—Review Article—

The value of artificial intelligence techniques in predicting pancreatic ductal adenocarcinoma with EUS images: A meta-analysis and systematic review

Hua Yin[1,2,3,*], Xiaoli Yang[1,2,3,*], Liqi Sun[2,*], Peng Pan[2], Lisi Peng[2], Keliang Li[4], Deyu Zhang[2], Fang Cui[2], Chuanchao Xia[2], Haojie Huang[2,#], Zhaoshen Li[2,#]

[1]Department of Gastroenterology, General Hospital of Ningxia Medical University, Yinchuan, China; [2]Department of Gastroenterology, Changhai Hospital, Second Military Medical University, Shanghai, China; [3]Postgraduate Training Base in Shanghai Gongli Hospital, Ningxia Medical University, Shanghai, China; [4]Department of Gastroenterology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan Province, China

ABSTRACT

Conventional EUS plays an important role in identifying pancreatic cancer. However, the accuracy of EUS is strongly influenced by the operator's experience in performing EUS. Artificial intelligence (AI) is increasingly being used in various clinical diagnoses, especially in terms of image classification. This study aimed to evaluate the diagnostic test accuracy of AI for the prediction of pancreatic cancer using EUS images. We searched the Embase, PubMed, and Cochrane Library databases to identify studies that used endoscopic ultrasound images of pancreatic cancer and AI to predict the diagnostic accuracy of pancreatic cancer. Two reviewers extracted the data independently. The risk of bias of eligible studies was assessed using a Deek funnel plot. The quality of the included studies was measured by the QUDAS-2 tool. Seven studies involving 1110 participants were included: 634 participants with pancreatic cancer and 476 participants with nonpancreatic cancer. The accuracy of the AI for the prediction of pancreatic cancer (area under the curve) was 0.95 (95% confidence interval [CI], 0.93–0.97), with a corresponding pooled sensitivity of 93% (95% CI, 0.90-0.95), specificity of 90% (95% CI, 0.8-0.95), positive likelihood ratio 9.1 (95% CI 4.4-18.6), negative likelihood ratio 0.08 (95% CI 0.06-0.11), and diagnostic odds ratio 114 (95% CI 56–236). The methodological quality in each study was found to be the source of heterogeneity in the meta-regression combined model, which was statistically significant ($P = 0.01$). There was no evidence of publication bias. The accuracy of AI in diagnosing pancreatic cancer appears to be reliable. Further research and investment in AI could lead to substantial improvements in screening and early diagnosis.

Key words: accuracy, AI, artificial intelligence, EUS, pancreatic cancer, predicting pancreatic ductal adenocarcinoma

Access this article online

Quick Response Code:

Website: www.eusjournal.com

DOI: 10.4103/EUS-D-21-00131

How to cite this article: Yin H, Yang X, Sun L, Pan P, Peng L, Li K, *et al.* The value of artificial intelligence techniques in predicting pancreatic ductal adenocarcinoma with EUS images: A meta-analysis and systematic review. Endosc Ultrasound 2023;12:50-8.

*Dr. Hua Yin, Dr. Xiaoli Yang and Dr. Liqi Sun contributed equally to this work.

#These authors share senior authorships.

Address for correspondence

Dr. Zhaoshen Li, Department of Gastroenterology, Changhai Hospital, Second Military Medical University, Shanghai, China.

E-mail: zhs_li@126.com

Received: 2021-06-11; **Accepted:** 2021-11-29; **Published online:** 2022-03-21

INTRODUCTION

The morbidity and mortality of pancreatic cancer are on the rise worldwide.[1] According to the data released by the American Cancer Society in 2020, pancreatic cancer ranks 10th in the number of new cases in men, ninth in women, and fourth in the number of deaths from in the United States.[2] The prognosis for pancreatic cancer is very poor, with a typical 1-year survival rate of only 24% and a 5-year survival rate of 9% after the diagnosis.[3] However, the 5-year survival rate of cases with PC smaller than 10 mm reaches 80.4% and that of cases with stage 0 is 85.8%.[4,5] It is very important to improve the diagnostic rate of pancreatic cancer, especially the early diagnostic rate.

EUS is now considered to be the most sensitive imaging method for the detection of pancreatic lesions. The median sensitivity of EUS in the detection of pancreatic tumors was 94%.[6] The diagnostic sensitivity of EUS was significantly better than that of computed tomography (CT) and magnetic resonance imaging for pancreatic tumors <30 mm.[7] EUS is also more sensitive than computed tomography CT, abdominal ultrasound and PET imaging in the diagnosis of early pancreatic cancer <10 mm.[8] The accuracy of EUS and CT for the detection of pancreatic cancer lesions measuring ≤1.5 cm has been reported to be 100% *vs.* 67%.[9] All of these are attributed to the high-resolution images that endoscopic ultrasonography can provide.

The accuracy of EUS depends on both operator- and lesion-related factors. The most important operator-related factor is the amount of experience performing EUS. Although new techniques for EUS, including the incorporation of elastography and contrast-enhanced EUS (CE-EUS), both have high sensitivity and specificity in the diagnosis of pancreatic cancer,[10,11] novice endoscopists require substantial time to perform this task efficiently. These new technologies are also influenced by the operator's experience and executive factors. Therefore, it is very significant to reduce the anthropogenic factors to improve the diagnosis of pancreatic cancer.

The image analysis capabilities of artificial intelligence (AI) have been studied in various diseases, some of which have been applied in clinical practice.[12,13] More advanced AI is already being used in the field of gastrointestinal endoscopy, including the detection of colon polyps.[14,15] Compared with other fields, the application of AI in pancreatic EUS is still in the development stage. AI technology has a promising application in endoscopic ultrasonography and is helpful for the diagnosis of pancreatic cancer. There has been no meta-analysis that has studied the diagnostic test accuracy of AI for the prediction of pancreatic cancer using EUS images. This study aimed to evaluate the diagnostic performance of AI for the diagnosis of pancreatic cancer using endoscopic ultrasound imaging.

METHODS

Trial registration

We identified the study protocol and registered the study in the International Systematic Review Prospective Registry (PROSPERO), CRD42021256916.

Search strategy and study selection

PubMed, Embase, and the Cochrane Library were used to retrieve potential eligible studies up until May 15, 2021. The search results, including title and abstract, were reviewed by two independent investigators (YH and YXL). Appendix 1 shows the detailed search strategy that combined common keywords relevant to pancreatic neoplasms and AI. Full-text reviews were conducted to determine whether the inclusion criteria were satisfied in all the studies. Disagreements between the evaluators were resolved by a consultation with a third evaluator (SLQ).

Studies were eligible if they met the following prespecified inclusion criteria: (1) studies were based on the analysis of B-mode images of endoscopic ultrasonography by AI; (2) application of the AI algorithm for the prediction of pancreatic cancer; (3) studies with data on the inability to construct a 2 × 2 contingency table; and (4) prospective or retrospective study design. The exclusion criteria included incomplete data, narrative reviews, letters, comments, editorials, protocol studies, and guidelines. Studies meeting at least one of the exclusion criteria were excluded from the analysis.

Data extraction and study quality assessment

Two evaluators independently used the same data form to extract the summary of TP (subjects with a positive finding using AI who had pancreatic cancer on the EUS images), FP (subjects with a positive finding using AI who did not have pancreatic cancer on the EUS images), FN (participants with a negative finding using

AI who had pancreatic cancer on the EUS images), and TN subjects with a negative finding using AI who did not have pancreatic cancer on the EUS images). For the studies with incomplete data, we calculated the values for TP, FP, FN, and TN using the formulas[16] or contacted the corresponding author of each study through E-mail to obtain the exact values of the following primary outcomes. We applied 2 × 2 tables whenever possible to the data of the original articles that contained various diagnostic performance indices.

Review Manager version 5.4 (RevMan for Windows 7, Nordic Cochrane Centre) was used to generate the summary figure of the methodological quality evaluation. The Quality Assessment of Diagnostic Accuracy Studies tool was used to assess the methodological quality of and potential bias of all of the studies by two independent reviewers.[17] Conflicts were resolved by discussion with and involvement of the third author.

Statistical methods

The meta-analysis used Stata Statistical Software version 15.1 (College Station, Texas, US) and relevant packages of metandi and midas. A forest plot of pooled sensitivity or specificity using a bivariate model, as well as a summary receiver operating characteristic (SROC) curve using a hierarchical SROC (HSROC) model, was generated and presented. Subgroup analyses and meta-regression were used to explore the heterogeneity among the studies. All covariates used in the subgroup analysis were predetermined. Sources of potential heterogeneity were included as covariates in the HSROC model, as recommended by the Cochrane Handbook for Diagnostic Tests Review. The heterogeneity of these studies was determined by the correlation coefficient between the sensitivity and specificity of the logit transformation using the bivariate model and the asymmetric parameter β (beta), where β=0 corresponds to the symmetric receiver operating characteristic (ROC) curve. According to the HSROC model, the diagnostic odds ratio (DOR) does not change along the curve. A positive correlation coefficient (>0) and the β and a *P* value were significant ($P < 0.05$), indicating heterogeneity between the studies. The risk of bias of eligible studies was assessed using a Deek funnel plot. Fagan's nomogram results showed clinical usability. In this nomogram, the left axis represents the pretest probability, the middle axis represents the likelihood ratio, and the right axis shows the posttest probability. Initially, we found and marked the pretest probability and likelihood ratio values on the left and middle axes, respectively. Then, a straight line was drawn from the two marked points along the right axis. The point at which the line crosses the left axis is the value of posttest probability.[18]

RESULTS

Identification of relevant studies

Among the 1221 studies retrieved, 987 were selected after removal of duplicates and 970 were excluded after screening the title and abstract. Full texts of the remaining 17 articles were thoroughly reviewed. Among these, 10 studies were excluded from the final analysis for the following reasons: review ($n = 2$), study with incomplete data ($n = 3$) and failure to meet the inclusion criteria ($n = 5$). Seven studies were selected for the inclusion in the meta-analysis. Figure 1 illustrates a flow diagram showing the process used to identify the relevant articles.

Characteristics of eligible studies

We identified 1110 subjects in seven studies[19-25] that were included for the prediction of pancreatic cancer using EUS images. There were 634 subjects with pancreatic cancer and 476 subjects with nonpancreatic cancer. All studies were retrospective.[19-25] According to the development history of AI, computer-aided diagnosis can be divided into deep learning-based and conventional types before deep learning. Among the studies, computer-aided diagnosis was of the conventional type in five studies.[19-23] Two computer-aided diagnosis studies used deep learning to predict pancreatic cancer.[24,25] Two studies[23,24] reported gender and age, and the age of the enrolled population ranged from a mean of 50.91 years to a mean of 64 years. Most of the studies[19,20,23-25] established AI algorithms based on neural networks, while only two studies[21,22] established support vector machines (SVM). Six studies[19-22,24,25] focused on pancreatic cancer versus chronic pancreatitis (including chronic pancreatitis, chronic pancreatitis and normal pancreas) to predict pancreatic cancer. Only one study[23] compared pancreatic cancer with a normal pancreas. There were five studies[21-25] with a total number of included patients greater than 100. Five studies[21-25] were published after 2010, and two studies[19,20] were published before 2010. In one study,[25] the data were extracted from the abstract, and some details were available from another study[26] that used the same database. These characteristics (modifiers) were evaluated as potential sources of heterogeneity through subgroup analysis and

meta-regression. Table 1 shows the detailed characteristics of the included studies.

Risk of bias assessment of eligible studies

The quality of the included studies and the risk of bias using the revised Quality Assessment of Diagnostic Accuracy Studies tool are presented in Figure 2. Regarding patient selection, the risk of bias was low in five studies, unclear in 1 study[20] due to insufficient information describing the sampling method, and high risk in 1 study.[19] One study[19] only selected histologically positive patients with chronic pancreatitis, which may have led to selection bias. This study was rated as high risk in the patient selection domain in the risk of bias evaluation. For

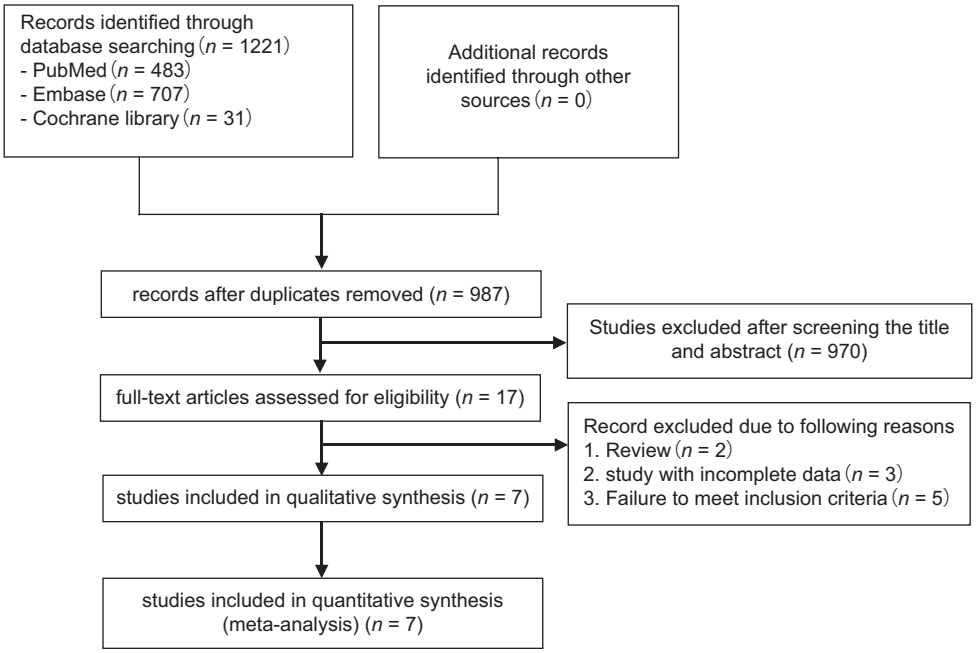

Figure 1. Flow diagram of identification of relevant research

Table 1. Clinical characteristics of the included studies

Study, format, nationality	Year diagnostic method of PC type of EUS	Number and type of patients	Number of training and validation group	Number of testing group	TP	FP	FN	TN	Algorithm of AI type of CAD
Norton ID *et al.* Retrospective USA	2001 Histologic B-mode image	21PC *versus* 14CP	-	21PC *versus* 14CP	21	7	0	7	Neural network Conventional
Das Ananya *et al.* Retrospective USA	2008 FNA B-mode image	22PC *versus* 12CP and 22NP	50PC *versus* 55CP/55NP	55PC *versus* 49CP and 55NP	51	8	4	96	ANN Conventional
Zhang M *et al.* Retrospective China	2010 FNA B-mode image	153PC *versus* 43CP and 20NP	76PC *versus* 32NC	77PC *versus* 31NC	73	1	4	30	SVM Conventional
Zhu M *et al.* Retrospective China	2013 FNA B-mode image	262PC *versus* 126 CP	131PC *versus* 63CP	131PC *versus* 63 CP	120	3	11	60	SVM Conventional
Ozkan M *et al.* Retrospective Turkey	2016 FNA B-mode image	93PC *versus* 79NP	160PC *versus* 100NP	42PC *versus* 30NP	35	2	7	28	ANN Conventional
Tonozuka *et al.* Retrospective Japan	2020 FNA or surgery B-mode image	77PC *versus* 34CP and 29NP	520PC *versus* 220CP/190NP	250PC *versus* 120CP and 100NP	231	35	19	185	CNN Deep-learning
Marya *et al.* Retrospective USA	2020 Histologic B-mode image	288PC *versus* 72CP or72NP or144AIP	230PC *versus* 72CP or72NP or144AIP	58PC *versus* 14CP	57	1	1	13	CNN
				58PC *versus* 14NP	53	1	5	13	Deep-learning
				58PC *versus* 29AIP	52	3	6	26	

PC: Pancreatic cancer; AI: Artificial intelligence; CP: Chronic pancreatitis; NP: Normal pancreas; NC: Noncancer cases; FNA: Guided fine-needle aspiration biopsy; ANN: Artificial neural network; CNN: Convolutional neural network; SVM: Support vector machine; TP: True positive; FP: False positive; FN: False negative; TN: True negative; CAD: Computer-aided diagnosis

publication bias, the test for Deek funnel plot asymmetry was insignificant (P =0.25). Figure 3 shows the Deek funnel plot of the studies to explore heterogeneity with Meta-Regression and Subgroup Analysis. For the prediction of pancreatic cancer using the EUS images, the shape of the SROC curve was symmetric [Figure 4]. The HSROC model obtained β estimates and 95% confidence intervals (CI) of 1.44 (-1.41–4.30), P = 0.32, suggesting that the SROC was symmetric [Figure 5]. However, the 95% prediction region in the SROC curve was wide. Methodological quality was found to be the source of heterogeneity in the joint meta-regression model, which was statistically significant (P = 0.01). There was no statistical significance in the type of AI algorithm (P = 0.16), year of publication (P = 0.29), type of computer-aided diagnosis (P = 0.8), total number of included patients (P = 0.29), or type of control (P = 0.11) [Figure 6].

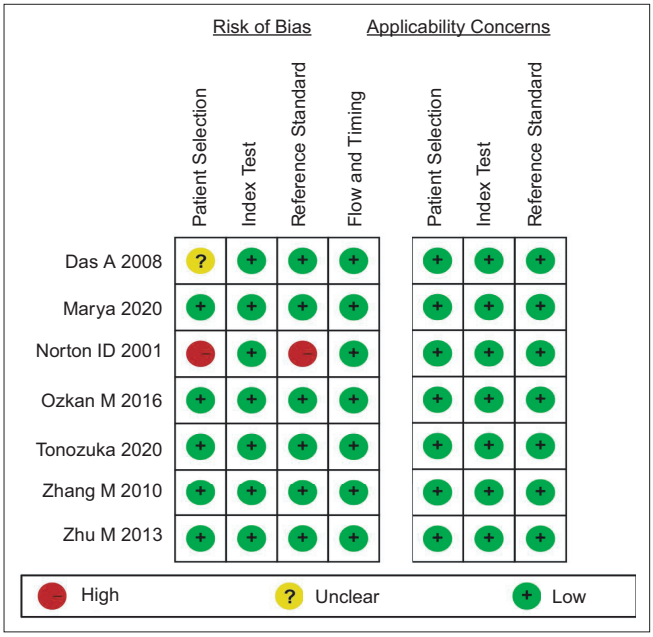

Figure 2. Quality assessment of diagnostic accuracy studies – 2 for the assessment of the methodological qualities of all the enrolled studies. (+) denotes low risk of bias, (?) denotes unclear risk of bias, (-) denotes high risk of bias

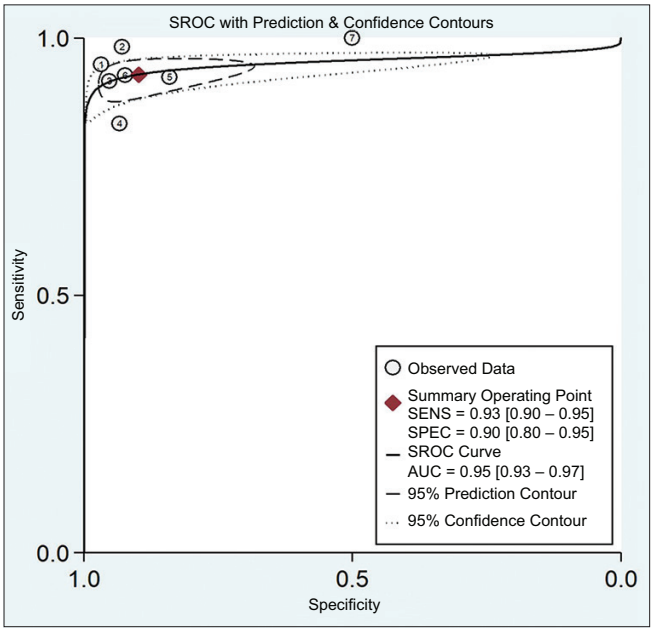

Figure 4. Summary receiver operating characteristic curve with 95% confidence region and prediction region for the prediction of pancreatic cancer in EUS images

Diagnostic test accuracy of AI for the prediction of pancreatic cancer

Among the 7 studies,[19-25] the sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, DOR, and area under the curve (AUC) with 95% CI

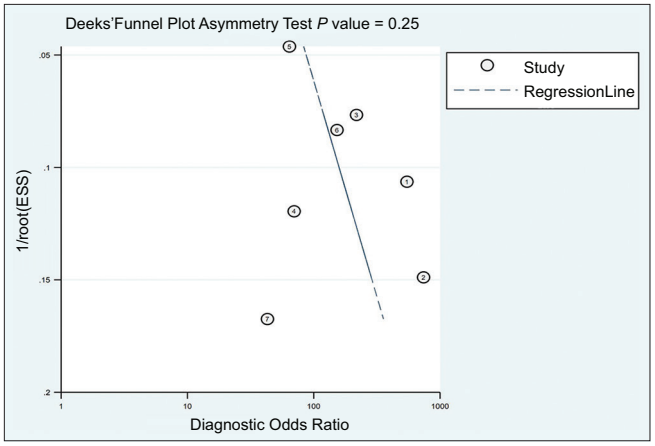

Figure 3. Deek funnel plot for studies

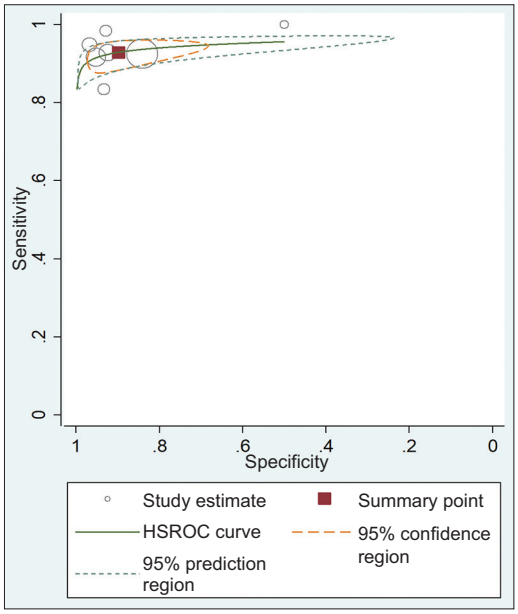

Figure 5. Hierarchical summary receiver operating characteristic curve

of AI for the prediction of pancreatic cancer were 93% (95% CI, 0.9–0.95), 90% (95% CI, 0.8–0.95),9.1 (95% CI 4.4–18.6), 0.08 (95% CI 0.06–0.11), 114 (95% CI 56–236), and 0.95 (95% CI 0.93–0.97), respectively [Figure 7]. The SROC curve, with a 95% confidence region and prediction region, is illustrated in Figure 4. We generated a Fagan diagram to study the clinical application of AI. Assuming a 20% prevalence of pancreatic cancer, the Fagan diagram shows a posterior probability of 69% for pancreatic cancer if the test is positive and 2% for a negative test [Figure 8].

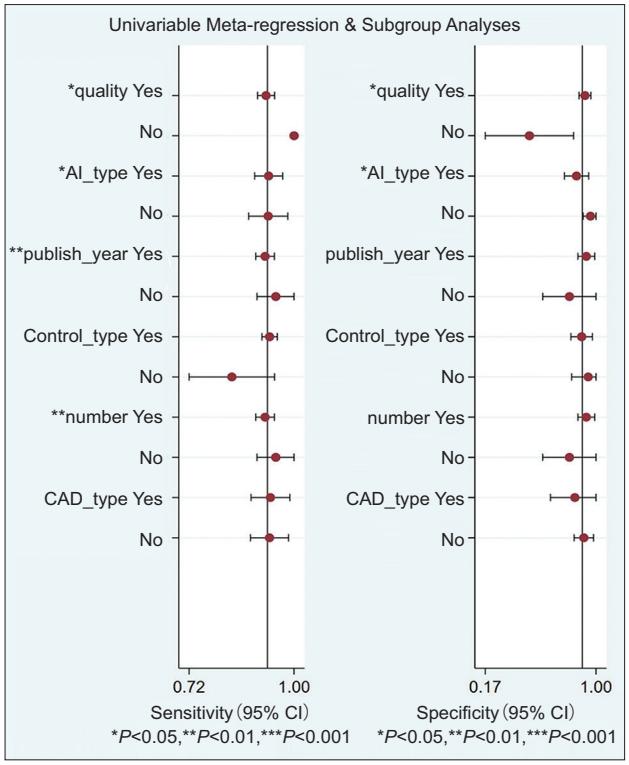

Figure 6. Meta-regression for the reason of heterogeneity in the diagnostic test accuracy meta-analysis

In five studies, computer-aided diagnosis fell into the traditional type.[19-23] The pooled sensitivity of pancreatic cancer prediction was 94% (95% CI, 86%–97%) with a specificity 91% (95% CI, 78%–97%). The AUC was 0.97 (95% CI, 95–98). Two studies of computer-aided diagnosis types used deep learning to predict pancreatic cancer.[24,25] Therefore, a meta-analysis was not possible. Pooled analysis of the crude values of TP, FP, FN, and TN revealed that the accuracy of the deep-learning type of AI reached 89.9%.

Five studies used neural network.[19,20,23-25] The pooled sensitivity of pancreatic cancer prediction was 94% (95% CI, 87%−97%), and the specificity was 85% (95% CI, 71%−93%). The AUC was 0.96 (95% CI, 94-97). Two studies used SVM to predict pancreatic cancer.[21,22] Therefore, a meta-analysis was not possible. Pooled analysis of the crude values of TP, FP, FN, and TN revealed that the accuracy of the AI algorithm reached 93.7%.

Six studies[19-22,24,25] focused on pancreatic cancer versus chronic pancreatitis (including chronic pancreatitis, chronic pancreatitis, and normal pancreas) to predict

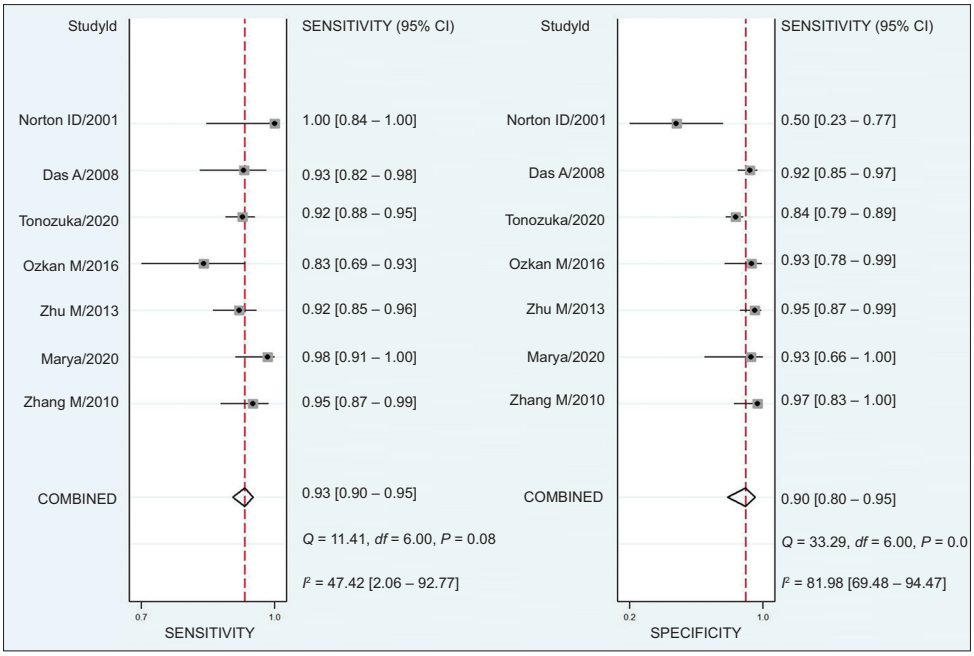

Figure 7. Forest plots of sensitivity and specificity of AI for the prediction of pancreatic cancer in EUS images

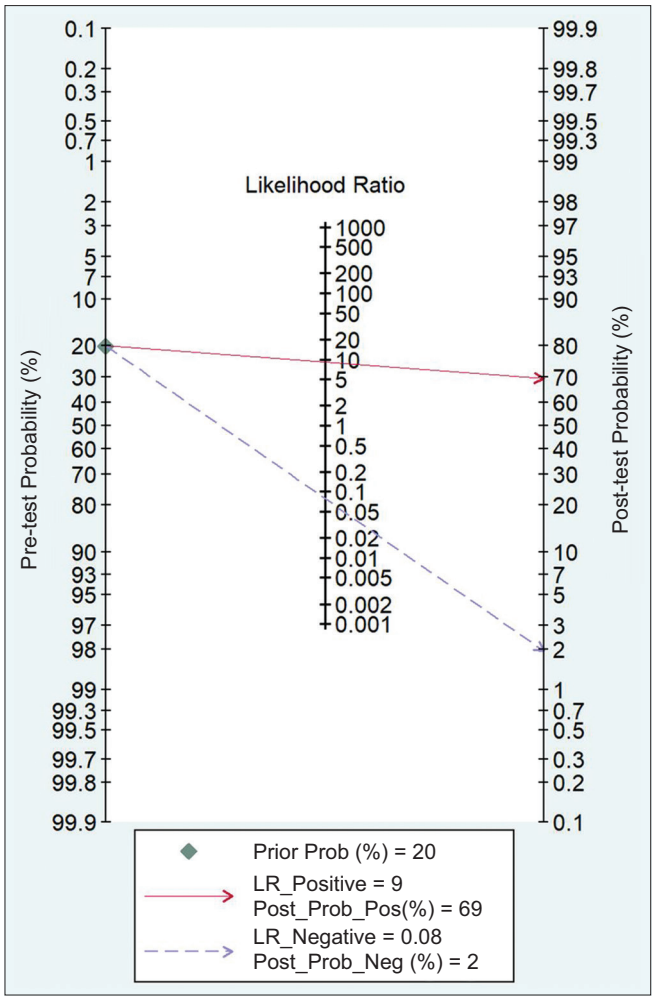

Figure 8. Fagan normogram for the prediction of pancreatic cancer in EUS images

pancreatic cancer. The pooled sensitivity of pancreatic cancer prediction was 93% (95% CI, 91%–95%), and the specificity was 89% (95% CI, 77%–95%). The AUC was 0.94 (95% CI, 92–96). One study compared pancreatic cancer with normal pancreas glands. A meta-analysis was not possible.

Six of the studies were of high quality.[20-25] The pooled sensitivity of pancreatic cancer prediction was 93% (95% CI, 90%–95%), and the specificity was 92% (95% CI, 86%–95%). The AUC was 0.93 (95% CI, 91–95). One study was low quality,[19] and a meta-analysis was not possible. The accuracy of the AI calculated from TP, FP, FN, and TN was 80%.

There were five studies[21-25] with a total number of included patients greater than 100. The pooled sensitivity in pancreatic cancer prediction was 92% (95% CI, 90%–95%), and the specificity was 95% (95% CI, 84%–96%). The AUC was 0.93 (95% CI, 91–95). There were two studies[19,20] with a total number of included patients less than 100. A meta-analysis was not possible. The accuracy of the AI calculated from TP, FP, FN, and TN was 90.2%.

Five studies[21-25] were published after 2010. The AUC was 0.93 (95% CI, 91–95), and the pooled sensitivity of pancreatic cancer prediction was 92% (95% CI, 90%−95%), and the specificity was 95% (95% CI, 84%−96%). Two studies[19,20] were published before 2010. Pooled analysis of the crude values of TP, FP, FN, and TN revealed that the accuracy of the AI algorithm reached 90.2%. Table 2 shows the detailed results of the subgroup analyses.

DISCUSSION

This is the first meta-analysis on the diagnostic accuracy of AI for the prediction of pancreatic cancer. Herein, we showed that AI was accurate for the diagnosis of pancreatic cancer, with a pooled sensitivity of 93%, specificity of 90%, and AUC of 0.95. The results show that AI has good performance, indicating that AI -assisted EUS may be applied in clinical practice. In general, the performance of the human brain differs from that of machines. In particular, when using EUS images for diagnosis, factors such as physician fatigue, stress or limited experience may lead to neglected or misdiagnosed diseases. In contrast, AI can continuously provide a reliable performance over a short period of time, has the potential to compensate for limited human capabilities, can prevent mistakes made by doctors in clinical practice, and can also facilitate the training and education of less experienced endoscopists.

EUS is the most suitable modality for the detection of pancreatic lesions due to its spatial resolution. Based on the early diagnosis of pancreatic cancer by EUS, the operator's experience and subjective factors have a great influence on the results, especially in the presence of chronic pancreatitis, and even experienced endoscopists may produce false negatives.[27] In addition, combining advanced imaging techniques such as elastography and contrast-enhanced ultrasound with EUS for the early detection may improve the diagnosis of early pancreatic cancer. AI-assisted ultrasound elastography and contrast-enhanced ultrasound also show high accuracy in the diagnosis of pancreatic cancer.[28-30] Săftoiu *et al.*[30] reported that the sensitivity and specificity of AI for distinguishing of pancreatic cancer and mass-forming

Table 2. Summary of diagnostic test accuracy and subgroup analysis of the included studies

Subgroup	Number of studies	Sensitivity (95% CI)	Specificity (95% CI)	PLR	NLR	DOR	AUC
Methodological quality of included studies							
High	6	0.93 (0.90-0.95)	0.92 (0.86-0.95)	11.5 (6.5-20.3)	0.08 (0.06-0.11)	141 (69-288)	0.93 (0.91-0.95)
Low	1	-	-	-	-	-	-
Total number of included patients							
>100	5	0.92 (0.90-0.95)	0.92 (0.84-0.96)	11.9 (5.7-24.7)	0.08 (0.06-0.11)	146 (60-355)	0.93 (0.91-0.95)
<100	2	-	-	-	-	-	-
Published year							
After 2010	5	0.92 (0.90-0.95)	0.92 (0.84-0.96)	11.9 (5.7-24.7)	0.08 (0.06-0.11)	146 (60-355)	0.93 (0.91-0.95)
Before 2010	2	-	-	-	-	-	-
Type of AI							
Neural network	5	0.94 (0.87-0.97)	0.85 (0.71-0.93)	6.2 (3.1-12.2)	0.07 (0.04-0.15)	83 (49-143)	0.96 (0.94-0.97)
SVM	2	-	-	-	-	-	-
Type of patient							
PC *versus* CP	6	0.93 (0.91-0.95)	0.89 (0.77-0.95)	8.8 (3.9-19.9)	0.07 (0.05-0.10)	120 (48-297)	0.94 (0.92-0.96)
PC *versus* NP	1	-	-	-	-	-	-
Type of CAD							
Convention	5	0.94 (0.86-0.97)	0.91 (0.78-0.97)	10.5 (4.0-27.1)	0.07 (0.03-0.15)	153 (63-370)	0.97 (0.95-0.98)
Deep learning	2	-	-	-	-	-	-

PLR: Positive likelihood ratio; NLR: Negative likelihood ratio; DOR: Diagnostic odds ratio; AUC: Area under the curve; SVM: Support vector machine; CAD: Computer-aided diagnosis; CI: Confidence interval

chronic pancreatitis based on contrast-enhanced were 94.64% and 94.44%, respectively.

The type of AI and computer-aided diagnosis is not suggested as a source of heterogeneity in the meta-regression joint model. Compared with the machine learning algorithm of the artificial neural network, the SVM system used is much more suited to a limited number of training samples. With the progress of technology, CNN is considered to be an optimal method of deep learning for image recognition because it can obtain not only an image without a large increase in the amount of information but also universality with respect to image position movement.[31] The performance of AI models with deep learning and CNN as the backbone will be further improved in the future. AI systems need vast amounts of reliable patient data to implement deidentification. This may be one reason why studies with more patients have been more accurate in AI diagnosing pancreatic cancer. Year of publication and methodological quality were also statistically significant in the subgroup analysis. Technological advances may account for these differences.

Our analysis has several inevitable limitations originating from the potential bias in each study. First, all included studies were retrospective in nature, and potential bias in the case selection could not entirely be eliminated in this meta-analysis. Second, the control group had a single disease type. Most of the study groups had only CP (a few patients with lump-forming pancreatitis) and NP, and no patients had other pancreatic diseases. In clinical practice, pancreatic cancer is often difficult to distinguish from mass type chronic pancreatitis, pancreatic neuroendocrine neoplasms, and local autoimmune pancreatitis, but this is important because the treatment of each disease is fundamentally different. Third, when a certain learning model is over clipped to the training dataset and the prediction cannot be well generalized to the new dataset, there will be modeling errors, so the overfitting of the AI algorithm cannot be excluded.[32] Fourth, these studies were collected from a single center. When patients were randomized into the development group, validation group, and test group, there was a potential generalizability problem with the data. External (prospective) validation of the model using unused data sets is the best and only way to prove the true performance of an AI algorithm, and it should be collected in a way that minimizes the spectrum bias. Therefore, further prospective,

multicenter studies are needed to demonstrate the role of AI in clinical practice.

CONCLUSION

The accuracy of AI in the diagnosis of pancreatic cancer has good performance, and further prospective, multicenter, external validation studies are needed to confirm the role of AI-assisted EUS in clinical practice.

Supplementary materials

Supplementary information is linked to the online version of the paper on the *Endoscopic Ultrasound* website.

Financial support and sponsorship

This work was supported in part by National Natural Science Foundation of China Grants 82020108005 (to L.Z.S).

Conflicts of interest

Zhaoshen Li is an Honorary Editor-in-Chief of the journal. This article was subject to the journal's standard procedures, with peer review handled independently of this editor and his research groups.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, *et al*. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209-49.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020;70:7-30.
3. Wild CP, Weiderpass E, Stewart BW, editors (2020). World Cancer Report: Cancer Research for Cancer Prevention. Lyon, France: International Agency for Research on Cancer. 2020: pp 367-70.
4. Egawa S, Toma H, Ohigashi H, *et al.* Japan pancreatic cancer registry; 30th year anniversary: Japan pancreas society. *Pancreas* 2012;41:985-92.
5. Egawa S, Takeda K, Fukuyama S, *et al*. Clinicopathological aspects of small pancreatic cancer. *Pancreas* 2004;28:235-40.
6. Kitano M, Yoshida T, Itonaga M, *et al*. Impact of endoscopic ultrasonography on diagnosis of pancreatic cancer. *J Gastroenterol* 2019;54:19-32.
7. Müller MF, Meyenberger C, Bertschinger P, *et al*. Pancreatic tumors: Evaluation with endoscopic US, CT, and MR imaging. *Radiology* 1994;190:745-51.
8. Okusaka T, Nakamura M, Yoshida M, *et al*. Clinical practice guidelines for pancreatic cancer 2019 from the japan pancreas society: A synopsis. *Pancreas* 2020;49:326-35.
9. Legmann P, Vignaux O, Dousset B, *et al.* Pancreatic tumors: Comparison of dual-phase helical CT and endoscopic sonography. *AJR Am J Roentgenol* 1998;170:1315-22.
10. He XK, Ding Y, Sun LM. Contrast-enhanced endoscopic ultrasound for differential diagnosis of pancreatic cancer: An updated meta-analysis. *Oncotarget* 2017;8:66392-401.
11. Lu Y, Chen L, Li C, *et al*. Diagnostic utility of endoscopic ultrasonography-elastography in the evaluation of solid pancreatic masses: A meta-analysis and systematic review. *Med Ultrason* 2017;19:150-8.
12. Jiang Y, Inciardi MF, Edwards AV, *et al*. Interpretation time using a concurrent-read computer-aided detection system for automated breast ultrasound in breast cancer screening of women with dense breast tissue. *AJR Am J Roentgenol* 2018;211:452-61.
13. Abràmoff MD, Lavin PT, Birch M, *et al*. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018;1:39.
14. Mori Y, Kudo SE, Misawa M, *et al.* Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: A prospective study. *Ann Intern Med* 2018;169:357-66.
15. Goyal H, Mann R, Gandhi Z, *et al*. Scope of artificial intelligence in screening and diagnosis of colorectal cancer. *J Clin Med* 2020;9:E3313.
16. Bang CS, Lee JJ, Baik GH. Prediction of chronic atrophic gastritis and gastric neoplasms by serum pepsinogen assay: A systematic review and meta-analysis of diagnostic test accuracy. *J Clin Med* 2019;8:E657.
17. Whiting PF, Rutjes AW, Westwood ME, *et al.* QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-36.
18. Safari S, Baratloo A, Elfil M, *et al*. Evidence based emergency medicine; part 4: Pre-test and post-test probabilities and Fagan's nomogram. *Emerg (Tehran)* 2016;4:48-51.
19. Norton ID, Zheng Y, Wiersema MS, *et al.* Neural network analysis of EUS images to differentiate between pancreatic malignancy and pancreatitis. *Gastrointest Endosc* 2001;54:625-9.
20. 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;67:861-7.
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;72:978-85.
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. Ozkan M, Cakiroglu M, Kocaman O, *et al*. Age-based computer-aided diagnosis approach for pancreatic cancer on endoscopic ultrasound images. *Endosc Ultrasound* 2016;5:101-7.
24. Tonozuka R, Itoi T, Nagata N, *et al*. Deep learning analysis for the detection of pancreatic cancer on endosonographic images: A pilot study. *J Hepatobiliary Pancreat Sci* 2021;28:95-104.
25. Marya N, Powers P, Chari S, et al. Tu1604 Development of an EUS convolutional neural network algorithm for the differentiation of pancreatic ductal adenocarcinoma from benign conditions of the pancreas. Gastroenterology 2020;158:1134. [doi: 10.1016/s0016-5085(20) 33503-4.
26. Marya NB, Powers PD, Chari ST, *et al.* Utilisation of artificial intelligence for the development of an EUS-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut* 2021;70:1335-44.
27. Fritscher-Ravens A, Brand L, Knöfel WT, *et al.* Comparison of endoscopic ultrasound-guided fine needle aspiration for focal pancreatic lesions in patients with normal parenchyma and chronic pancreatitis. *Am J Gastroenterol* 2002;97:2768-75.
28. 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;68:1086-94.
29. 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* 2012;10:84-90.e81.
30. 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.
31. Munir K, Elahi H, Ayub A, *et al*. Cancer diagnosis using deep learning: A bibliographic review. *Cancers (Basel)* 2019;11:E1235.
32. Yang YJ, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019;25:1666-83.

SUPPLEMENTARY MATERIALS

Multimedia Appendix 1. Searching strategy to find the relevant articles

Database: MEDLINE (through PubMed)

#1 “artificial intelligence”[MeSH terms] OR “AI”[title/abstract] OR “deep learning”[title/abstract] OR “computational intelligence”[title/abstract] OR “computer aided”[title/abstract] OR “digital image”[title/abstract] OR “machine learning”[title/abstract] OR “neural network”[title/abstract] OR “CNN”[title/abstract]: 222,238
#2 “pancreatic neoplasms”[MeSH terms] OR “pancreatic mass”[title/abstract] OR “pancreatic cancer”[title/abstract] OR “pancreatic carcinoma”[title/abstract]: 93,699
#3 #1 AND #2: 483

Database: Embase

#1 “artificial intelligence”/exp OR ai: ti, ab, kw OR “deep learning”:ti, ab, kw OR “computational intelligence”:ti, ab, kw OR “computer aided”:ti, ab, kw OR “digital image”:ti, ab, kw OR “machine learning”:ti, ab, kw OR “neural network”:ti, ab, kw OR cnn: ti, ab, kw: 210496
#2 “pancreatic neoplasms”/exp OR “pancreatic mass”:ti, ab, kw OR “pancreatic cancer”:ti, ab, kw OR “pancreatic carcinoma”:ti, ab, kw: 171319
#3 #1 AND #2: 707

Database: Cochrane library

#1 “MeSH descriptor: [artificial intelligence] explode all trees OR ai: ti, ab, kw OR “deep learning”:ti, ab, kw OR “computational intelligence”:ti, ab, kw OR “computer aided”:ti, ab, kw OR “digital image”:ti, ab, kw OR “machine learning”:ti, ab, kw OR “neural network”:ti, ab, kw OR cnn: ti, ab, kw: 10306
#2 MeSH descriptor: [pancreatic neoplasms] explode all trees OR pancreatic mass’:ti, ab, kw OR “pancreatic cancer”:ti, ab, kw OR “pancreatic carcinoma”:ti, ab, kw: 5830
#3 #1 AND #2: 31