Systematic review and/or Meta-analysis

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Artificial intelligence assisted endoscopic ultrasound for detection of pancreatic space-occupying lesion: a systematic review and meta-analysis

Arkadeep Dhali, MBBS[a,*], Vincent Kipkorir, BS[c], Bahadar S. Srichawla, MBBS[d], Harendra Kumar, MBBS[e], Roger B. Rathna, MBBS[b], Ibsen Ongidi, BS, MBChB[c], Talha Chaudhry, BS[c], Gisore Morara, BS[c], Khulud Nurani, BS[c], Doreen Cheruto, MBChB[c], Jyotirmoy Biswas, MBBS[f], Leonard R. Chieng, MBBS[a], Gopal Krishna Dhali, MBBS, MD, DM[g]

Background: Diagnosing pancreatic lesions, including chronic pancreatitis, autoimmune pancreatitis, and pancreatic cancer, poses a challenge and, as a result, is time-consuming. To tackle this issue, artificial intelligence (AI) has been increasingly utilized over the years. AI can analyze large data sets with heightened accuracy, reduce interobserver variability, and can standardize the interpretation of radiologic and histopathologic lesions. Therefore, this study aims to review the use of AI in the detection and differentiation of pancreatic space-occupying lesions and to compare AI-assisted endoscopic ultrasound (EUS) with conventional EUS in terms of their detection capabilities.

Methods: Literature searches were conducted through PubMed/Medline, SCOPUS, and Embase to identify studies eligible for inclusion. Original articles, including observational studies, randomized control trials, systematic reviews, meta-analyses, and case series specifically focused on AI-assisted EUS in adults, were included. Data were extracted and pooled, and a meta-analysis was conducted using Meta-xl. For results exhibiting significant heterogeneity, a random-effects model was employed; otherwise, a fixed-effects model was utilized.

Results: A total of 21 studies were included in the review with four studies pooled for a meta-analysis. A pooled accuracy of 93.6% (CI 90.4–96.8%) was found using the random-effects model on four studies that showed significant heterogeneity ($P < 0.05$) in the Cochrane's Q test. Further, a pooled sensitivity of 93.9% (CI 92.4–95.3%) was found using a fixed-effects model on seven studies that showed no significant heterogeneity in the Cochrane's Q test. When it came to pooled specificity, a fixed-effects model was utilized in six studies that showed no significant heterogeneity in the Cochrane's Q test and determined as 93.1% (CI 90.7–95.4%). The pooled positive predictive value which was done using the random-effects model on six studies that showed significant heterogeneity was 91.6% (CI 87.3–95.8%). The pooled negative predictive value which was done using the random-effects model on six studies that showed significant heterogeneity was 93.6% (CI 90.4–96.8%).

Conclusion: AI-assisted EUS shows a high degree of accuracy in the detection and differentiation of pancreatic space-occupying lesions over conventional EUS. Its application may promote prompt and accurate diagnosis of pancreatic pathologies.

Keywords: artificial intelligence, diagnosis, endoscopic ultrasound, pancreatic lesion

[a]*NIHR Academic Clinical Fellow in Gastroenterology, University of Sheffield; Internal Medicine Trainee, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK,* [b]*University Hospitals Birmingham, Birmingham, UK,* [c]*School of Medicine, Faculty of Health Sciences, University of Nairobi, Nairobi, Kenya,* [d]*University of Massachusetts Chan Medical School, Worcester, MA,* [e]*Dow University of Health Sciences, Karachi, Pakistan,* [f]*College of Medicine and Sagore Dutta Hospitals and* [g]*School of Digestive and Liver Diseases, Institute of Postgraduate Medical Education and Research, Kolkata, India*

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article

**Corresponding author. Address: Sheffield Teaching Hospitals NHS Foundation Trust, S57AU, UK. E-mails: arkadipdhali@gmail.com; A.Dhali@sheffield.ac.uk (A. Dhali).*

International Journal of Surgery (2023) 109:4298–4308

Received 13 June 2023; Accepted 21 August 2023

Published online 4 October 2023

http://dx.doi.org/10.1097/JS9.0000000000000717

Introduction

The pancreas is a retroperitoneal organ that has both digestive and hormonal functions. Pathologies, including acute and chronic pancreatitis, autoimmune pancreatitis, and pancreatic cancer, affect the pancreas. These diseases are pretty lethal and have significant morbidity. For example, pancreatic cancer is the seventh leading cause of cancer-related deaths worldwide[1,2]. More so, pancreatic cancer's overall 5-year survival rate stands at 11.5%. In addition, diagnosing chronic pancreatitis, autoimmune pancreatitis, and pancreatic cancer is challenging as they closely resemble each other. This mimicry has led to late diagnosis of these diseases, affecting overall patient outcomes. Additionally, there is a considerable risk of confusion between autoimmune pancreatitis from pancreatic cancers, two pathologies with very different management strategies[3,4].

Imaging modalities for diagnosing pancreatic pathologies include computed tomography scans, MRIs, and endoscopies (EUS). EUS is the gold standard in diagnosing pancreatic pathologies due to its high specificity, sensitivity, and negative

predictive value. In differentiating between the three disease entities (pancreatic cancer, chronic pancreatitis, and autoimmune pancreatitis), cytological analysis is preferred. Due to this, EUS with fine needle aspirates or biopsies has been developed[1,5]. However, EUS/FNA/B requires additional training and is quite challenging due to its steep learning curve. The equipment cost and the need for anaesthesia make this procedure difficult. It also relies heavily on the operator leading to quite significant interobserver variability[1]. These disadvantages are substantial in resource-poor settings due to the scarcity of skilled personnel and the operating costs.

Artificial intelligence (AI) integrates computer systems and software designs to display the properties of critical thinking and intelligence. AI strives to replicate human intelligence with learning abilities and complex problem-solving skills[6]. As a result, AI has been incorporated into clinical practice with the advent of computer-aided diagnosis (CAD)[7]. There are three branches of AI beneficial to clinical practice. These are machine learning, deep learning, and expert systems. Over the years, a shift has shifted from machine learning to deep learning (artificial neuronal networks and convolutional neuronal networks), whose functioning resembles human neurophysiology[8].

Deep learning is a machine-learning technique miming the human neuronal network. It uses multiple layers of nonlinear processing units to abstract data hierarchically, extracting abstract features for tasks like target detection, classification, or segmentation. Artificial neural networks (ANNs) imitate the structure and functioning of biological neural networks. They consist of interconnected neurons organized into layers, learning from data and making predictions based on patterns. Convolutional neural networks (CNNs) are specialized ANNs for image recognition and processing. They excel at processing pixel data using convolutional layers, extracting features from local regions. By stacking these layers, CNNs capture local and global image information for tasks like image generation and description[3,8].

Expert systems, on the other hand, are designed to solve complex problems by utilizing reasoning based on existing knowledge. They aim to emulate human experts by capturing their expertise in a computer program. Expert systems typically consist of a knowledge base that stores relevant information and a reasoning engine that uses this knowledge to draw conclusions or make recommendations. The findings or decisions made by expert systems are often expressed as probabilities based on input data[1].

AI is slowly being incorporated into clinical practice since it has many benefits. It can analyze large data sets with increased accuracy, decrease interobserver variability, decrease the rate of misdiagnosis, and standardize the interpretation of radiologic and histopathologic lesions. These benefits have come in handy in aiding the diagnosis of pancreatic pathologies. A study by Marya *et al.*[3] depicted the benefits of utilizing convolutional neuronal networks in diagnosing autoimmune pancreatitis. Dahiya *et al.*[1] carried out a systematic review portraying the help of AI in diagnosing pancreatic cancer and differentiating it from chronic pancreatitis and autoimmune pancreatitis. The benefits of utilizing AI are valuable, especially in resource-poor settings, as it helps mitigate the number of gastroenterology centres by reducing the level of specialized knowledge needed to detect ambiguous results[7].

AI, however, comes with its disadvantages and shortcomings. These include inadequate standardization of input data used to train the AI algorithm. Currently, there is no standardized protocol; for data collection, processing, and storage for the AI-assisted model[7]. In addition to this, the quality of input data utilized is not optimum. Therefore, this can lead to selection bias since most of the input focuses on only a particular population[1]. There is also the issue of a black box whereby the user cannot interpret and determine the reasoning behind how a specific variable was weighed within the AI algorithm[1]. Lastly, there is the issue of ethics, where input data acquisition can prove challenging.

We reviewed the use of AI in the detection of pancreatic space-occupying lesions in our study. We also compared the AI-assisted EUS over conventional EUS in the detection of pancreatic space-occupying lesions and the efficacy of differentiating between different pancreatic pathologies.

HIGHLIGHTS

- Endoscopic ultrasound (EUS) is the gold standard in diagnosing pancreatic pathologies.
- It has high specificity, sensitivity, and negative predictive value.
- AI-assisted EUS shows a high degree of accuracy in the detection and differentiation of pancreatic space-occupying lesions over conventional EUS.

Methodology

Study protocol and registration

This systematic review and meta-analysis were conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) (Fig. 1)[9] and Assessing the methodological quality of systematic reviews

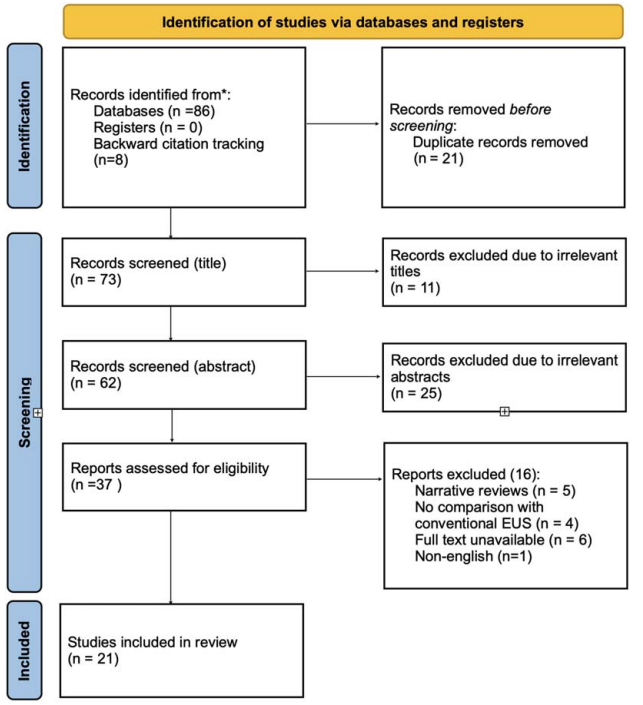

Figure 1. Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) flowchart for included studies. EUS, endoscopic ultrasound.

Table 1

Search strategy for the databases utilized in the study.

Database	Search strategy
Pubmed	((AI OR "artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network"OR "digital image analysis") AND ("pancreatic carcinoma" OR "pancreatic ca" OR "cystic neoplasm" OR SPEN OR "Solid pseudopapillary epithelial neoplasm" OR mass OR masses) AND (endoscopic OR endoscopy OR EUS) AND (ultrasound OR ultrasonography) AND (detection OR diagnosis OR diagnosing) AND (pancreas OR pancreatic))
Scopus	TITLE-ABS-KEY (((ai OR "artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "digital image analysis") AND ("pancreatic carcinoma" OR "pancreatic ca" OR "cystic neoplasm" OR spen OR "Solid pseudopapillary epithelial neoplasm" OR mass OR masses) AND (endoscopic OR endoscopy OR eus) AND (ultrasound OR ultrasonography) AND (detection OR diagnosis OR diagnosing) AND (pancreas OR pancreatic)))
Embase	((AI OR "artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network"OR "digital image analysis") AND ("pancreatic carcinoma" OR "pancreatic ca" OR "cystic neoplasm" OR SPEN OR "Solid pseudopapillary epithelial neoplasm" OR mass OR masses) AND (endoscopic OR endoscopy OR EUS) AND (ultrasound OR ultrasonography) AND (detection OR diagnosis OR diagnosing) AND (pancreas OR pancreatic))

AI, artificial intelligence; EUS, endoscopic ultrasound.

(AMSTAR)[10] guidelines. The protocol for the study was registered in the International Prospective Register of Systematic Reviews (PROSPERO).

Data sources and search strategy

Literature searches were performed through PubMed/Medline, SCOPUS, and Embasse to identify studies eligible for inclusion. All publications up to April 2023, the latest search date, were included. Search terms used for the three databases are outlined in Table 1. No restrictions on language, or study type were specified on the search protocol. The PubMed function 'related articles' was used to extend the search to provide a reference list of all included studies. A backward citation was used when appropriate to include pertinent articles. The following PICOS criteria were used as a framework to design the study question and formulate the literature search strategies to ensure comprehensive and bias-free searches:

P (Population): adults (> 18) with pancreatic lesions.

I (Intervention): AI-assisted endoscopic ultrasound.

C (Comparison): conventional endoscopic ultrasound.

O (Outcomes): detection of pancreatic carcinoma, cystic neoplasms, SPEN.

S (Studies): original articles (including observational studies, randomized control trials) systematic reviews, meta-analyses, and case series.

Eligibility criteria and screening of articles

Rayyan citation manager was used to facilitate the screening of articles obtained from the search process. Duplicate citations were cross-checked manually and removed after careful evaluation of the data. The title and abstract of the remaining articles were screened for relevance and full texts were obtained for those that passed the inclusion criteria. For repeat articles from the same group containing a search period overlap and similar data sets, only the most recent article was included to avoid duplication of data.

Studies were considered eligible for inclusion if they contained relevant information on the use of AI-assisted machine-learning algorithms in endoscopic ultrasound for the detection of pancreatic space-occupying lesions in adults. Particularly, the following criteria were used to establish the eligibility of studies. Inclusion criteria: original articles (including observational studies, randomized control trials) systematic reviews, meta-analyses, and case series specific to AI-assisted EUS in adults. Exclusion criteria: narrative reviews, editorials, short communications, case studies, and articles for which full text was not retrievable. Non-English articles were excluded at this stage, as were studies with incomplete or irrelevant information. Any disagreements about eligibility were settled through consensus.

Data extraction and outcomes of interest

All relevant articles that passed the screening and inclusion criteria were considered for analysis. Data extraction was conducted by two independent reviewers. Data extraction was done using a standard template based on the Cochrane Consumers and Communication Review group's extraction template for quality assessment and evidence synthesis.

From each study, the following information was extracted: Study characteristics: authors, database, journal, DOI, original title, full article abstract, publication year, country and continent, study design, sample size, and study period; Participant demographics: age, sex, and clinical characteristics (e.g. symptoms, risk

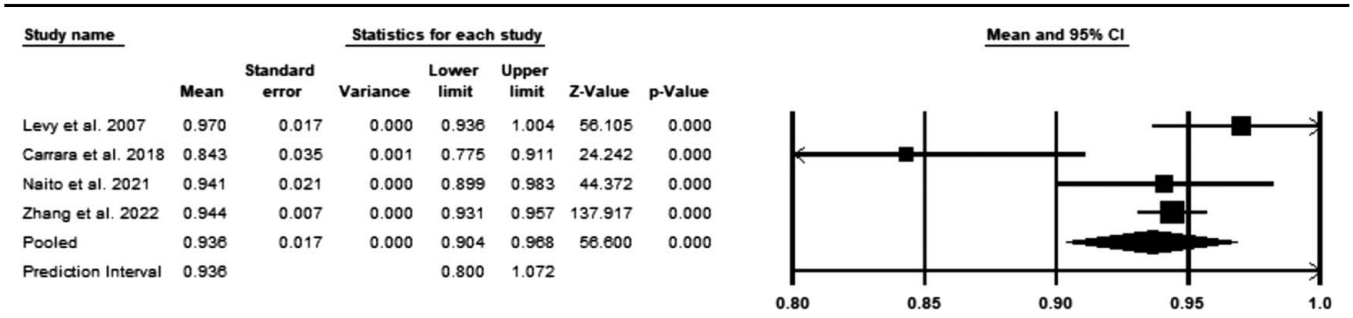

Figure 2. Pooled accuracy of studies.

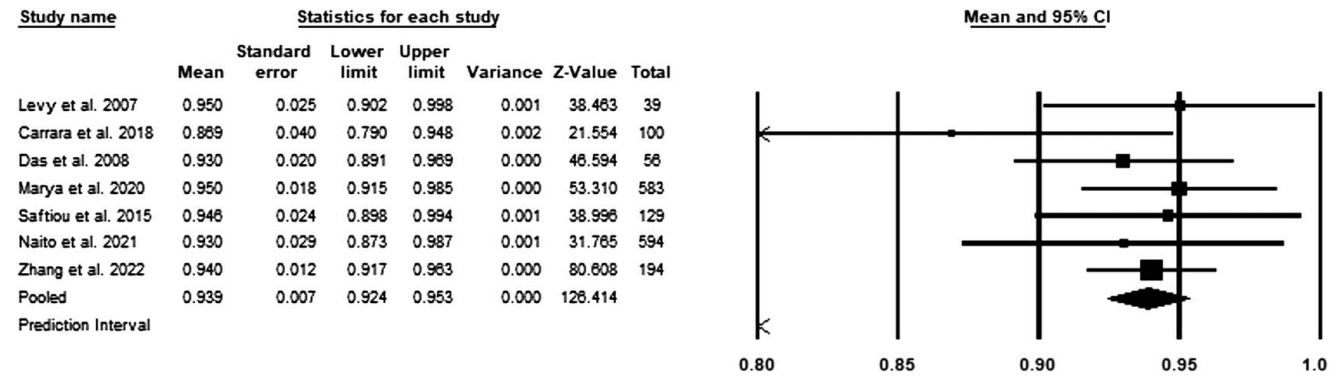

Figure 3. Pooled sensitivity of studies.

factors, and comorbidities); Intervention details: Description of the AI-assisted EUS system (e.g. type of algorithm, training data) and the standard EUS procedures; Outcome measures: Diagnostic accuracy (sensitivity, specificity, positive predictive value, and negative predictive value), procedure time, complications, and interobserver agreement

Data summary and synthesis

The data were entered in an Excel sheet for cleaning, validation, and coding. The information will be classified into pancreatic cancer detection, cystic neoplasm (including IPMN) detection, SPEN detection, and false negative rates. Studies containing other information other than the mentioned groups will be included in the miscellaneous category. Data were presented using a summary of findings table and variables assessed for their suitability for a meta-analysis. The extracted data were pooled, and a meta-analysis was performed for the appropriate variables, considering the clinical and methodological heterogeneity among the included studies. Meta-xl was used for the analysis of the data.

Meta-analysis of diagnostic test accuracy

Pooled accuracy, sensitivity (se), specificity (sp), positive predictive value, and negative predictive value were determined for all AI-assisted EUS procedures. Meta-analysis was conducted only on full-text articles that provided complete descriptive statistical data including confidence intervals. Forest plots with a 95% CI were calculated and pooled as well as pooled interval data were assessed. Heterogeneity among the outcomes of included studies in this meta-analysis was evaluated using Cochrane's Q test. Significant heterogeneity was indicated by P less than 0.05 in Cochrane's Q test. For results with significant heterogeneity,

a random-effects model was utilized. And those with non-significant heterogeneity a fixed-effects model was performed. Statistical analyses were performed using Python programming language *v3.4* (Python Software Foundation, Wilmington, Delaware). Data analysis and visualization were completed using Comprehensive Meta-Analysis *v4.0* (Biostat Inc.).

Risk of bias assessment

The quality of the included studies will be assessed using the appropriate tools for each study design. For observational studies, the Newcastle-Ottawa Scale will be used, while the Cochrane Risk of Bias tool (ROB2) will be employed for randomized controlled trials (RCTs). Two independent reviewers (V.K. and C.D.) will assess the quality of each study, with disagreements resolved through discussion or consultation with a third reviewer if necessary.

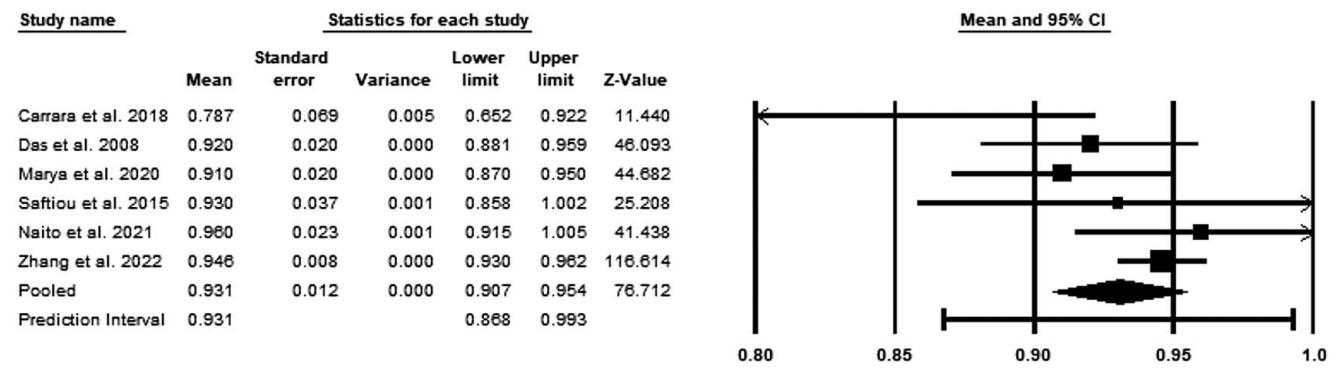

Figure 4. Pooled specificity of studies.

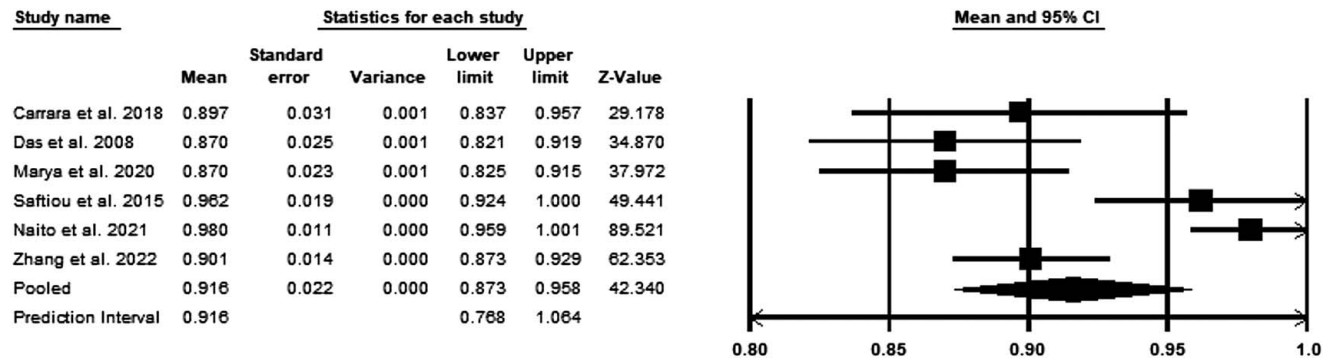

Figure 5. Pooled positive predictive value.

Results

Summary of study characteristics

A Total of 21 studies out of the 94 retrieved were included in the review after identification and screening using the PRISMA guidelines (Fig. 1). A random-effects model was utilized to assess the pooled accuracy. A total of 4 studies were included in the meta-analysis[11–14]. Significant heterogeneity was observed with a *P* less than 0.05 in Cochrane's Q test. The pooled accuracy was 93.6% (CI 90.4–96.8%) (Fig. 2).

A fixed-effects model was utilized to assess the pooled sensitivity. A total of seven studies were included in the meta-analysis[3,12–16]. No significant heterogeneity was observed with a *P* greater than 0.05 in the Cochrane's Q test. The pooled accuracy was 93.9% (CI 92.4–95.3%) (Fig. 3).

A fixed-effects model was utilized to assess the pooled specificity. A total of six studies were included in the meta-analysis[3,12–16]. No significant heterogeneity was observed with a *P* greater than 0.05 in the Cochrane's Q test. The pooled accuracy was 93.1% (CI 90.7–95.4%) (Fig. 4).

A random-effects model was utilized to assess the pooled positive predictive value. A total of six studies were included in the meta-analysis[3,12–16]. Significant heterogeneity was observed with a *P* less than 0.05 in Cochrane's Q test. The pooled accuracy was 91.6% (CI 87.3–95.8%) (Fig. 5).

A random-effects model was utilized to assess the pooled negative predictive value. A total of six studies were included in the meta-analysis[3,12–16]. Significant heterogeneity was observed with a *P* less than *0.05* in Cochrane's Q test. The pooled accuracy was 93.6% (CI 90.4–96.8%) (Fig. 6).

Quality assessment

Non-randomized studies were assessed using the RoBANS tool. Figure 7 provides a visual depiction of the risk of bias analysis of non-randomized trials. Overall, the risk of bias for non-randomized trials was low. The domain with the highest risk of bias was in the intervention (exposure) measurement while selective outcome reporting was the domain with the lowest risk of bias. The revised Cochrane risk-of-bias tool for randomized trials (RoB2) was used to assess the risk of bias in RCT and results are provided in Table 2. Some concerns in domains 2 and 4 were noted in 2 out of the five RCTs.

Discussion

Utilization of AI in endoscopic ultrasound for the detection of pancreatic space-occupying lesion (SOL)

Pancreatic masses consist of various types, including pancreatic adenosquamous carcinoma, pancreatic acinar cell carcinoma, metastatic pancreatic tumour, neoplasm, solid pseudopapillary neoplasm, as well as benign causes such as chronic and autoimmune pancreatitis (Table 3)[29].

EUS is an important diagnostic tool for pancreatic diseases, but its specificity for diagnosing pancreatic malignancies is limited, reaching as low as 58%[30]. Traditional EUS requires additional

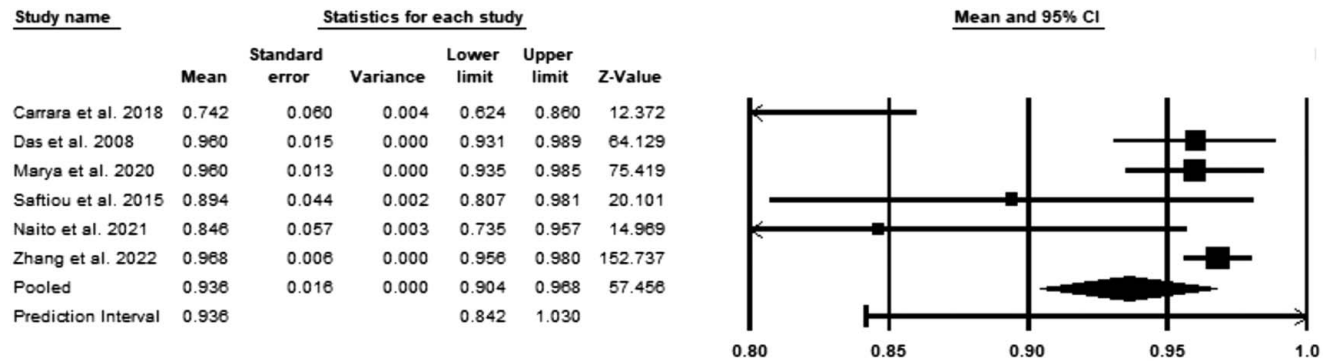

Figure 6. Pooled negative predictive value.

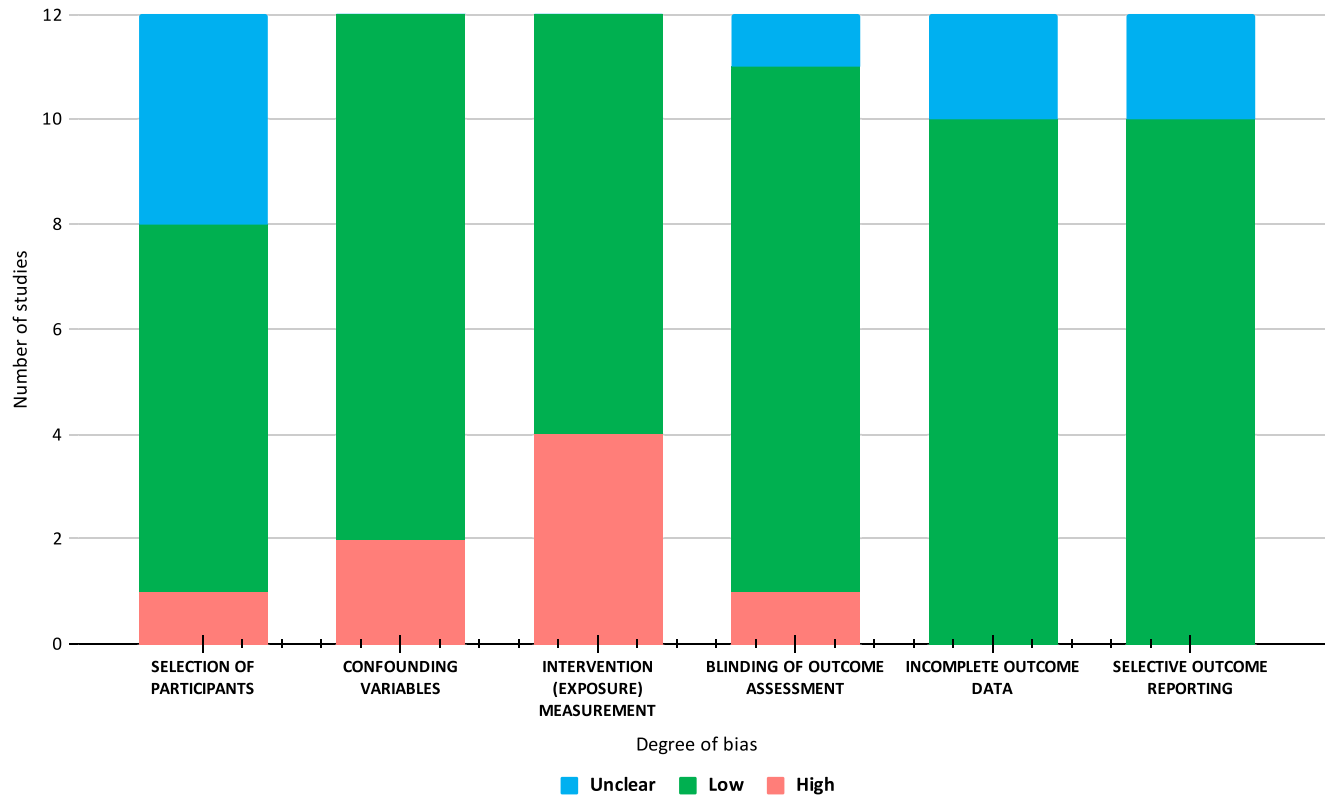

Figure 7. Risk of bias assessment output using the ROBANS tool.

training and hence is operator dependent leading to quite significant interobserver variability, more pronounced in resource-poor settings, due to the scarcity of skilled personnel and the high operating costs. AI-based EUS, however, shows significantly higher sensitivities and specificities of up to 0.93 and 0.78, respectively, with diagnostic odds ratio of 36.74 and area under the receiver operating characteristic curve of 0.94[31]. Other studies comparing traditional and AI-assisted EUS further report superior performance of AI with sensitivities of 0.93 versus 0.71, specificities of 0.81 versus 0.69, and area under the curve of 0.94 versus 0.75, respectively[32].

Standard EUS has limitations in diagnosing pancreatic malignancies, such as low specificity and operator dependence, leading to increased interest in AI-assisted EUS. It has been shown that AI-based EUS improves diagnostic accuracy and reduces interobserver variability. Studies have shown that AI algorithms are capable of achieving significantly higher sensitivity and specificity than traditional EUS, with diagnostic odds ratios and area under the receiver operating characteristic curve indicating superior performance[33,34]. Advances in diagnostic capabilities could revolutionize pancreatic lesion detection and diagnosis, especially in settings lacking skilled personnel and resources. EUS can be

Table 2

Risk of bias assessment output using the ROB2 tool.

Study ID	Experimental	Comparator	Outcome	Weight	D1	D2	D3	D4	D1	Overall
Carrara *et al.*[12]	Fractal-based quantitative analysis	EUS elastography	Differentiation of SPL	1	+	!	+	+	+	+
Săftoiu *et al.*[17]	Real-time EUS elastography using CAD by artificial neural network analysis	Positive Cytology	Accuracy	1	+	+	+	+	+	+
Tang *et al.*[14]	Images with AI annotation	Images without AI annotation	Accuracy of CH-EUS diagnosis system	1	+	+	+	+	+	+
Tang *et al.*[14]	Contrast-enhanced harmonic endoscopic ultrasound MASTER	Histopathology	Accuracy of Contrast-enhanced harmonic endoscopic ultrasound MASTER	1	+	+	+	+	+	+
Zhu *et al.*[18]	Computer-Aided Diagnosis of EUS images	Positive Cytology	Accuracy in differentiation of pancreatic cancer (PC), chronic pancreatitis (CP).	1	+	+	+	!	+	+

D1: Randomisation process, D2: Deviations from the intended interventions, D3: Missing outcome data, D4: Measurement of the outcome, D5: Selection of the reported result, +: Low risk, !: Some concern, –: High risk.

AI, artificial intelligence; CP, chronic pancreatitis CAD, computer-aided diagnosis; EUS, endoscopic ultrasound; PC, pancreatic cancer.

Table 3

Comprehensive review of the studies included in our review.

Author, year	Study design	Sample size (*n*)	Image type	Type of algorithm	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
Lee *et al.*, 2023[19]	Retrospective cross-sectional	22 424 nCLE video frames (50 videos) as the training/validation set and 11 047 nCLE video frames (18 videos) as the test set	NR	Deep learning algorithm, U-Net. Deep learning algorithm, VGG19	NR	46	94.3	CNN1 Pseudocyst 70.1, CNN2 Pseudocyst 43.69, CNN3 Pseudocyst 39.50	CNN1 Pseudocyst 93.26, CNN2 Pseudocyst 83.24, CNN3 Pseudocyst 84.38
Levy *et al.*, 2007[11]	Cohort	39	EUS images	digital image analysis and fluorescence in situ hybridization	98 (93–100)	97 (90–100)	100 (100–100)	No false-positive results occurred for DIA or FISH.	1 failed diagnosis for DIA/FISH in a patient with a malignant GI stromal tumour.
Carrara *et al*, 2018[12]	RCT	100	EUS elastography	Fractal-based quantitative analysis	84.31 (76.47–90.20)	86.96 (78.26–94.20)	78.79 (63.64–90.91)	89.71 (83.10–95.38)	74.29 (62.86–86.67)
Das, 2008[15]	Retrospective, cross-sectional	*n*=56; 11 099 images	EUS images	Neural network	100	93 (89-97)	92 (88-96)	87 (82–92)	96 (93-99)
Marya, 2020[3]	Cohort	*n*=583; 1 174 461 (EUS images), 955 (EUS frames per second) (video data)	EUS images/ videos	Neural network	NR	95 (91-98)	91 (86-94)	87 (82-91)	97 (93-98)
Norton *et al.*, 2001[20]	Retrospective, cross-sectional	35	EUS images	Neural network	80	100	50	75	100
Ozkan *et al.*, 2016[21]	Retrospective, cross-sectional	*n*=172; 332 images (202 cancer and 130 noncancer)	EUS images	Neural network	87.5±0.04	83.3±0.11	93.33±0.07	NR	NR
Săftoiu *et al.*, 2008[22]	Prospective, cross-sectional	68	EUS elastography	Neural network	89.70	91.40	87.90	88.90	90.60
Săftoiu *et al.*, 2012[17]	RCT	*n*=258; 774 images	EUS elastography	Neural network	84.27(83.09-85.44)	87.59	82.94	96.25	57.22
Saftoiu, 2015[16]	Prospective, observational trial	*n*= 129; 167 videos	Contrast-enhanced harmonic EUS	Neural network	NR	94.64 (88.22-97.8)	94.44 (83.93-98.58)	97.24 (91.57-99.28)	89.47 (78.165-95.72)
Tonozuka *et al.*, 2020[23]	Prospective, cross-sectional	*n*= 139; 920 images (endosonographic images), 470 (images were independently tested)	EUS images	Neural network	NR	92.40	84.10	86.80	90.70
Zhang *et al.*, 2010[24]	retrospective cross-sectional	216	EUS images	SVM support vector machine.	97.98±1.23	94.32±0.03	99.45±0.01	98.65±0.02	97.77±0.01
Zhu *et al.*, 2013[18]	RCT	388	EUS images	SVM	94.20±0.17	96.25±0.4	93.38±0.2	92.21±0.42	96.68±0.14
Naito 2021[13]	Retrospective cross-sectional	594	NR	deep learning model	94.17(89.17–97.5)	93.02(86.02–975.3)	97.06(90.91–100)	98.77 (95.71–100)	84.62 (72.97–95.12)
Nguon *et al.* 2021[25]	Cross-sectional	47 MCN and 31 SCN patients at the 1st hospital and 13 MCN and 18 SCN patients at the 2nd hospital. MCN, SCN	EUS images	deep learning network model.	82.76	81.46	84.36	NR	NR
Tang, 2023[26]	RCT	4530 images and 270 videos	Contrast-enhanced harmonic EUS	deep learning	93.80	90.90	100	100	83.30

Tang 2023[26]	RCT	39	Contrast-enhanced harmonic EUS	deep convolutional neural networks and Random Forest algorithm	93.80	90.90	100	100		83.30
Udristoiu 2021[27]	Cross-sectional	*n*=65, 1300 images	NR	machine-learning algorithms: RMSProp optimization algorithm	98.26	98.6	97.4	98.7		97.4
Vilas boas 2022[28]	Retrospective crosssectional	*n*= 28; 5505 images	EUS images and videos	convolutional neural network	98.50	98.3	98.90	99.50		96.4
Zhang,2022[14]	Retrospective crosssectional	*n*=194, 5345 cytopathological slide images	EUS images	Deep convolutional neural network	94.4 (92.9–95.6)	94.0 (91.7–96.3)	94.6 (93.0–96.2)	90.1 (87.3–93.0)		96.8, (95.5–98.0)

CNN, convolutional neural network; DIA, Digital image analysis; EUS, endoscopic ultrasound; FISH, fluorescence in situ hybridization; MCN, mucinous cystic neoplasm; NR, not reported; nCLE, needle-based confocal laser endomicroscopy; RCT, randomized controlled trial; SCN, serous cystic neoplasm; SVM, support vector machine.

further optimized with AI to provide more accurate and reliable assessments, leading to improved patient outcomes[35].

AI is a mathematical technique used for classification or regression, and deep learning, which is an advanced machine-learning method utilizing neural networks, falls under the category of AI algorithms[33,34]. AI has been successfully applied in the detection and classification of various tumours, such as oesophageal tumours[34], gastric tumours[33], colon polyps[34], and subepithelial lesions[35]. Its potential for improving the diagnostic accuracy of pancreatic masses has also been explored.

In our study, we observed a significant improvement in the overall diagnostic accuracy when using AI assistance for diagnosing pancreatic masses compared to conventional EUS, achieving a rate of 93%. However, there were no significant differences in sensitivity and specificity. Additionally, AI-assisted EUS showed better positive and negative predictive values compared to conventional EUS.

The development of an AI system that can accurately diagnose pancreatic masses may have the potential to replace EUS-FNA/B in the future, reducing adverse events and decreasing dependence on operator expertise in diagnosing pancreatic masses. However, according to Kuwahara *et al*., the current 90% accuracy of AI may not be high enough to fully replace EUS-FNA, but it can still be valuable in diagnosing pancreatic masses[36,37].

Utilization of AI in endoscopic ultrasound to differentiate pancreatic SOL from chronic pancreatitis

Marya *et al*.[3] in 2020, demonstrated that an AI model using a convolutional neural network on EUS images effectively differentiated chronic pancreatitis from other pancreatic masses. The model achieved a sensitivity and specificity of 81% and had an area under receiver operating characteristic curve of 0.847 (95% CI 0.770 to 0.911) when distinguishing chronic pancreatitis from all other pancreatic masses.

Utilization of AI in endoscopic ultrasound to differentiate pancreatic SOL from autoimmune pancreatitis

AI has been used in endoscopic ultrasound (EUS) to aid in distinguishing pancreatic SOL from autoimmune pancreatitis. Distinguishing autoimmune pancreatitis from other SOLs is particularly challenging due to overlapping clinical and radiological features. Misdiagnosis can lead to unnecessary interventions, delayed treatment, or inappropriate management plans. However, this differentiation is crucial since the treatment approaches for these diseases vary significantly.

Marya *et al*.[3] 2020 developed an AI model using CNN on EUS images that effectively differentiated autoimmune pancreatitis from pancreatic ductal adenocarcinoma. The AI model achieved a sensitivity of 93%, a specificity of 90%, and an area under the receiver operating characteristic curve of 0.95. The sensitivity and specificity of the AI model were significantly higher compared to conventional diagnostic methods. This suggests that AI can serve as a valuable adjunct tool to endoscopy in making more accurate and timely diagnoses, with the potential for better patient outcomes.

The utilization of AI in EUS can enhance the efficiency and workflow of endoscopy units. With the increasing demand for EUS procedures, AI can assist in streamlining the interpretation process and reduce the burden on practitioners. By providing rapid and accurate analyses of EUS images, AI can save time and

resources, allowing clinicians to focus on patient care and complex decision-making.

Limitations of AI in endoscopic ultrasound for the detection of pancreatic SOL

Despite the promising results and potential benefits of AI in EUS, there are several challenges and limitations that need to be addressed. First, the development and validation of AI models require large and diverse data sets that adequately represent the target population. The availability of such data sets can be a limitation, particularly for rare pancreatic conditions. Most studies that have been done are retrospective and use a limited number of images from single-centre studies[19,26,36–38]. Due to limited data availability and concerns related to model over-fitting, the effectiveness of these systems remains insufficient, emphasizing the need for comprehensive external validation[39]. Collaborative efforts and data sharing among hospitals and academic institutions are necessary to overcome this limitation and ensure the robustness and generalizability of AI models. Robust multicenter trials are necessary to increase the sample size and increase the clinical significance of study results. Well-designed randomized controlled trials (RCTs) are warranted to provide higher-quality evidence and enhance the level of confidence in the findings. The inclusion of more RCTs in future systematic reviews and meta-analyses would contribute to a more robust evidence base and further elucidate the potential benefits of AI-assisted EUS in clinical practice.

Second, the diagnosis performance of AI algorithms may be limited in data sets where there is heterogeneity of image contents[36]. Although deep learning models can achieve high accuracy, there is potential for selection bias and misclassification resulting in suboptimal performance of CNNs. Additionally, there is a lack of studies that perform external validation of the AI models used in the EUS of the pancreas. In the absence of external validation, there is a lack of assurance regarding the model's generalizability, which may result in the possibility of overestimating the outcomes[27,28,40]. Efforts are underway to develop techniques and methods that enhance the reliability and interpretability of AI models, allowing technicians and clinicians to understand and trust the results generated by AI algorithms.

Lastly, the integration of AI into clinical workflows requires careful consideration of ethical, legal, and regulatory aspects. The use of real-time training for learning models has been difficult due to the possibility of ethical and safety issues[41,42]. Data privacy and patient consent are critical concerns that need to be addressed before adopting the use of AI in clinical practice. Transparent guidelines and regulations should be established to govern the use of AI in healthcare and ensure its responsible and ethical implementation.

Future directions of AI in endoscopic ultrasound for pancreatic SOL

The use of AI in clinical practice is still in its preliminary stages. There is a lot of promise in utilizing AI and incorporating computers in aiding clinical diagnosis. Various AI algorithms can be used as a second set of eyes by specialists to diagnose multiple pancreatic pathologies. AI-assisted EUS (Fig. 8), from our study, has shown to have higher diagnostic accuracy than conventional EUS. However, both methods do not have a diagnostic accuracy of 100%. However, AI can be used by specialists, and in so doing, it can help reduce interobserver variability while also learning from them. Over time, standardization in diagnosis can be achieved, improving patient outcomes[1].

AI can also be used to augment other diagnostic techniques. For example, in endoscopies, AI can augment capsule endoscopies, increasing their efficiency. In diagnosing pancreatic pathologies, AI-assisted algorithms can merge the use of different imaging modalities such as computed tomography scans, MRI and EUS. Lastly, AI algorithms can be utilized to interpret biomarkers and diagnostic enzymology to differentiate further

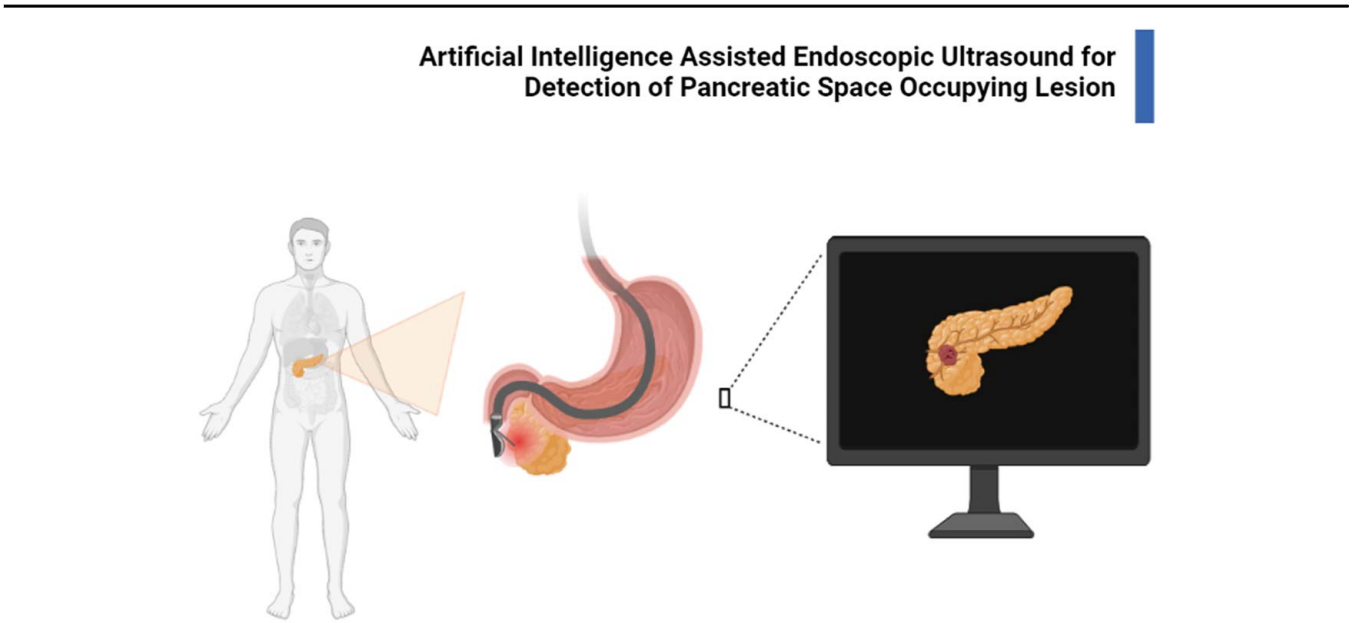

Figure 8. An illustration of the use artificial intelligence assisted endoscopic ultrasound for detection of pancreatic space-occupying lesion.

chronic pancreatitis, autoimmune pancreatitis, and pancreatic cancer[7].

Developing an expert system in clinical practice will go a long way in improving patient outcomes. Expert systems solve problems with reasoning based on current knowledge, emulating a human expert. This algorithm can also draw conclusions based on the input data. Incorporation of this into diagnosing pancreatic pathologies can help augment the current challenges as well as augment the scarcity of skilled personnel[1].

Conclusion

AI-assisted EUS has emerged as a highly accurate method for detecting and differentiating pancreatic space-occupying lesions, surpassing the capabilities of conventional EUS. By leveraging advanced computational algorithms, AI enables clinicians to achieve a prompt and precise diagnosis of various pancreatic pathologies. The integration of AI in EUS holds great promise in revolutionizing the field of pancreatic imaging, enhancing the efficiency of diagnostic workflows, and ultimately improving patient outcomes. However, the current meta-analysis is limited based on the few studies included. Future studies including high-quality RCTs and implementation of AI-assisted EUS in clinical practice can potentially unlock new avenues for early detection, personalized treatment strategies, and improved prognostication in pancreatic diseases.

Ethical approval statement

Ethics approval was not required for this review.

Consent

Informed consent was not required for this review.

Source of funding

None.

Author contribution

Conceptualization: A.D. Data curation: A.D., V.K., D.C., T.C. Data analysis: B.S. Manuscript Writing: A.D., H.K., V.K., B.S., R.B.R, J.B., G.M., K.N., L.R.C., G.K.D. All authors agreed to the final version of the manuscript.

Conflicts of interest disclosure

None.

Research registration unique identifying number (UIN)

The protocol for the study was registered in the International Prospective Register of Systematic Reviews (PROSPERO) and assigned registration number CRD42023416731.

Guarantor

Prof. (Dr.) Gopal Krishna Dhali acts as guarantor of the article.

Data statement

No primary data were generated in this systematic review article.

Peer review

Not commissioned. Externally peer-reviewed.

References

[1] Dahiya DS, Al-Haddad M, Chandan S, *et al*. Artificial intelligence in endoscopic ultrasound for pancreatic cancer: where are we now and what does the future entail? J Clin Med 2022;11:7476.
[2] 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.
[3] 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 2020; 70:1335–44.
[4] Miura F, Takada T, Amano H, *et al*. Diagnosis of pancreatic cancer. HPB 2006;8:337–42.
[5] Iglesias-Garcia J, Poley J-W, Larghi A, *et al*. Feasibility and yield of a new EUS histology needle: results from a multicenter, pooled, cohort study. Gastrointest Endosc 2011;73:1189–96.
[6] Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. J Glob Health 2018;8:020303.
[7] Mehta A, Kumar H, Yazji K, *et al*. Effectiveness of artificial intelligence-assisted colonoscopy in early diagnosis of colorectal cancer: a systematic review. Int J Surg 2023;109:946–52.
[8] Cai L, Gao J, Zhao D. A review of the application of deep learning in medical image classification and segmentation. Ann Transl Med 2020;8:713.
[9] Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg 2021;88: 105906.
[10] Shea BJ, Reeves BC, Wells G, *et al*. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:j4008. doi:10.1136/bmj.j4008. PMID: 28935701; PMCID: PMC5833365.
[11] Levy MJ, Clain JE, Clayton A, *et al*. Preliminary experience comparing routine cytology results with the composite results of digital image analysis and fluorescence in situ hybridization in patients undergoing EUS-guided FNA. Gastrointest Endosc 2007;66:483–90.
[12] Carrara S, Di Leo M, Grizzi F, *et al*. EUS elastography (strain ratio) and fractal-based quantitative analysis for the diagnosis of solid pancreatic lesions. Gastrointest Endosc 2018;87:1464–73.
[13] Naito Y, Tsuneki M, Fukushima N, *et al*. A deep learning model to detect pancreatic ductal adenocarcinoma on endoscopic ultrasound-guided fine-needle biopsy. Sci Rep 2021;11:8454.
[14] Zhang S, Zhou Y, Tang D, *et al*. A deep learning-based segmentation system for rapid onsite cytologic pathology evaluation of pancreatic masses: a retrospective, multicenter, diagnostic study. EBioMedicine 2022;80:104022.
[15] 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.
[16] 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.
[17] Nagtegaal ID, Odze RD, Klimstra D, *et al*. The 2019 WHO classification of tumours of the digestive system. Histopathology 2020;76:182–8.
[18] Nakai Y, Takahara N, Mizuno S, *et al*. Current status of endoscopic ultrasound techniques for pancreatic neoplasms. Clin Endosc 2019;52: 527–32.
[19] Ye XH, Zhao LL, Wang L. Diagnostic accuracy of endoscopic ultrasound with artificial intelligence for gastrointestinal stromal tumors: a meta-analysis. J Digest Dis 2022;23:253–61.

[20] Liu XY, Song W, Mao T, *et al*. Application of artificial intelligence in the diagnosis of subepithelial lesions using endoscopic ultrasonography: a systematic review and meta-analysis. Front Oncol 2022;12:915481.
[21] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44.
[22] Horie Y, Yoshio T, Aoyama K, *et al*. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. Gastrointest Endosc 2019;89:25–32.
[23] Hirasawa T, Aoyama K, Tanimoto T, *et al*. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. Gastric Cancer 2018;21:653–60.
[24] Byrne MF, Chapados N, Soudan F, *et al*. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. Gut 2019;68:94–100.
[25] Hirai K, Kuwahara T, Furukawa K, *et al*. Artificial intelligence-based diagnosis of upper gastrointestinal subepithelial lesions on endoscopic ultrasonography images. Gastric Cancer 2022;25:382–91.
[26] Kuwahara T, Hara K, Mizuno N, *et al*. Artificial intelligence using deep learning analysis of endoscopic ultrasonography images for the differential diagnosis of pancreatic masses. Endoscopy 2023;55:140–9.
[27] Lee T-C, Angelina CL, Kongkam P, *et al*. Deep-learning-enabled computer-aided diagnosis in the classification of pancreatic cystic lesions on confocal laser endomicroscopy. Diagnostics 2023;13:1289.
[28] Tang A, Gong P, Fang N, *et al*. Endoscopic ultrasound diagnosis system based on deep learning in images capture and segmentation training of solid pancreatic masses. Med Phys 2023;50:4197–205.
[29] 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.
[30] Udriştoiu AL, Cazacu IM, Gruionu LG, *et al*. Real-time computer-aided diagnosis of focal pancreatic masses from endoscopic ultrasound imaging based on a hybrid convolutional and long short-term memory neural network model. PLOS ONE 2021;16:e0251701.
[31] Vilas-Boas F, Ribeiro T, Afonso J, *et al*. Deep learning for automatic differentiation of mucinous versus non-mucinous pancreatic cystic lesions: a pilot study. Diagnostics 2022;12:2041.
[32] Qureshi TA, Javed S, Sarmadi T, *et al*. Artificial intelligence and imaging for risk prediction of pancreatic cancer: a narrative review. Chin Clin Oncol 2022;11:1.
[33] Dumitrescu EA, Ungureanu BS, Cazacu IM, *et al*. Diagnostic value of artificial intelligence-assisted endoscopic ultrasound for pancreatic cancer: a systematic review and meta-analysis. Diagnostics 2022;12:309.
[34] Mohan B, Facciorusso A, Khan S, *et al*. Pooled diagnostic parameters of artificial intelligence in EUS image analysis of the pancreas: a descriptive quantitative review. Endosc Ultrasound 2022;11:156.
[35] Hirasawa T, Aoyama K, Tanimoto T, *et al*. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. Gastric Cancer 2018; 21:653–60.
[36] Byrne MF, Chapados N, Soudan F, *et al*. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. Gut 2019;68:94–100.
[37] Hirai K, Kuwahara T, Furukawa K, *et al*. Artificial intelligence-based diagnosis of upper gastrointestinal subepithelial lesions on endoscopic ultrasonography images. Gastric Cancer 2022;25:382–91.
[38] Kuwahara T, Hara K, Mizuno N, *et al*. Artificial intelligence using deep learning analysis of endoscopic ultrasonography images for the differential diagnosis of pancreatic masses. Endoscopy 2023; 55:140–9.
[39] 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.
[40] Qureshi TA, Javed S, Sarmadi T, *et al*. Artificial intelligence and imaging for risk prediction of pancreatic cancer: a narrative review. Chin Clin Oncol 2022;11:1–.
[41] Dumitrescu EA, Ungureanu BS, Cazacu IM, *et al*. Diagnostic value of artificial intelligence-assisted endoscopic ultrasound for pancreatic cancer: a systematic review and meta-analysis. Diagnostics 2022;12:309.
[42] Mohan B, Facciorusso A, Khan S, *et al*. Pooled diagnostic parameters of artificial intelligence in EUS image analysis of the pancreas: a descriptive quantitative review. Endosc Ultrasound 2022;11:156.