas, CP, and PC groups were analyzed to determine the accuracy of EUS-based images. Digital image analysis was applied with principal component analysis to create a neural network as a predictive model. The ANN model was very accurate in classifying ROI belonging to the PC group with sensitivity, specificity, PPV, and NPV of 93% (95% confidence interval [CI], 89–97%), 92% (95% CI, 88–96%), 87% (95% CI, 82–92%), and 96% (95% CI, 93–99%), respectively. The area under the characteristic operating curve (AUC) showed an accuracy of 0.93 in differentiating PC.[23]

Ozkan *et al.* also used a high-performance CAD system with image processing and pattern recognition with ANN with EUS images to detect PC. They used the ROI of EUS images of 202 PC and 130 non-cancer patients. They further divided them into three age groups of <40, 40–60, and >60 years. These images were tested via 200 random tests and provided the combined results of all three groups with accuracy, sensitivity, and specificity of 87.5%, 83.3%, and 93.3 %, respectively. The model was found to have the accuracy, sensitivity, and specificity of 92%, 87.5%, and 94.1%, respectively, for PC detection when age <40 years was also included. For the age group 40–60 years, accuracy, sensitivity, and specificity of 88.5%, 85.7%, and 94.7%, respectively, and for >60 years, accuracy, sensitivity, and specificity of 91.7%, 93.3%, and 88.9%, respectively, for PC detection. Therefore, the authors concluded that considering the patients' age could provide better results while using CAD to differentiate PC.[14]

EUS elastography is a newer diagnostic technique for differentiating pancreatic masses. Historically, EUS elastography has been used for differentiating between benign and malignant lymph nodes due to the ability to assess the differences in the hardness of normal and pathological tissues. However, it has not taken off in the United States, unlike Europe and Asia, mainly due to the

Table 1. Application of Artificial Intelligence in diagnosis of Pancreatobiliary diseases.

Author, year, and reference	Diagnostic modality	AI system	Data set	Results	Conclusion	Limitation
Artificial intelligence in endoscopic ultrasound						
Săftoiu *et al.*[10]	EUS elastography	ENN model	22 normal pancreas, 11 chronic pancreatitis, 32 pancreatic adenocarcinoma, 3 pancreatic neuroendocrine tumors	The sensitivity, specificity, and accuracy of differentiation between benign and malignant masses were 91.4%, 87.9%, and 89.7%, respectively. The PPV and NPV were 88.9% and 90.6%, respectively. Multilayer neural network system with high training had 97% accuracy.	Artificial neural network processing of EUS elastography enabled an optimal prediction of the different types of pancreatic lesions.	A lack of surgical standard in all cases.
Săftoiu *et al.*[11]	EUS elastography	ANN model	47 chronic pancreatitis 211 pancreatic adenocarcinoma	The neural computing approach had 91.14% training accuracy and 84.27% testing accuracy. Test showed sensitivity of 87.59%, a specificity of 82.94%, a PPV of 96.25%, and an NPV of 57.22%. AUC was 0.94.	Artificial intelligence methodology via artificial neural networks supports the medical decision process, providing fast and accurate diagnoses.	The study sample was unbalanced and small.
Zhu *et al.*[12]	EUS	SVM model	262 PC patients 126 CP patients	The total cases were randomly divided into a training set and a testing set. After 200 trials of randomized experiments, the average accuracy, sensitivity, specificity, the positive, and negative predictive values of pancreatic cancer were 94%, 96%, 93%, 92% and 96%, respectively.	Showed digital image processing and computer-aided EUS image differentiation technologies are highly accurate and noninvasive.	Small sample size
Săftoiu *et al.*[13]	Contrast-enhanced harmonic EUS	ANN model	112 cases of PC 55 cases of CP	For the ANN, sensitivity was 94.64%, specificity 94.44%, PPV 97.24%, and NPV 89.47%.	Parameters obtained through TIC analysis can differentiate between PC and CP cases and can be used in an automated CAD system with good diagnostic results	Only PC and CP lesions were included.

(Continued)

Table 1. (Continued)

Author, year, and reference	Diagnostic modality	AI system	Data set	Results	Conclusion	Limitation
Ozkan et al.[14]	EUS	ANN model	202 PC patients and 130 non-cancer patients divided into <40, 40–60, and >60 years age group.	Images classified under three age groups (in years; <40, 40–60 and >60) were tested via 200 random tests with accuracy: 92%, 88.5%, and 91.7%, respectively; sensitivity: 87.5%, 85.7%, and 93.3%, respectively; and specificity: 94.1%, 91.7%, and 88.9%, respectively. All groups together have accuracy: 87.5%, sensitivity: 83.3%, and specificity: 93.3%.	PC was better diagnosed with CAD system with age classification compared to without age classification.	Lack of comparison with other pancreatic diseases such as chronic pancreatitis, pancreatic pseudocysts
Kuwahara et al.[15]	EUS	DL method	3970 EUS still images were collected	The sensitivity, specificity, and accuracy of AI malignant probability were 95.7%, 92.6%, and 94.0%, respectively. The area under ROC was 0.98.	AI may be a more accurate and objective method to diagnose malignancies of IPMNs in comparison to human diagnosis and conventional EUS features.	A small study sample and only surgical sample were used.
Artificial intelligence in diagnostic and therapeutic ERCP						
Golub et al.[7]	ERCP	Back-propagation NN	Training set: 140 Testing set: 16	The trained network could predict CBDS in 100% of the patients in both the training and test sets.	Screening of high-risk patients for CBDS by neural network analysis is highly accurate. This promising new, noninvasive, and inexpensive technique can potentially decrease the need for preoperative ERCP by 50%	Authors recognized a chance of false-negative (up to 2%).
Yeaton et al.[16]	ERCP	DT method	Training set: 34 Testing set: 15	Pancreatic adenocarcinoma was diagnosed in the training data set of 34 patients during a leave-one-out process with an estimated sensitivity of 91% and specificity of 87%. Both sensitivity and specificity were 80% in the independent test set of 15 patients.	Inflammatory and malignant pancreatic epithelia exhibit distinct morphological features distinguished by decision tree-based classifiers employing image-cytometric numerical data.	Small sample size

(Continued)

Table 1. (Continued)

Author, year, and reference	Diagnostic modality	AI system	Data set	Results	Conclusion	Limitation
Jovanovic *et al.*[17]	ERCP	ANN model	291 patients for ERCP	There were 80.4% of patients with positive findings on ERCP. The AUC with the multivariate logistic regression model was 0.787, whereas the AUC with the ANN model was 0.884.	An ANN model has better discriminant ability and accuracy than a multivariate logistic regression model in selecting patients for therapeutic ERCP	Only those variables believed to be related to the outcome of interest were included. The majority of patients in our sample had positive findings on ERCP.

Artificial intelligence in non-endoscopic diagnosis and treatment of pancreatic cancer

Author, year, and reference	Diagnostic modality	AI system	Data set	Results	Conclusion	Limitation
Kurita *et al.*[8]	Cystic fluid analysis	DL	85 patient samples were used.	AUC for the diagnostic ability of malignant cystic lesions were 0.719 (CEA), 0.739 (cytology), and 0.966 (AI). In the diagnostic ability of malignant cystic lesions, sensitivity, specificity, and accuracy of AI were 95.7%, 91.9%, and 92.9%, respectively. AI sensitivity and accuracy was higher than that of CEA (60.9%, p = 0.021, 71.8%, p < 0.001) and cytology (47.8%, p = 0.001, 85.9%, p = 0.210) respectively.	AI may improve the diagnostic ability in differentiating malignant from benign pancreatic cystic lesions.	The study sample was too small.
Liu *et al.*[18]	CT	Faster-R CNN	Training set: 4385 CT images from 238 PC patients Clinical verification set: 1699 images from 100 PC patients	The mean average precision was 0.7664. Sequential contrast-enhanced CT images of 100 PC patients were used for clinical verification. The area under the ROC curve was 0.9632. It took approximately 0.2 s for the Faster R-CNN AI compared to an imaging specialist.	Faster R-CNN AI is an effective and objective method with high accuracy for the diagnosis of pancreatic cancer.	It is a retrospective study no control group.

(Continued)

Table 1. (Continued)

Author, year, and reference	Diagnostic modality	AI system	Data set	Results	Conclusion	Limitation
Wei et al.[19]	Multi-detector row CT	SVM model	Cross-validation cohort: 200 patients independent validation cohort: 60 patients	The diagnostic scheme in cross-validation showed the AUC = 0.767, sensitivity = 0.686, and specificity = 0.709. In the independent validation cohort, we acquired similar results with AUC = 0.837, sensitivity = 0.667, and specificity = 0.818.	The proposed radiomics-based computer-aided diagnosis scheme could increase preoperative diagnostic accuracy and assist clinicians in making accurate management decisions	The study sample was too small and needs enhanced classification accuracy of the tumors.
Artificial intelligence in cholangiocarcinoma						
Logeswaran et al.[20]	MRCP	MLP	Testing set:593 validation set: 55	The test results achieved was 94% when differentiating only healthy and tumor images, and 88% in a robust multi-biliary disease test.	This system uses MLP to perform automated preliminary detection of cholangiocarcinoma in 2D MRCP images.	Only a single MRCP image was selected and criteria need to be worked upon.
Shao et al.[21]	Early occlusion of bilateral plastic stent placement	BP-ANN model	Training set: 230 patients Validation set: 58 patients	In the training cohort, BP-ANN had larger AUC than the multivariate logistic regression model ($p = 0.00049$). In the internal testing cohort, the AUC of the BP-ANN had larger AUC than the multivariate logistic regression model ($p = 0.02142$).	AI can be a helpful tool for prediction of early occlusion of bilateral stent placement for inoperable hilar cholangiocarcinoma	There exists a data barrier in terms of mathematical predictions that needs to be considered.
Zhou et al.[22]	MBO	ANN model	Training set: 182 Validation set: 61	The c-index values showed good predictive performance in the training and validation cohorts (0.792 and 0.802, respectively) with independent risk factors. The optimum cut-off value of risk was 0.25.	AI can help in early and accurate prediction of EBI in patients with MBO who underwent PTBS.	Patient sample was non-randomized and had potential selection bias.

ANN, artificial neural network; AUC, area under the receptor; BP-ANN, back-propagation artificial neural network; CAD, computer-aided diagnosis; CP, chronic pancreatitis; DL, deep learning; DT, decision tree; EBI, early biliary infection; ENN, extended neural network; EUS, endoscopic ultrasound; IPMN, intraductal papillary mucinous neoplasm; MBO, malignant biliary obstruction; MLP, multilayer perceptron; MRCP, magnetic retrograde cholangiopancreatography; PC, pancreatic cancer; ROC, receiver operating curve; SVM, support vector machine; TIC, time-intensity curve; R-CNN, region-based convoluted neural network; PTBS, percutaneous transhepatic biliary stent.

difficulty in interpretation of the elastographic images compared to conventional ones. It has also been shown to have better results than traditional EUS and EUS-FNA, which is more invasive.[10,24,25] Saftiou *et al.* studied the application of extended neural network analysis with EUS elastography images to differentiate between normal pancreatic tissue, CP, and pancreatic adenocarcinoma (PAC). In this prospective study, 68 patients were included with normal pancreas (*n* = 22), PAC (*n* = 32), CP (*n* = 11), and pancreatic neuroendocrine tumors (*n* = 3). The study showed an accuracy of 89.7%, the sensitivity of 91.4%, specificity of 87.9%, PPV of 88.9%, and NPV of 90.6% for differentiation of benign (normal pancreas and CP) and malignant (PAC and neuroendocrine tumors) tissue based on hue histogram classification.[10] With the promising results of EUS elastography, a prospective, blinded, and the multicentric study was conducted with 258 patients with CP (*n* = 47) and PAC (*n* = 211). Images were analyzed with extended neural network analysis after performing EUS elastography. The neural network analysis showed 91.14% training accuracy and 84.27% testing accuracy. The analysis had corresponding sensitivity, specificity, PPV, and NPV of 87.59%, 82.94%, 96.25%, and 57.22%, respectively. Also, the AUC was 0.94, whereas the mean hue histogram analysis had an AUC of 0.85.[11]

Contrast-enhanced EUS and contrast-enhanced harmonic EUS (CEH-EUS) are the new techniques in the field of EUS, which help to better delineate between different focal pancreatic masses and also differentiate between CP and PC.[13,26] A multi-center observational study was conducted in 2015 comparing EUS-FNA and CEH-EUS and the use of parameters from the time-sensitivity curve (TIC) analysis in ANN processing. A total of 167 cases with 112 cases of PC and 55 cases of CP were analyzed. The sensitivity, specificity, PPV and NPV were 84.82% vs 87.5% vs 94.64%, 100% vs 92.72% vs 94.44%, 100% vs 96.07% vs 97.24%, 76.63% vs 78.46% vs 89.47% for EUS-FNA vs CEH-EUS vs with TIC analysis in ANN system, respectively. This study showed that the TIC analysis used and CAD could further enhance the diagnostic value of EUS-FNA and CEH-EUS.[13]

A single-center, retrospective study was conducted to analyze whether preoperative AI can recognize data patterns to predict malignancy using EUS images of intraductal papillary mucinous neoplasms (IPMNs). A total of 206 patients with 3970 images who had EUS before pancreatectomy for confirmed IPMNs were included in the study. EUS images of IPMNs were provided to DL AI algorithm to predict malignancy. The mean AI value of malignant IPMNs was statistically significantly higher than benign IPMNs (0.808 vs 0.104, *p* < 0.001). The area under the receiver operating curve (ROC) for the ability to diagnose malignancy of IPMNs with AI malignant probability was 0.98 (*p* < 0.001). The diagnostic ability to predict malignant IPMN was much higher for AI with an accuracy of 94% when compared to endoscopists' diagnosis with an accuracy of 56.0%. This study recommends the use of AI preoperatively in patients with IPMN to assess malignancy potential. The results of this study need to be validated further with prospective multicentric studies as it was a small retrospective study.[15]

A recent study by Zhang *et al.* constructed a DL system, BP MASTER (pancreaticobiliary master), for EUS training and quality control as although EUS is one of the most sensitive methods to detect pancreatic lesions, but it is still operator dependent and requires unique skills. In this study, the standard procedure of pancreatic EUS was divided into six stations according to the Japan Gastroenterological Endoscopy Society and European and American experts' experience. Station classification model and a pancreas/abdominal aorta/portal confluence segmentation model developed using 19,486 and 2207 EUS images, respectively. Similarly, classification and segmentation internal validation developed using 1920 and 700 images, respectively, from 311 eligible EUS videos collected retrospectively. From two other hospitals, 768 images were collected for external validation. In addition, an independent data set of 180 images was applied to compare the model and the EUS expert's performance. This model achieved an accuracy of 94.2%, 82.4%, and 86.2% in station classification at internal validation, external validation, and video test, respectively. Both the BP MASTER model and EUS expert achieved equal accuracy of 90% in classification. Similarly, dice in pancreatic and blood vessel segmentation were comparable in both groups with 0.77, 0.813 for the model, and 0.706 and 0.813 for the expert group. With augmentation, trainees' station recognition accuracy improved 67.2–78.4% (*p* value = 0.02) in the cross-over study. This study showed that BP MASTER, a DL system, can help EUS training by improving the accuracy of detecting PC.[27]

AI in diagnostic and therapeutic ERCP

ERCP is an important tool to access the bile ducts for diagnostic and therapeutic purposes.[28] However, due to its invasive nature and universal unavailability, modalities such as magnetic retrograde cholangiopancreatography (MRCP) are used for diagnostic uses. Due to these reasons, methods to reliably predict the need for ERCP are actively being sought. Conventional neural network (CNN) technologies could be used in these conditions to assess the need for ERCP.[7,16]

In a study of 49 patients who underwent ERCP for patients with differential diagnosis of PAC and CP, brush cytology samples with image cytometry were collected to understand the morphological differences in PAC and CP. The AI data classification methods, decision tree, and the production rule system were used to perform data analysis. Patients were divided into two data sets, one data set of 34 patients with a definitive diagnosis labeled as a training set, and another independent data set of 15 patients labeled as the test set. A classification system that can distinguish between PAC and CP was created using the training set with the decision tree method. PAC was diagnosed with sensitivity and specificity of 91% vs 80% and 87% vs 80% in the training and test sets, respectively. These results showed that although chronic inflammation in CP can complicate PAC diagnosis, benign and malignant pancreatic diseases have distinct morphological differences that can be used by AI classifiers like a decision tree and a production rule system to help in cancer diagnosis.[16]

ERCP plays a vital role in preoperative diagnosis and management (with endoscopic sphincterotomy (ES)) of high-risk patients with choledocholithiasis (CDL) due to technical difficulties and often lack of expertise with the laparoscopic exploration of the common bile duct in these patients. A single-center, prospective observational study was conducted on 291 patients referred for therapeutic ERCP for suspected CDL to evaluate if ANN will help predict the need for therapeutic ERCP for CDL patients. The diagnostic performance of ANN for CDL was compared with the previously reported prediction model by the multivariate logistic regression model. AUC for the multivariate logistic regression model was 0.787 (95% CI, 0.720–0.854; $p < 0.001$). When the performance of ANN was compared to the multivariate logistic regression using pairwise AUC, AUC for multivariate logistic regression was 0.752

(95% CI, 0.593–0.847) with a mean difference between both AUCs as 0.137 (95% CI, −0.012 to 0.276; $p = .07$). The ANN model correctly classified 92.3% of patients with positive findings on ERCP and 69.6% of patients with negative findings on ERCP, showing the superiority of the ANN model in predicting CDL.[17]

A retrospective study was performed to determine whether the use of ANN can predict suspected CDL. A total of 156 patients (140 patients in the training set and 16 in the testing set) who had cholecystectomy with either a preoperative ERCP or an intraoperative cholangiogram were included. The back-propagation neural network was trained using 140 patients in the training set and then applied to the test set. It was able to predict CDL incidence in 100% of patients in both the training and test set of 16 patients. Thus, back-propagation NNA was found to be highly accurate in predicting CDL to reduce the need for invasive ERCP, which is futile in half of the suspicious patients.[7]

In addition, the role of AI technologies like natural language processing (NLP) has been studied to measure ERCP quality metrics. In a single-center study from 2006 to 2014, ERCP from six providers were analyzed with NLP for quality measures based on 2014 ASGE/ACG quality indicators for ERCP. It showed NLP precision between 84% and 100% and accuracy between 90% and 100%. This study showed that NLP and data mining could be used for tracking ERCP quality measures. These results are based on a single-center study, so further multi-center studies are needed to validate these results.[29]

AI in non-endoscopic diagnosis and treatment of PC

Due to its retroperitoneal positioning, the best diagnostic model is a dual-phase or triple-phase intravenous contrast helical CT of the abdomen to diagnose and stage cancer.[18,19,30]

CT scan

Most of the pancreatic serous cystic neoplasms (SCNs) are indolent and behave almost as benign tumors compared to other pancreatic cystic neoplasms (PCNs). SCNs usually do not need further follow-up or are managed conservatively with serial imaging. SCNs are found to be misdiagnosed as cancerous, with imaging leading to

surgery instead of conservative management and vice versa. Surgical resection is the treatment of choice for other PCNs like IPMNs, mucinous cystic neoplasms (MCNs), and solid pseudopapillary neoplasms (SPNs). A retrospective study was conducted with the use of a radiomics-based CAD system to improve preoperative differentiation of pancreatic SCNs from other PCNs. A total of 260 patients who underwent resection of PCNs were included in this study. Out of 260 patients, 200 patients were included in the cross-validation cohort and 60 in the independent validation cohort. All patients underwent abdominal multi-detector-row CT (MDCT), which were read by an experienced radiologist. The design of a radiomics system contained 24 guideline-based features and 385 radiomics high-throughput features. Clinicians were able to identify only 31 out of 102 SCNs correctly before surgery. Radiomics-based CAD scheme improved the diagnostic accuracy, with AUC increased from 0.707 to 0.767 and 0.774 to 0.837 in the cross-validation and independent validation, respectively.[19]

Liu *et al.* conducted a study to establish an AI diagnosis system with a faster region-based convolutional neural network (Faster R-CNN), which could read CT scans correctly to quickly diagnose PC quickly. A total of 338 patients were included, and the AI system was composed of two data sets: training and verification. Patients' characteristics such as sex, age, tumor location, tumor-node-metastasis stage, and differentiation grade between training and verification groups were not significant. A total of 4385 CT images of 238 PC patients were included in the training process and applied to a faster region-based convolution network (Faster R-CNN) model, and 1699 images from the remaining 100 PC patients were included in the verification process. The training set of Faster R-CNN had a mean average precision of 0.7664, and AUC with a trapezoidal rule of 0.9632, reflecting the accuracy of results. Time taken by Faster R-CNN to analyze an image was 3 s compared to 8 min taken by imaging specialists. This was a retrospective single-center study, so further prospective studies are needed.[18]

Cystic fluid analysis

Pancreatic cystic fluid can be analyzed for different markers like carcinoembryonic antigen (CEA) and amylase/lipase levels. These fluid analysis markers can be used to differentiate between mucinous and non-mucinous pancreatic lesions

but not benign versus malignant cystic lesions.[31–33] Diagnostic accuracy of cyst fluid analysis, cytology, and AI combining cystic fluid analysis and clinical data to differentiate between benign and malignant pancreatic lesions were tested in a retrospective analysis. A total of 85 patients with pancreatic cystic lesions were included in this study who underwent cyst fluid analysis, collected either surgically or by EUS-FNA. The area under ROC for the diagnostic ability of malignant from benign pancreatic cystic lesion were 0.739 ($p = 0.001$), 0.719 ($p = 0.002$), 0.956 ($p < 0.001$ and 0.966 ($p < 0.001$) for cytology, CEA, AI using only CEA and AI, respectively. Similarly, AI sensitivity (95.7%) and accuracy (92.9%) were higher compared to the cyst diagnosis of EUS by cytology, CEA, and AI using only CEA. This study showed that AI using DL could improve diagnostic accuracy to differentiate benign and malignant cystic pancreatic lesions. It was a small sample retrospective study, so larger studies are needed, especially with a new AI algorithm that can analyze data with the missing item.[8]

Radiation therapy

Another treatment modality for PC is radiation therapy. Given that pancreas is soft tissue, image guidance strategies for accurate delivery of plan dose distribution like kV x-ray imaging cannot localize the target area accurately. In these cases, metallic fiducials are often implanted through EUS or by interventional radiologist or gastroenterologists into tumor target or adjacent normal tissue. Fiducials implantation is an invasive procedure and sometimes interferes with treatment by causing metal artifact or obscuring target tissue. They can also cause pancreatitis and infection leading to delay in treatment. Zhao *et al.* applied the DL method to image guidance strategies to recognize pancreatic tumor targets for image-guided radiation therapy. A DL model was trained using planning CT image data sets to localize the planning target volume for pancreatic image-guided radiation therapy. Once model training was completed, the accuracy of the model was evaluated retrospectively by studying the patients who underwent PC radiation therapy. Mean absolute differences (MADs) between the actual positions and the model predicted were found to be less than 2.60 mm in three different planes—anteroposterior, lateral, and oblique directions for both the axes in the detector plane. For comparison studies with and without reference markers, MADs were less than 2.49 mm.[34]

AI in cholangiocarcinoma

MRCP is a diagnostic test of choice for cholangiocarcinoma. CAD systems for MRCP in cholangiocarcinoma have not been studied extensively, especially due to its rarity and the many variations in the MRCP images. Furthermore, surrounding tissue structures and organs further affects the quality of images needed to develop the CAD system.[20] Logeswaran *et al.* presented a CAD system, the multi-layer perceptron (MLP), for automated preliminary detection of cholangiocarcinoma using a single MRCP image. MLP, a form of ANN, was used to differentiate images with and without cholangiocarcinoma. The multi-layer system employed algorithms that corresponded to the radiological diagnosis characteristics like a radiologist.[20] MLP was compared to previously available automated tumor detection systems using MRCP, a region growing scheme.[35] When compared to this reference system, MLP was found to be better in the detection of cholangiocarcinoma (88.03% vs 86.17%), healthy tissue (83.64% vs 76.90%), and non-cholangiocarcinoma disease (90.14% vs 80.99%). This tool is proposed to aid in diagnosing cholangiocarcinoma using single MRCP images to improve the detection rate with the system analyzing multiple images.[20]

Hilar cholangiocarcinoma (HCA) is most commonly associated with malignant biliary obstruction (MBO). More than half of the cases are detected when they need biliary drainage or stent placement, either surgical, percutaneous, or endoscopy guided. A retrospective single-center study of 288 patients with bilateral plastic stent placement for inoperable HCA of bismuth type 2 to study whether back-propagation ANN (BP-ANN) can accurately predict the early occlusion of bilateral plastic stent placement in these patients. Patients were randomly assigned to either a training cohort ($n = 231$) and internal testing cohort ($n = 57$). Results from BP-ANN were compared to multivariate logistic regression analysis. Results showed that early stent occlusion was significantly associated with cancer staging ($p = 0.002$), the diameter of the stent ($p = 0.012$), and the bismuth stage ($p = 0.017$). The AUC for the BP-ANN (0.8763 (95% CI, 0.8164–0.9364)) was larger than the multivariate logistic regression model 0.9648 (95% CI, 0.9375–0.9922; $p = 0.00049$) in the training cohort. Similarly, the AUC of the BP-ANN in the internal testing cohort was larger than the multivariate logistic regression model ($p = 0.02142$). This study showed that the BP-ANN is a better predictor model than a multivariate logistic regression model.[21] This predictive model will help plan future ERCP for stent exchange at an appropriate time prior to developing cholangitis.

Percutaneous transhepatic biliary stent (PTBS) is a palliative treatment for MBO, which can be due to PC, metastatic lesion, or obstruction of the biliary tract due to malignancy. The most common complication associated with it is early biliary infection (EBI). A multi-center retrospective study with 243 MBO patients treated with PTBS (training cohort = 182; validation cohort = 61) conducted to study risk factors associated with EBI and establish a "Risk Prediction for Early Biliary Infection" (RPEBI) nomogram for these patients. Results were analyzed using multivariate logistic regression analysis. It showed that length of obstruction (odds ratio (OR), 1.061; 95% CI, 1.013–1.111; $p = 0.012$), diabetes (OR, 5.070; 95% CI, 1.917–13.412; $p = 0.001$), location of obstruction (OR, 2.283; 95% CI, 1.012–5.149; $p = 0.047$), and previous surgical or endoscopic intervention (OR, 3.968; 95% CI, 1.709–9.217; $p = 0.001$) are independent risk factor for EBI in patients with MBO after PTBS. Based on this data, the RPEBI nomogram was developed to evaluate the risk for EBI in MBO patients who had PTBS. The internal and external validation showed a c-index of 0.792 in the training cohort and 0.802 in the validation cohort, translating to good predictive performance. According to the maximized Youden index, the cut-off value for risk was found to be 0.25, so patients with a risk score value (0.25) should have close monitoring and prophylactic antibiotics.[22]

AI prospects in pancreatobiliary diseases

The application of AI in pancreatobiliary diseases and the field of endoscopy as a whole hold a bright future. Ever-expanding imaging databases and digital information continue to expand, which could be used for AI. We believe that several other opportunities will significantly improve the incorporation of AI. AI has started to find its way to technical integration with electronic medical records and endoscopy platforms. It will optimize clinical workflow while improving the detection, treatment, and prognosis of cancerous lesions using the various machine and DL techniques for image characterization, analysis, and risk stratification. Moreover, AI training for endoscopists and its seamless integration in the endoscopy suite is important. Future research should include extensive prospective studies with

comprehensive data involving a multidisciplinary team of physicians and data professionals to develop AI methods to aid in the diagnosis, clinical decision-making, and risk stratification, thus improving the patient's outcomes.

Conclusion

Over the last decade, significant progress has been made on AI and its applications in the gastrointestinal field, especially for pancreaticobiliary disease. AI-based methods can assist physicians in making quicker and more accurate decisions. In conjunction with various endoscopic and non-endoscopic methods, AI and CAD systems are showing promising results in the diagnosis of PC. Multiple studies showed that ANN, an advanced AI technique, could assist in diagnosing and staging cholangiocarcinoma and predicting CDL, thus avoiding diagnostic ERCP. AI is making a rapid shift in our approach to medicine. More research is needed to identify areas where AI-based methods can further enhance diagnostic, therapeutic, and prognostic models and enhance patient care.

Author contributions

H.G. and R.M. contributed in conception and design of the study. H.G., R.M., Z.G., and A.P. contributed in literature search. R.M. and Z.G. produced the first draft. All authors contributed in critical revision and editing of the manuscript. All authors gave final approval for the article.

Conflict of interest statement

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Ethical approval

This manuscript, as submitted or its essence in another version is not under consideration for publication elsewhere and will not be published elsewhere while under review by *Therapeutic Advances in Gastrointestinal Endoscopy*. All authors have made substantive contributions to the study and endorse the data and conclusions.

ORCID iD

Hemant Goyal [iD] https://orcid.org/0000-0002-9433-9042

References

1. Ebigbo A, Palm C, Probst A, *et al.* A technical review of artificial intelligence as applied to gastrointestinal endoscopy: clarifying the terminology. *Endosc Int Open* 2019; 7: E1616–E1623.

2. Russell S and Norvig P. *Artificial intelligence: a modern approach*. Upper Saddle River, NJ: Prentice Hall Press, 2009.

3. Ruffle JK, Farmer AD and Aziz Q. Artificial intelligence-assisted gastroenterology- promises and pitfalls. *Am J Gastroenterol* 2019; 114: 422–428.

4. Hamet P and Tremblay J. Artificial intelligence in medicine. *Metabolism* 2017; 69S: S36–S40.

5. Yang YJ and Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019; 25: 1666–1683.

6. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw* 2015; 61: 85–117.

7. Golub R, Cantu R Jr and Tan M. The prediction of common bile duct stones using a neural network. *J Am Coll Surg* 1998; 187: 584–590.

8. Kurita Y, Kuwahara T, Hara K, *et al.* Diagnostic ability of artificial intelligence using deep learning analysis of cyst fluid in differentiating malignant from benign pancreatic cystic lesions. *Sci Rep* 2019; 9: 6893.

9. Le Berre C, Sandborn WJ, Aridhi S, *et al.* Application of artificial intelligence to gastroenterology and hepatology. *Gastroenterology* 2020; 158: 76–94.

10. Săftoiu A, Vilmann P, Gorunescu F, *et al.* Neural network analysis of dynamic sequences of EUS elastography used for the differential diagnosis of chronic pancreatitis and pancreatic cancer. *Gastrointest Endosc* 2008; 68: 1086–1094.

11. Săftoiu A, Vilmann P, Gorunescu F, *et al.* Efficacy of an artificial neural network- based approach to endoscopic ultrasound elastography in diagnosis of focal pancreatic masses. *Clin Gastroenterol Hepatol* 2011; 10: 84.e1–90.e1.

12. Zhu M, Xu C, Yu J, *et al.* Differentiation of pancreatic cancer and chronic pancreatitis using computer-aided diagnosis of endoscopic ultrasound (EUS) images: a diagnostic test. *PLoS ONE* 2013; 8: e63820.

13. Săftoiu A, Vilmann P, Dietrich CF, *et al.* Quantitative contrast-enhanced harmonic EUS in differential diagnosis of focal pancreatic masses (with videos). *Gastrointest Endosc* 2015; 82: 59–69.

14. Ozkan M, Cakiroglu M, Kocaman O, *et al.* Age-based computer-aided diagnosis approach for pancreatic cancer on endoscopic ultrasound images. *Endosc Ultrasound* 2016; 5: 101–107.

15. Kuwahara T, Hara K, Mizuno N, *et al.* Usefulness of deep learning analysis for the diagnosis of malignancy in intraductal papillary mucinous neoplasms of the pancreas. *Clin Transl Gastroenterol* 2019; 10: 1–8.

16. Yeaton P, Sears RJ, Ledent T, *et al.* Discrimination between chronic pancreatitis and pancreatic adenocarcinoma using artificial intelligence-related algorithms based on image cytometry-generated variables. *Cytometry* 1998; 32: 309–316.

17. Jovanovic P, Salkic NN and Zerem E. Artificial neural network predicts the need for therapeutic ERCP in patients with suspected choledocholithiasis. *Gastrointest Endosc* 2014; 80: 260–268.

18. Liu SL, Li S, Guo YT, *et al.* Establishment and application of an artificial intelligence diagnosis system for pancreatic cancer with a faster region-based convolutional neural network. *Chin Med J (Engl)* 2019; 132: 2795–2803.

19. Wei R, Lin K, Yan W, *et al.* Computer-aided diagnosis of pancreas serous cystic neoplasms: a radiomics method on preoperative MDCT images. *Technol Cancer Res Treat* 2019; 18: 1533033818824339.

20. Logeswaran R. Cholangiocarcinoma—an automated preliminary detection system using MLP. *J Med Syst* 2009; 33: 413–421.

21. Shao F, Huang Q, Wang C, *et al.* Artificial neural networking model for the prediction of early occlusion of bilateral plastic stent placement for inoperable hilar cholangiocarcinoma. *Surg Laparosc Endosc Percutan Tech* 2018; 28: e54–e58.

22. Zhou HF, Huang M, Ji JS, *et al.* Risk prediction for early biliary infection after percutaneous transhepatic biliary stent placement in malignant biliary obstruction. *J Vasc Interv Radiol* 2019; 30: 1233–1241.

23. Das A, Nguyen CC, Li F, *et al.* Digital image analysis of EUS images accurately differentiates pancreatic cancer from chronic pancreatitis and normal tissue. *Gastrointest Endosc* 2008; 67: 861–867.

24. Săftoiu A, Vilmann P, Hassan H, *et al.* Analysis of endoscopic ultrasound elastography used for characterisation and differentiation of benign and malignant lymph nodes. *Ultraschall Med* 2006; 27: 535–542.

25. Giovannini M, Hookey LC, Bories E, *et al.* Endoscopic ultrasound elastography: the first step towards virtual biopsy? Preliminary results in 49 patients. *Endoscopy* 2006; 38: 344–348.

26. Seicean A, Badea R, Stan-Iuga R, *et al.* The added value of real-time harmonics contrast-enhanced endoscopic ultrasonography for the characterisation of pancreatic diseases in routine practice. *J Gastrointestin Liver Dis* 2010; 19: 99–104.

27. Zhang J, Zhu L, Yao L, *et al.* Deep learning–based pancreas segmentation and station recognition system in EUS: development and validation of a useful training tool (with video). *Gastrointest Endosc* 2020; 92: 874–885.

28. Johnson KD, Perisetti A, Tharian B, *et al.* Endoscopic retrograde cholangiopancreatography-related complications and their management strategies: a "scoping" literature review. *Dig Dis Sci* 2020; 65: 361–375.

29. Imler TD, Sherman S, Imperiale TF, *et al.* Provider-specific quality measurement for ERCP using natural language processing. *Gastrointest Endosc* 2018; 87: 164–173.

30. Zhu L, Dai MH, Wang ST, *et al.* Multiple solid pancreatic lesions: prevalence and features of non-malignancies on dynamic enhanced CT. *Eur J Radiol* 2018; 105: 8–14.

31. Nagashio Y, Hijioka S, Mizuno N, *et al.* Combination of cyst fluid CEA and CA 125 is an accurate diagnostic tool for differentiating mucinous cystic neoplasms from intraductal papillary mucinous neoplasms. *Pancreatology* 2014; 14: 503–509.

32. Park WG, Mascarenhas R, Palaez-Luna M, *et al.* Diagnostic performance of cyst fluid carcinoembryonic antigen and amylase in histologically confirmed pancreatic cysts. *Pancreas* 2011; 40: 42–45.

33. Ngamruengphong S and Lennon AM. Analysis of pancreatic cyst fluid. *Surg Pathol Clin* 2016; 9: 677–684.

34. Zhao W, Shen L, Han B, *et al.* Markerless pancreatic tumor target localization enabled by deep learning. *Int J Radiat Oncol Biol Phys* 2019; 105: 432–439.

35. Logeswaran R and Eswaran C. Discontinuous region growing scheme for preliminary detection of tumor in MRCP images. *J Med Syst* 2006; 30: 317–324.