


# Application of artificial intelligence in diagnosis of pancreatic malignancies by endoscopic ultrasound: a systemic review

Hemant Goyal<sup>\*</sup> , Syed Ali Amir Sherazi<sup>\*</sup>, Shweta Gupta, Abhilash Periseti, Ikechukwu Achebe, Aman Ali, Benjamin Tharian, Nirav Thosani and Neil R. Sharma

## Abstract

**Background:** Pancreatic cancer (PC) is a highly fatal malignancy with a global overall 5-year survival of under 10%. Screening of PC is not recommended outside of clinical trials. Endoscopic ultrasonography (EUS) is a very sensitive test to identify PC but lacks specificity and is operator-dependent, especially in the presence of chronic pancreatitis (CP). Artificial Intelligence (AI) is a growing field with a wide range of applications to augment the currently available modalities. This study was undertaken to study the effectiveness of AI with EUS in the diagnosis of PC.

**Methods:** Studies from MEDLINE and EMBASE databases reporting the AI performance applied to EUS imaging for recognizing PC. Data were analyzed using descriptive statistics. The Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool was used to assess the quality of the included studies.

**Results:** A total of 11 articles reported the role of EUS in the diagnosis of PC. The overall accuracy, sensitivity, and specificity of AI in recognizing PC were 80–97.5%, 83–100%, and 50–99%, respectively, with corresponding positive predictive value (PPV) and negative predictive value (NPV) of 75–99% and 57–100%, respectively. Types of AI studied were artificial neural networks (ANNs), convolutional neural networks (CNN), and support vector machine (SVM). Seven studies using other than basic ANN reported a sensitivity and specificity of 88–96% and 83–94% to differentiate PC from CP. Two studies using SVM reported a 94–96% sensitivity, 93–99% specificity, and 94–98% accuracy to diagnose PC from CP. The reported sensitivity and specificity of detection of malignant from benign Intraductal Papillary Mucinous Neoplasms (IPMNs) was 96% and 92%, respectively.

**Conclusion:** AI reported a high sensitivity with high specificity and accuracy to diagnose PC, differentiate PC from CP, and differentiate benign from malignant IPMN when used with EUS.

**Keywords:** artificial intelligence, artificial neural network, convolutional neural network, endoscopic ultrasonography, pancreatic cancer, support vector machine

Received: 30 July 2021; revised manuscript accepted: 1 March 2022.

## Introduction

Pancreatic cancer (PC) is one of the most fatal cancers globally, with a 5-year overall survival rate of 9% for all stages and only 3% for Stage IV disease. There has been an over 50% increase in incidence and mortality over the last 25 years, and the burden may double in the next four decades.<sup>1</sup>

It is the third leading cause of cancer mortality in the United States, with more than 45,000 deaths annually and is expected to become the second leading cause of cancer death. Despite these trends, there is insufficient evidence for current practice guidelines to recommend PC screening in asymptomatic individuals. However, certain

*Ther Adv Gastroenterol*

2022, Vol. 15: 1–11

DOI: 10.1177/  
17562848221093873

© The Author(s), 2022.  
Article reuse guidelines:  
sagepub.com/journals-  
permissions

Correspondence to:

**Hemant Goyal**  
The Wright Center  
for Graduate Medical  
Education, 501 S.  
Washington Avenue,  
Scranton, PA 18503, USA  
[goyalh@thewrightcenter.org](mailto:goyalh@thewrightcenter.org)

**Syed Ali Amir Sherazi**  
**Ikechukwu Achebe**  
Department of Medicine,  
John H. Stroger Jr.  
Hospital of Cook County,  
Chicago, IL, USA

**Shweta Gupta**  
Division of Hematology-  
Oncology, John H. Stroger  
Jr. Hospital of Cook  
County, Chicago, IL, USA

**Abhilash Periseti**  
Advanced Interventional  
Oncology and Surgical  
Endoscopy, Parkview  
Cancer Institute, Parkview  
Health, Fort Wayne, IN,  
USA

**Aman Ali**  
Commonwealth Medical  
College, Scranton, PA,  
USA; Wilkes Barre General  
Hospital, Wilkes-Barre,  
PA, USA; Digestive Care  
Associates, Kingston,  
PA, USA

**Benjamin Tharian**  
University of Arkansas for  
Medical Sciences, Little  
Rock, AR, USA

**Nirav Thosani**  
Center for Interventional  
Gastroenterology at  
UTHealth (iGUT), Houston,  
TX, USA

Division of  
Gastroenterology,  
Hepatology & Nutrition,  
McGovern Medical School,  
UTHealth, Houston, TX,  
USA

**Neil R. Sharma**

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

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

\*Hemant Goyal and Syed  
Ali Amir Sherazi share  
equal authorship

high-risk groups, for example, patients with germline mutations in BRCA1, BRCA2, TP53, and Lynch syndrome with mismatch repair genes, may benefit from screening if highly sensitive and specific non-invasive tests were available.<sup>2</sup>

Multiple modalities, such as CT scans, magnetic resonance imaging (MRI), and endoscopic ultrasound (EUS), are currently being used to diagnose PC. Among these, EUS is considered superior due to its ability to obtain high-quality images.<sup>3</sup> However, it has low sensitivity in differentiating benign from malignant Intraductal Papillary Mucinous Neoplasms (IPMNs).<sup>3</sup> In the presence of chronic pancreatitis (CP), EUS and endoscopic sonoelastography have a high sensitivity but a low specificity making the differentiation challenging, and cytologic diagnosis remains the gold standard.<sup>4</sup> Furthermore, EUS is operator-dependent, and less experienced endoscopists could miss the subtle difference between CP and PC because of the presence of concomitant scarring and calcification due to chronic inflammation.<sup>5</sup>

Artificial Intelligence (AI) is a term applied to any computer system that has been developed to learn and emulate the biological brain. Machine learning (ML) is a form of AI that uses large amounts of data to find various patterns from it (Table 1). There are three types of ML – supervised learning, unsupervised learning, and reinforced learning. Supervised learning has been studied and applied to medicine, especially in diagnostics. Two types of supervised learning methods have been explored in EUS, which are artificial neural networks (ANNs), also called neural networks (NN), and support vector machine (SVM).<sup>6,7</sup> Deep Learning (DL) is an advanced concept that stems from ANN and uses multiple complex layers of ANN, inspired by the neurons of the human brain. Convolutional neural networks (CNNs) are an example of a DL program that is based on visual signal processing by humans.<sup>7</sup> SVM is a type of supervised ML method where large amounts of data are inputted to generate two or more categories divided by boundaries. Although simpler and more generalizable than ANN, SVM takes longer to develop and requires a very large amount of data for ML.<sup>6,8</sup>

During the past decade, significant advances have been made on the application of AI in the diagnosis and management of gastrointestinal diseases.

Some of them include application of AI for increased accuracy and predicting survival in esophageal cancer, improved prognostic prediction in inflammatory bowel disease (IBD), prediction of nodal metastasis in early-stage colorectal cancer, and prognostic evaluation of PCs.<sup>9,10</sup> Hence in this article, we performed a systemic review of the current published literature to assess the application of AI programs for EUS-mediated recognition of primary pancreatic malignancies, including differentiation between PC and CP.

## Methods

### *Study selection*

We searched MEDLINE and EMBASE databases (inception to 10 April 2021) using keywords and/or Medical Subject Headings (MeSH) for ‘pancreatic cancer’, ‘artificial intelligence’, ‘computer-aided’, ‘neural networks’, ‘deep learning’, ‘machine learning’, ‘computer-assisted diagnosis’ and ‘support vector machine’. The detailed search strategy is listed in the Supplementary file. Two authors (S.G. and H.G.) independently reviewed the references and selected studies for full-text screening. An additional search was done by reference screening of the selected articles. Studies were deemed fit for inclusion if they reported the performance of any form of AI applied to EUS imaging for recognizing PC. Studies were excluded if they used AI to gauge human performance only or did not report diagnostic sensitivity and specificity of AI in recognition of pancreatic malignancies. Diagnostic odds ratio (DOR) was calculated from the available data.

### *Data extraction and quality assessment*

We extracted data regarding study design, patient population, type of AI used, reported outcomes, and limitations (Table 2). Data were extracted by one author (S.A.A.S.) and reviewed by a second author (S.G.). The Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool was used to assess the quality/risk of bias of studies included in this review.<sup>11</sup>

## Results

A total of 1669 studies were identified from the initial search of all databases and reference screening, of which 24 were selected for full-text review

**Table 1.** Various types of AI terminologies and their explanation.

Terminology	Explanation
ML	ML is a type of AI where a machine is taught to give output by processing input as an intelligent being would.
Supervised ML	Supervised ML uses known datasets created by humans to train a machine to make decisions within the defined parameters of the known dataset.
Unsupervised ML	In unsupervised ML, the machine is fed known datasets; however, it learns to find new patterns, previously unknown, to generate new output based on the newly identified patterns. It can self-improve without human input.
Reinforced ML	Reinforced ML uses known knowledge, such as supervised learning, combined with unknown input to generate an output at the time of an unknown encounter. It emulates the decision-making capacity of an intelligent being in uncharted territory with known knowledge to come up with the best action plan in the unknown scenario.
SVM	Type of supervised ML where a very large amount of data already trained with input and output is fed in the machine. The machine uses the data to create categories, and any subsequent data input is classified in those categories. It cannot self-learn to make any more categories without more trained input.
ANN	Type of supervised ML where the machine can identify more complex patterns based on input features of the data. Unlike SVM, which can create defined categories, ANN can emulate a biological brain to recognize intricate patterns to produce an output. However, it cannot learn unsupervised and develop new algorithms.
DL	DL is a type of ML that can be supervised, unsupervised, or reinforced. When combined with NNs, it can form sophisticated supervised learning algorithms but can also learn without human supervision to create the best output based on the data inputted. It can create new outputs that are not already defined.
CNN	CNN is a type of DL combined with ANN that emulates the visual cortex of the biological brain. The various visual inputs/images are processed by complex neuronal connections in the machine to create the best output. It is a type of supervised learning method but can also be programmed for unsupervised learning, which can learn and improve its output accuracy over time.

(after removing duplicates) and 11 were included in the final analysis (Figure 1). There were a total of 2292 patients across all studies consisting of 1409 pancreatic malignancies with 1383 patients with PC, 3 with pancreatic neuroendocrine tumors (PNET), and 27 with malignant IPMN.<sup>12–22</sup> Study characteristics are listed in detail in Table 2.

### Study design

Eight studies<sup>12–16,20–22</sup> were retrospective in nature using images (still or video) from already performed EUS procedures. Three studies<sup>17–19</sup> were conducted to collect images in real time to be fed into an AI system. Overall, 10 studies<sup>12,14–22</sup> assessed the performance of AI to recognize PC,

while one<sup>13</sup> studied the recognition of malignant IPMNs. All patients in all studies had a confirmed cytologic diagnosis of the condition studied, including CP and malignancy. Most studies lacked detailed data to create a  $2 \times 2$  contingency table; hence a formal meta-analysis could not be performed.

### EUS images used

Seven studies<sup>12,13,15,16,20–22</sup> used still EUS images, of which two studies used data augmentation to increase the number of images several folds to be fed in the AI system. Three studies<sup>17–19</sup> used video images, and one<sup>14</sup> used both still and video images.

**Table 2.** Study characteristics.

Author	Study design	AI system studied	Patient population	Outcomes of AI in detection of PC	Author conclusion	Limitations
Das et al. <sup>1,2</sup> USA	Retrospective review of EUS images divided in regions of interest (ROIs) groups from patients with NP, CP, and PC	ANN (MLP)	NP 22 patients, 110 ROIs CP 12 patients, 99 ROIs PC 22 patients, 110 ROIs (confirmed by EUS-FNA) Still images used	<i>PC recognition</i> Sn = 93% Sp = 92% PPV = 87% NPV = 9.6% AUC = 0.93 CP recognition from NP Sn = 100% Sp = 100%	Digital image analysis (DIA) of EUS images is accurate in differentiating PC from chronic inflammation and normal pancreas	1. EUS images acquired from echoendoscopes with fixed settings; impact on other equipment is unknown 2. ROI selection not blinded 3. DIA not real time, retrospective design 4. Single institution
Kuwahara et al. <sup>13</sup> Japan	Retrospective review of EUS images of IPMN patients to differentiate benign from malignant IPMN	DL (CNN)	Total IPMN 50 patients 3970 still images 508,160 still images generated by data augmentation Benign IPMN 27 patients Malignant IPMN 23 patients	<i>Malignant IPMN recognition</i> Sn = 95.7% Sp = 92.6% PPV = 91.7% NPV = 9.6,2% Accuracy = 94% AUC = 0.98	AI using DL algorithm may be more accurate and objective to diagnose malignant IPMN compared to human diagnosis and conventional EUS features	1. Retrospective design 2. Single institution 3. Only surgical cases included
Marya et al. <sup>14</sup> USA	Retrospective review of still image and video data to differentiate NP, CP, AIP, and PDAC using AI	CNN	Total 583 patients Total 643 EUS exams Total 1,174,461 images NP 73 patients CP 72 patients AIP 146 patients PDAC 292 patients	<i>PDAC recognition versus all others (video)</i> Sn = 95% Sp = 91% PPV = 87% NPV = 97% AUC = 0.976 <i>PDAC recognition versus AIP</i> All images Sn = 90% Sp = 87% AUC = 0.95 Video-only Sn = 90% Sp = 93% AUC = 0.96	EUS-based CNN model successfully differentiated AIP from PDAC and all other conditions	1. Single center 2. Retrospective 3. Images not processed for AI in real time
Norton et al. <sup>15</sup> USA	Retrospective study of single EUS image from each procedure compared to EUS diagnosis reported on actual procedure	ANN (basic NN)	PC 21 patients Focal pancreatitis 14 patients Still image used	<i>PC recognition by AI</i> Sn = 100% Sp = 50% PPV = 75% NPV = 100% Accuracy = 80% <i>PC recognition by EUS</i> Sn = 89% Sp = 79% PPV = 89% NPV = 79% Accuracy = 85% <i>PC recognition by human</i> Sn = 73% Sp = 100% PPV = 100% NPV = 70% Accuracy = 83%	Analysis of EUS images with computer software programs compares favorably with human interpretation	1. Small sample 2. Only single image studied 3. Old AI model used compared to current times

(Continued)

Table 2. (Continued)

Author	Study design	AI system studied	Patient population	Outcomes of AI in detection of PC	Author conclusion	Limitations
Ozkan et al. <sup>16</sup> Turkey	Retrospective study of EUS images from patients with PC and non-cancer patients	ANN	PC 202 patients Non-cancer = 130 Total still images 332	PC recognition (all ages) Sn = 83.3% Sp = 93.33% Accuracy = 87.5% PC recognition (>60 years) Sn = 93.3% Sp = 88.88% Accuracy = 91.66% PC recognition (40–60 years) Sn = 85.7% Sp = 91.66% Accuracy = 88.46% PC recognition (<40 years) Sn = 87.5% Sp = 94.11% Accuracy = 92%	CAD system performed better in diagnosing PC based on EUS images when patients were classified by age	1. Small image numbers in the younger age groups (<60 years) 2. Single center 3. Retrospective 4. Other chronic pancreatic diseases not studied
Saftoiu et al. <sup>17</sup> Europe	Cross-sectional study to assess the accuracy of real-time EUS elastography subjected to extended NN malignancy from benign patterns	ANN	PC 32 patients PNET 3 patients CP 11 patients NP 22 patients Video images	PC and PNET recognition Sn = 91.4% Sp = 87.9% PPV = 88.9% NPV = 90.6% Accuracy = 89.7% AUC = 0.932	ANN processing of digitalized EUS elastography movies enabled optimal prediction of pancreatic lesion types	1. Small study in two centers 2. Dimensionality may have affected outcomes
Saftoiu et al. <sup>18</sup> Europe	Prospective multicenter-blinded analysis of real-time EUS elastography images to differentiate PC from CP using ANN model	ANN (MLP)	Total patients 258 Total 774 recordings PC 211 patients CP 47 patients Video images	PC recognition Sn = 87.59% Sp = 82.94% PPV = 96.25% NPV = 57.22% AUC = 0.94	AI methodology using ANN supports the medical decision-making process providing fast and accurate diagnosis	1. Uneven distribution with much less CP patients 2. Interobserver variability due to the use of recorded videos
Saftoiu et al. <sup>19</sup> Europe	Prospective multicenter observational study to assess utility of ANN to differentiate PC from CP	ANN	Total 167 patients PC 112 patients CP 55 patients Video images	PC recognition with AI Sn = 94.64% Sp = 94.44% PPV = 97.24% NPV = 89.47% PC recognition – contrast-enhanced EUS (no AI) Sn = 87.5% Sp = 92.72% PPV = 96.07% NPV = 78.46%	Computer-aided diagnostic system can differentiate PC and CP with good diagnostic results	1. Recorded videos analyzed, may cause inter-observer variability 2. Other focal pancreatic conditions excluded

(Continued)

Table 2. (Continued)

Author	Study design	AI system studied	Patient population	Outcomes of AI in detection of PC	Author conclusion	Limitations
Tonozuka <i>et al.</i> <sup>20</sup> Japan	Cross-sectional study of ability of DL system of EUS images in differentiation of PC from CP and NP	CNN	Total 139 patients PDAC 76 patients CP 34 patients NP 29 patients Total 920 still images PDAC 510 images CP 220 images NP 190 images Total 88,320 images (with data augmentation)	<i>PC recognition validation set</i> Sn = 90.2% Sp = 74.9% PPV = 80.1% NPV = 88.7% AUC = 0.924 <i>PC recognition test set</i> Sn = 92.4% Sp = 84.1% PPV = 86.8% NPV = 90.7% AUC = 0.940	EUS-CAD model can detect PDAC with good results	1. Single center 2. Retrospective design 3. Low number of patients 4. Other pancreatic disorders excluded
Zhang <i>et al.</i> <sup>21</sup> China	Retrospective study to assess ability to recognize PC from normal tissue using SVM of EUS images	SVM	Total 216 patients PC 153 patients Non-cancer 63 patients Still images	<i>PC recognition</i> Sn = 94.32% Sp = 99.45% PPV = 98.65% NPV = 97.77 Accuracy = 97.98%	SVM is a useful method to classify EUS images with application to PC and can be used for rapid non-invasive screening of pancreatic disorders	1. Single institution 2. Retrospective study 3. SVM not performed real time 4. Comparison group with no pancreatic disease
Zhu <i>et al.</i> <sup>22</sup> China	Retrospective study using CAD techniques to extract EUS image parameters in differentiating PC from CP	SVM	Total 388 patients PC 262 patients CP 126 patients Still images used	<i>PC recognition</i> Sn = 96.25% Sp = 93.38% PPV = 92.21% NPV = 96.68% Accuracy = 94.2%	Computer-aided EUS image differentiation is highly accurate and non-invasive for clinical determination of PC	1. Single institution study 2. Retrospective design 3. Only used a simple SVM classifier

AIP, autoimmune pancreatitis; ANN, artificial neural network; CAD, computer-aided diagnosis; CNN, convolutional neural network; DL, deep learning; EUS, endoscopic ultrasound; FNA, fine needle aspiration; IPMN, intraductal papillary mucinous neoplasm; MLP, multilayer perceptron; NP, normal pancreas; NN, neural network; NPV, negative predictive value; PC, pancreatic cancer; PDAC, pancreatic ductal adenocarcinoma; PNET, pancreatic neuroendocrine tumor; PPV, positive predictive value; Sn, sensitivity; Sp, specificity; SVM, support vector machine.

### Type of AI studied

SVM was used in two studies<sup>21,22</sup> while nine studies used NNs<sup>12–20</sup> – one used basic NN,<sup>15</sup> five used ANN<sup>12,16–19</sup> [with two using multilayer perceptron (MLP)]<sup>12,18</sup> while three studies used CNN.<sup>13,14,20</sup>

### Overall performance of AI in pancreatic malignancy recognition

Among the 11 studies,<sup>12–22</sup> the overall reported sensitivity of AI in recognizing malignancy of the pancreas ranged 83–100%, while specificity ranged 50–99% and accuracy 80–97.5%. The reported positive predictive value (PPV) and negative predictive value (NPV) ranged 75–99% and 57–100%, respectively. The DOR could be calculated for 10 studies, and it ranged 34–3003 (Table 3).

### AI to differentiate PC from CP

Seven studies<sup>12,14,15,17–19,22</sup> reported the diagnostic value of AI in differentiating PC from CP with a sensitivity and specificity ranging 88–100% and 50–94%, respectively. Excluding the study by Norton *et al.*<sup>15</sup> that used basic ANN, the sensitivity and specificity ranged 88–96% and 83–94%, respectively. One study<sup>22</sup> that used SVM to differentiate PC from CP reported the highest sensitivity of 96%, with 93% specificity and 94% accuracy. This study used a simple SVM classifier.

One study trained a CNN model with images from patients with PC, CP, and normal pancreas to identify PC with 90% sensitivity, 75% specificity with an AUC of 0.92.<sup>20</sup>

### AI to differentiate malignant from benign IPMNs

One study<sup>13</sup> reported using AI to differentiate benign from malignant IPMNs. It included 50 patients with IPMN with 23 malignant and 27 benign cases consisting of 3970 still images. Data augmentation was used to generate 508,160 images that were analyzed using the CNN system. The system reported 94% accuracy, 95.7% sensitivity, and 92.6% specificity to identify malignant IPMNs.

### SVM and PC recognition

Two studies<sup>21,22</sup> used SVM and reported accuracy, sensitivity, and specificity ranging 94–98%,

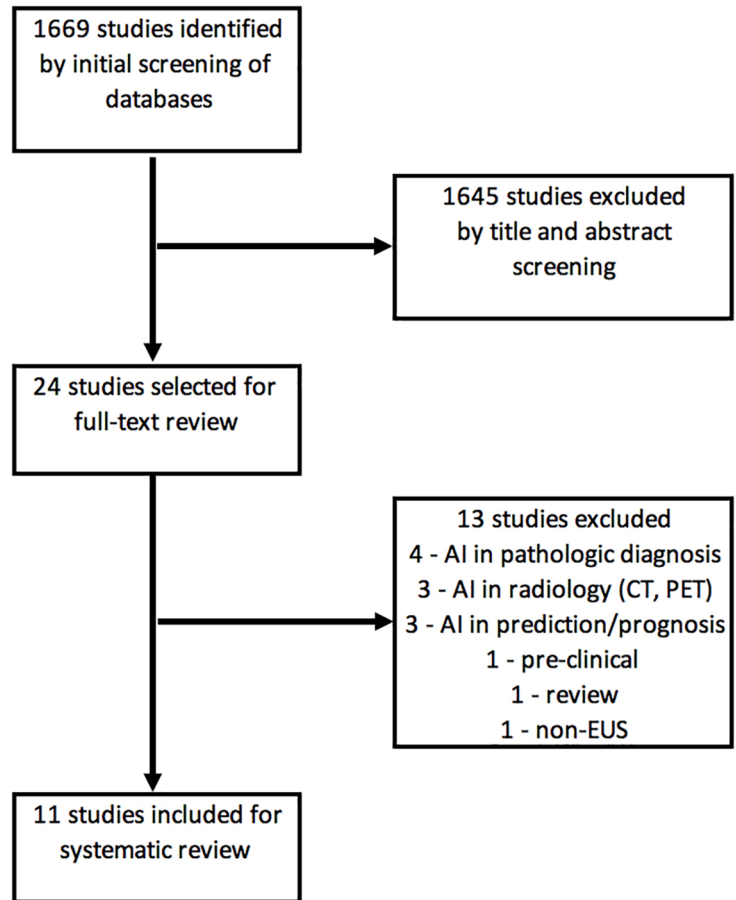


Figure 1. PRISMA flowchart.

94–96%, and 93–99%, respectively, to recognize PC, with the highest DOR of all studies. The corresponding PPV and NPV ranged 92–99% and 97–98%, respectively. Both these studies were retrospective and used still images in the AI model.

### CNN and PC recognition

Of the three studies<sup>13,14,20</sup> that used CNN, two<sup>14,20</sup> studied the diagnostic value of AI to recognize PC, while one<sup>13</sup> studied the differentiation of benign *versus* malignant IPMNs. The two studies<sup>14,20</sup> evaluating PC recognition reported sensitivity, specificity, PPV, and NPV ranging 92–95%, 84–91%, 87%, and 91–97%, respectively. The third study<sup>13</sup> reported sensitivity, specificity, PPV, and NPV of 96%, 93%, 92%, and 96%, respectively. All three studies used still images from the EUS, while one<sup>14</sup> also included video images. Two<sup>13,20</sup> of the three studies had small sample sizes of 50 and 139 patients and used data

**Table 3.** Performance of AI in diagnosis of pancreatic malignancies.

Study	AI type	Accuracy	Sn	Sp	PPV	NPV	DOR calculated
Das <i>et al.</i> <sup>12</sup>	ANN	n/a	0.93	0.92	0.87	0.96	153
Kuwahara <i>et al.</i> <sup>13</sup>	CNN	0.94	0.96	0.93	0.92	0.96	278
Marya <i>et al.</i> <sup>14</sup>	CNN	n/a	0.95	0.91	0.87	0.97	192
Norton <i>et al.</i> <sup>15</sup>	ANN	0.8	1.00	0.50	0.75	1.00	n/a
Ozkan <i>et al.</i> <sup>16</sup>	ANN	0.875	0.83	0.93	n/a	n/a	70
Saftoiu <i>et al.</i> <sup>17</sup>	ANN	0.897	0.91	0.88	0.89	0.91	77
Saftoiu <i>et al.</i> <sup>18</sup>	ANN	n/a	0.88	0.83	0.96	0.57	34
Saftoiu <i>et al.</i> <sup>19</sup>	ANN	n/a	0.95	0.94	0.97	0.89	300
Tonozuka <i>et al.</i> <sup>20</sup>	CNN	n/a	0.92	0.84	0.87	0.91	64
Zhang <i>et al.</i> <sup>21</sup>	SVM	0.975	0.94	0.99	0.99	0.98	3003
Zhu <i>et al.</i> <sup>22</sup>	SVM	0.942	0.96	0.93	0.92	0.97	362

AI, Artificial Intelligence; ANN, artificial neural network; CNN, convolutional neural network; DOR, diagnostic odds ratio; NPV, negative predictive value; PPV, positive predictive value; SVM, support vector machine.

augmentation to generate a large dataset of images from the original images to create the AI algorithm.

#### ANN and PC recognition

Six studies<sup>12,15–19</sup> using ANN (without DL) reported sensitivity, specificity, PPV, and NPV ranging 83–100%, 50–94%, 75–97%, and 57–100%, respectively. Three studies<sup>12,15,16</sup> used still images to be fed in the AI system, while three others<sup>17–19</sup> used video images for analysis. One study<sup>16</sup> divided patients by age and showed a higher sensitivity (93.3%) of AI in detecting PC in patients above 60 years of age compared to patients below 60 years of age where the sensitivity fell to 85.7% (age 40–60 years) and 87.5% in patients below 40 years of age.

#### Quality of included studies

The overall quality of studies was graded using the QUADAS-2 tool<sup>11</sup> (Figure 2). Four studies collected the index and standard test at the same time, while others were retrospective review of images. Although there is a low risk of bias for the performance of the reference test, there was a high risk of bias in several studies for the index test being performed unblinded. Being

retrospective in design, several studies suggested high risk of patient selection bias. Overall, the quality of evidence was low to moderate, mostly due to a high or unclear risk of bias in patient selection.

#### Discussion

EUS is a superior diagnostic modality than CT or MRI in the diagnosis of PC with high sensitivity (95%) but a rather low specificity (53%), especially in the presence of CP.<sup>4</sup> The cytologic diagnosis remains the mainstay of differentiating CP from PC.<sup>23</sup> Our systematic review suggests that AI may be used as a unique tool to augment the performance of EUS and improve its diagnostic ability for recognizing pancreatic malignancies even in the presence of CP, with improved sensitivity and specificity. Furthermore, the AI can be trained with still or video images or a combination of both. However, the performance of AI can vary depending on the type of the AI model used.

Different AI systems vary in complexity (Table 1). For instance, SVM is a type of supervised ML method where the data fed in the system are classified into two or more categories separated by a linear line for two categories and plans for more than two categories. The calculation for two



	Risk of Bias				Applicability Concerns		
	Patient Selection	Index test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Das et al 2008 [12]	?	-	+	+	?	+	+
Kuwahara et al 2019 [13]	?	-	+	+	+	+	+
Marya et al 2020 [14]	+	-	+	+	+	+	+
Norton et al 2001 [15]	?	+	+	+	?	+	+
Ozkan et al 2016 [16]	?	-	+	+	?	+	+
Saftoiu et al 2008 [17]	+	+	+	+	+	+	+
Saftoiu et al 2012 [18]	+	+	+	+	+	+	+
Saftoiu et al 2015 [19]	+	+	+	+	+	+	+
Tonozuka et al 2020 [20]	?	+	+	+	?	+	+
Zhang et al 2010 [21]	-	-	+	+	-	+	+
Zhu et al 2013 [22]	-	+	+	+	?	+	+

**Figure 2.** QUADAS-2 analysis of study quality/risk of bias.

categories separated by a linear line requires a fairly simple calculation. In addition, the dividing line may be soft to allow accommodation of anomalous readings. Although it is fairly simpler than NNs and more generalizable, SVM requires a large amount of data for ML.<sup>6,8</sup>

ANN is a form of AI where the system tries to mimic neural circuits in the human brain. The data pass through multiple layers connected by nodes, and each connection is given a certain weight, indicating the strength of the connection that the system can adjust as it learns. The system can also adjust for bias and provides output by making necessary corrections through forward- and back-propagation of the data through the layers.<sup>6,7</sup>

CNN is the most sophisticated AI system since it is designed to emulate the visual signals processed by the biological brain. It is more independent in its learning compared to SVM, which is supervised ML. Simplistically, the system extracts distinct features from the data, creates classifications, and applies specific filters to create multiple feature maps. Each image is subject to filtering

giving it the name convolutional. A final layer combines the all-filter layers in a fully connected layer giving the final result.<sup>6,7</sup>

In our systematic review, the SVM methodology revealed the highest sensitivity, specificity, and diagnostic accuracy to distinguish PC from CP and normal pancreas with 94–98% accuracy, 94–96% sensitivity, and 93–99% specificity, respectively.<sup>21,22</sup> CNN was also effective in making that determination, but the specificity ranged 84–87%.<sup>14,20</sup> However, in the differentiation of benign *versus* malignant IPMNs, CNN performs better (with accuracy, sensitivity, and specificity of 94%, 95.7%, and 92.6%, respectively).<sup>13</sup> These numbers are higher compared to EUS alone per historic data.<sup>23</sup> ANN was more variable in performance, but when the study applying basic ANN was excluded, the specificity increased to 83–93%, again better than conventional EUS alone.

Few limitations exist for our systematic review. Most of the published literature included a small number of patients with a retrospective, non-randomized design. It remains to be seen if the performance of AI-assisted EUS in real time can

match or exceed these numbers. Another drawback is that most images selected were recorded and reviewed by extremely experienced endoscopists before being fed in the AI system; hence, the generalized applicability of AI to endoscopists with all levels of experience also remains to be seen. In addition, studies were heterogeneous in the types and methodologies of AI studied. Nevertheless, AI performed better than conventional EUS overall in differentiating PC from CP and non-cancer patients. SVM method turned out to be a simpler system than CNN and with its high performance seems promising in recognizing cancer in the presence of chronic pancreatic inflammation or screening for PC especially in high-risk individuals. However, further studies are needed, especially prospective and real time, to establish the role of AI in routine EUS procedures for endoscopists of all training levels.

If AI development continues at the current pace, it may be possible to use AI in the future to accurately differentiate PC from CP and other non-cancer conditions with EUS imaging alone without the need for pathological diagnosis. This would be especially helpful in PC screening in high-risk patients with germline mutations and genetic syndromes that places them at high risk of PC but currently have no consensus on effective screening.

### Conclusion

AI technology is a promising adjunct to conventional EUS for recognizing PC, especially in the presence of CP. Of all the AI modalities under development, SVM reported the highest sensitivities, specificities, and DOR for recognition of PC. Being a simpler system than ANN and CNN, the SVM system seems worthy of further exploration in prospective studies, possibly as a quick screening tool, especially in high-risk individuals. However, further studies are needed for its refinement and for use in daily practice.

### Authors statement

The article has been read and approved by all the authors that the requirements for authorship as stated earlier in this document have been met and that each author believes that the article represents honest work.

### Ethical statement

As submitted or its essence in another version, this article is not under consideration for publication elsewhere. It will not be published elsewhere while under review by Therapeutic Advances in Gastroenterology. All authors have made substantive contributions to the study, and all authors endorse the data and conclusions.

### Author contributions

**Hemant Goyal:** Conceptualization; Methodology; Project administration.

**Syed Ali Amir Sherazi:** Data curation; Investigation; Resources; Writing – original draft.

**Shweta Gupta:** Formal analysis; Methodology; Writing – review & editing.

**Abhilash Perisetti:** Resources; Supervision; Writing – review & editing.

**Ikechukwu Achebe:** Resources; Visualization; Writing – review & editing.

**Aman Ali:** Project administration; Supervision; Writing – review & editing.

**Benjamin Tharian:** Project administration; Resources; Supervision; Writing – review & editing.

**Nirav Thosani:** Supervision; Writing – review & editing.

**Neil R. Sharma:** Project administration; Supervision; Writing – review & editing.

### ORCID iD

Hemant Goyal  <https://orcid.org/0000-0002-9433-9042>

### Funding

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

### Conflict of interest statement

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Hemant Goyal – Consultant for Aimloxy LLC.

### Supplemental material

Supplemental material for this article is available online.

## References

1. Lippi G and Mattiuzzi C. The global burden of pancreatic cancer. *Arch Med Sci* 2020; 16: 820–824.
2. Lucas AL and Kastrinos F. Screening for pancreatic cancer. *JAMA* 2019; 322: 407–408.
3. Ang TL, Kwek ABE and Wang LM. Diagnostic endoscopic ultrasound: technique, current status and future directions. *Gut Liver* 2018; 12: 483–496.
4. Harmsen F, Domagk D, Dietrich CF, *et al.* Discriminating chronic pancreatitis from pancreatic cancer: contrast-enhanced EUS and multidetector computerized tomography in direct comparison. *Endosc Ultrasound* 2018; 7: 395–403.
5. Shahidi N, Ou G, Lam E, *et al.* When trainees reach competency in performing endoscopic ultrasound: a systematic review. *Endosc Int Open* 2017; 5: E239–E243.
6. Tonozuka R, Mukai S and Itoi T. The role of artificial intelligence in endoscopic ultrasound for pancreatic disorders. *Diagnostics* 2021; 11: 18.
7. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw* 2015; 61: 85–117.
8. Cortes C and Vapnik V. Support vector networks. *Machine Learning* 1995; 20: 273–297.
9. Yang YJ and Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019; 25: 1666–1683.
10. Xu W, Liu Y, Lu Z, *et al.* A new endoscopic ultrasonography image processing method to evaluate the prognosis of pancreatic cancer treated with interstitial brachytherapy. *World J Gastroenterol* 2013; 19: 6479–6484.
11. Whiting PF, Rutjes AWS, Westwood ME, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–536.
12. 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.
13. 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.
14. Marya NB, Powers PD, Chari ST, *et al.* Utilization of artificial intelligence for the development of an EUS-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut* 2021; 70: 1335–1344.
15. Norton ID, Zheng Y, Wiersema MS, *et al.* Neural network analysis of EUS images to differentiate between pancreatic malignancy and pancreatitis. *Gastrointest Endosc* 2001; 54: 625–629.
16. 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.
17. Saftoiu 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.
18. Saftoiu A, Vilmann P, Gorunescu F, *et al.* European EUS Elastography Multicentric Study Group. Efficacy of an artificial neural network-based approach to endoscopic ultrasound elastography in diagnosis of focal pancreatic masses. *Clin Gastroenterol Hepatol* 2012; 10: 84–90.
19. Saftoiu 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.
20. Tonozuka R, Itoi T, Nagata N, *et al.* Deep learning analysis for the detection of pancreatic cancer on endosonographic images: a pilot study. *J Hepato-Biliary Pancreat Sci* 2021; 28: 95–104.
21. Zhang MM, Yang H, Jin ZD, *et al.* Differential diagnosis of pancreatic cancer from normal tissue with digital imaging processing and pattern recognition based on a support vector machine of EUS images. *Gastrointest Endosc* 2010; 72: 978–985.
22. 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.
23. Brand B, Pfaff T, Binmoeller KF, *et al.* Endoscopic ultrasound for differential diagnosis of focal pancreatic lesions, confirmed by surgery. *Scand J Gastroenterol* 2000; 35: 1221–1228.