



Endoscopic ultrasound artificial intelligence-assisted for prediction of gastrointestinal stromal tumors diagnosis: A systematic review and meta-analysis

Rômulo Sérgio Araújo Gomes, Guilherme Henrique Peixoto de Oliveira, Diogo Turiani Hourneaux de Moura, Ana Paula Samy Tanaka Kotinda, Carolina Ogawa Matsubayashi, Bruno Salomão Hirsch, Matheus de Oliveira Veras, João Guilherme Ribeiro Jordão Sasso, Roberto Paolo Trasolini, Wanderley Marques Bernardo, Eduardo Guimarães Hourneaux de Moura

Specialty type: Gastroenterology and hepatology

Provenance and peer review: Invited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): B
Grade C (Good): C, C
Grade D (Fair): 0
Grade E (Poor): 0

P-Reviewer: Mijwil MM, Iraq;
Naganuma H, Japan

Received: March 16, 2023

Peer-review started: March 16, 2023

First decision: April 20, 2023

Revised: June 15, 2023

Accepted: July 24, 2023

Article in press: July 24, 2023

Published online: August 16, 2023



Rômulo Sérgio Araújo Gomes, Guilherme Henrique Peixoto de Oliveira, Diogo Turiani Hourneaux de Moura, Ana Paula Samy Tanaka Kotinda, Carolina Ogawa Matsubayashi, Bruno Salomão Hirsch, Matheus de Oliveira Veras, João Guilherme Ribeiro Jordão Sasso, Wanderley Marques Bernardo, Eduardo Guimarães Hourneaux de Moura, Department of Gastroenterology, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo 05403-010, Brazil

Roberto Paolo Trasolini, Division of Hepatology and Endoscopy, Department of Gastroenterology, Harvard Medical School, Brigham and Women's Hospital, Boston, MA 02115, United States

Corresponding author: Guilherme Henrique Peixoto de Oliveira, MD, Doctor, Medical Assistant, Department of Gastroenterology, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, Dr Enéas de Carvalho Aguiar, 225, 6o Andar, Bloco 3, Cerqueira Cesar ZIP, São Paulo 05403-010, Brazil. dr.guilhermehpoliveira@gmail.com

Abstract

BACKGROUND

Subepithelial lesions (SELs) are gastrointestinal tumors with heterogeneous malignant potential. Endoscopic ultrasonography (EUS) is the leading method for evaluation, but without histopathological analysis, precise differentiation of SEL risk is limited. Artificial intelligence (AI) is a promising aid for the diagnosis of gastrointestinal lesions in the absence of histopathology.

AIM

To determine the diagnostic accuracy of AI-assisted EUS in diagnosing SELs, especially lesions originating from the muscularis propria layer.

METHODS

Electronic databases including PubMed, EMBASE, and Cochrane Library were searched. Patients of any sex and > 18 years, with SELs assessed by EUS AI-assisted, with previous histopathological diagnosis, and presented sufficient data values which were extracted to construct a 2 × 2 table. The reference standard was

histopathology. The primary outcome was the accuracy of AI for gastrointestinal stromal tumor (GIST). Secondary outcomes were AI-assisted EUS diagnosis for GIST *vs* gastrointestinal leiomyoma (GIL), the diagnostic performance of experienced endoscopists for GIST, and GIST *vs* GIL. Pooled sensitivity, specificity, positive, and negative predictive values were calculated. The corresponding summary receiver operating characteristic curve and post-test probability were also analyzed.

RESULTS

Eight retrospective studies with a total of 2355 patients and 44154 images were included in this meta-analysis. The AI-assisted EUS for GIST diagnosis showed a sensitivity of 92% [95% confidence interval (CI): 0.89-0.95; $P < 0.01$], specificity of 80% (95%CI: 0.75-0.85; $P < 0.01$), and area under the curve (AUC) of 0.949. For diagnosis of GIST *vs* GIL by AI-assisted EUS, specificity was 90% (95%CI: 0.88-0.95; $P = 0.02$) and AUC of 0.966. The experienced endoscopists' values were sensitivity of 72% (95%CI: 0.67-0.76; $P < 0.01$), specificity of 70% (95%CI: 0.64-0.76; $P < 0.01$), and AUC of 0.777 for GIST. Evaluating GIST *vs* GIL, the experts achieved a sensitivity of 73% (95%CI: 0.65-0.80; $P < 0.01$) and an AUC of 0.819.

CONCLUSION

AI-assisted EUS has high diagnostic accuracy for fourth-layer SELs, especially for GIST, demonstrating superiority compared to experienced endoscopists' and improving their diagnostic performance in the absence of invasive procedures.

Key Words: Subepithelial lesions; Ultrasound endoscopy; Artificial intelligence

©The Author(s) 2023. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Artificial intelligence (AI) has shown itself as a promising tool in diagnostic endoscopic ultrasound. This systematic review and meta-analysis analyze the diagnostic performance of endoscopy ultrasound with AI for subepithelial lesions and compare it with experienced endoscopists. Based on our meta-analysis, the endoscopy ultrasound assisted for AI has high diagnostic accuracy with superiority over experienced endoscopists.

Citation: Gomes RSA, de Oliveira GHP, de Moura DTH, Kotinda APST, Matsubayashi CO, Hirsch BS, Veras MO, Ribeiro Jordão Sasso JG, Trasolini RP, Bernardo WM, de Moura EGH. Endoscopic ultrasound artificial intelligence-assisted for prediction of gastrointestinal stromal tumors diagnosis: A systematic review and meta-analysis. *World J Gastrointest Endosc* 2023; 15(8): 528-539

URL: <https://www.wjgnet.com/1948-5190/full/v15/i8/528.htm>

DOI: <https://dx.doi.org/10.4253/wjge.v15.i8.528>

INTRODUCTION

Gastrointestinal subepithelial lesions (SELs) are tumors that originate from the muscular mucosa, submucosa, or muscular propria, with the stomach being the most common location where they are identified[1]. Although most SELs are benign and asymptomatic at presentation, up to 15% present malignant potential and may cause symptoms such as bleeding and abdominal pain[1,2]. The most common histological types are gastrointestinal stromal tumor (GIST) and gastrointestinal leiomyoma (GIL), with GIST having malignant potential[3]. One major diagnostic challenge is differentiating between GIST and leiomyoma considering that both commonly originate from the muscular propria and have overlapping features on imaging evaluation[4]. The differentiation among them is imperative due to the difference in prognosis and therapeutic strategy[5,6]. Surgical resection is recommended after GISTs diagnosis due to the risk of malignancy and requires prior histological confirmation, even in small lesions[1].

Endoscopic ultrasonography (EUS) is a valuable tool for SELs because it can characterize them by size, vascularity, internal structure, location, echogenicity, shape, and the layer of origin[6,7]. However, the gold standard for diagnosis is histopathological evaluation, which is indicated in suspected GIST, size > 20 mm, high-risk malignancy, surgical indication, or oncological treatment[1]. In uncertain cases, auxiliary procedures such as fine needle aspiration or fine needle biopsy can be performed for tissue sampling acquisition and immunohistochemical analysis, leading to more accurate results, especially in lesions > 20 mm[8,9]. For lesions < 20 mm with malignancy risk, further analysis with a contrast-enhanced technique can stratify risk to help determine the need for and safety of biopsy[10,11].

Artificial intelligence (AI) has emerged as a powerful and exciting technology impacting many aspects of health care and promoting changes in daily clinical practice, especially for early, accurate, and real-time diagnosis[12]. Since the 1960s, AI systems have been applied in radiology for the recognition and interpretation of images and subsequently expanded to other areas including ophthalmology, cardiology, and neurology[13,14]. Between 2017 and 2018, the Food and Drug Administration approved more than 20 AI tools for medical use, including the endoscopy field[14]. In 2020, AI helped in the rational management of the coronavirus disease 2019 pandemic in various scenarios and countries, from predicting diagnostic imaging, manufacturing vaccines, and preventing viral spread[15].

The use of AI systems able to recognize specific patterns in EUS began in early 2000 and was initially applied to the evaluation of pancreatic disorders, especially differential diagnoses between chronic pancreatitis and pancreatic neoplasms[16]. Subsequently, the excellent results contributed to the development of studies that explored diagnosis and malignancy prediction for gastrointestinal SELs, particularly for GIST[17,18]. Although EUS AI-assisted has promising results, the real benefits, ethical implications, and clinical relevance need scientific evidence that supports the use in diverse clinical settings[19].

Considering the deficiency of research and the need for quality evidence to support the application of AI assistance in the subepithelial tumors EUS evaluation, this systematic review aims to perform an analysis of endoscopic ultrasound with AI assistance for GIST diagnosis. The main outcome was the diagnostic accuracy of AI-assisted EUS for GIST. Furthermore, we evaluated the AI capability that distinguishes between GIST and GIL and the experienced endoscopists' performance.

MATERIALS AND METHODS

Protocol and registration

This study was structured according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement [20] and recent recommendations for diagnostic test accuracy reviews[21]. This study was registered in the International Prospective Register of Systematic Reviews under the file number CRD42023418987.

Literature search

A comprehensive literature search was performed in the following databases up to December 2022: EMBASE, Cochrane Library, and MEDLINE. Two reviewers screened titles and abstracts of all the identified articles that evaluated the performance of AI for the diagnosis of SELs using EUS. Divergent opinions were resolved by a third reviewer. The MESH Terms for searches used were: ("endoscopic" OR "endoscopy") AND ("ultrasound" OR "endosonography" OR "echoendosonography") AND ("artificial intelligence" OR "neural network" OR "computer neural network" OR "deep learning") AND ("GIST" OR "subepithelial tumor" OR "subepithelial lesion" OR "stromal tumor" OR "gastrointestinal subepithelial tumor") present in titles, abstracts or full-text articles.

Inclusion and exclusion criteria

The studies included were performed in adults patients (> 18 years) with SELs assessed by EUS AI-assisted, with histopathological diagnosis established, and presented true-negative, true-positive, false-negative, and false-positive values which were extracted to construct a 2 × 2 table. Case reports, systematic reviews, reviews, editorials, conference abstracts, articles with algorithms different from convolutional neural network (CNN), and articles with incomplete data were excluded.

Data extraction

Using a standardized form, the relevant data from eligible studies were extracted and organized using the following main data: First author, year of publication, study type, geographical setting, number of patients, gender, number of GIST tumors, number of GIL tumors, number of other SELs, number images, tumor location, AI model, external validation, endoscopists comparison, and histopathologic analysis. All relevant texts, tables and figures were reviewed for data extraction.

Risk of bias assessment and quality of evidence

The risk of bias and quality assessment were assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool[22]. The quality of the studies was evaluated by two authors independently, and disagreement was resolved by consensus in consultation with the third author.

Measured outcomes

The main outcomes evaluated were the pooled accuracy, sensitivity, positive likelihood, negative likelihood, and specificity of AI-assisted EUS for the diagnosis of GIST based on analysis of images obtained by EUS of gastrointestinal SELs. The positive post-test probability and negative post-test probability were calculated based on likelihood ratios and GIST mean prevalence values from each article. The accuracy was defined as the area under the summary receiver operating characteristic (SROC) curve. Secondary outcomes were performed to evaluate the diagnostic performance of AI for GIST *vs* GIL, the diagnostic performance of experienced endoscopists for GIST, and GIST *vs* GIL. Experienced endoscopists were those who performed more than 500 EUS examinations or had at least 5 years of experience evaluating gastrointestinal SELs.

Statistical analysis

The pooled data of sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio, were meta-analyzed with a 95% confidence interval (CI) using the random effect model for the accuracy of EUS AI-assisted and experienced endoscopists. A SROC curve was drawn and the area under the curve (AUC) was calculated to estimate the accuracy.

Forest plots were made to show the point estimates in each study in relation to the summary pooled estimate. The width of the point estimates in the forest plots indicated the assigned weight for that study. For 0 values, 0.5 was added, as described by Cox and Snell[23]. The heterogeneity of likelihood ratios and diagnostic odds ratios were tested using Cochran's *Q* test based on inverse variance weights. The heterogeneity of the sensitivities and specificities was tested using the likelihood ratio test. Heterogeneity among studies was also tested using SROC curves. Heterogeneity was assessed and data were analyzed using Meta-DiSc (Clinical Biostatistics HRC, Madrid, Spain)[24]. The Bayes model was used to calculate the post-test probability and elaborate Fagan's Nomogram[25] using estimated mean prevalence data from each article for GIST.

RESULTS

Search results and characteristics of the included studies

A total of 163 studies were extracted after the search strategy which was shown in Figure 1. After the exclusion of 150 titles, based on the selection criteria, 13 studies were eligible for full-text examination. Of those, 4 were removed for being review articles and one[26] for not using a CNN model. Thus, eight relevant articles were selected for the present meta-analysis (Table 1) with a total of 2355 patients and 44154 images[27-34]. All articles were retrospective studies. The characteristics of the included studies are summarized in Table 1. A total of 1436 patients were diagnosed with GIST, 725 were GIL, and 194 were non-GIST/non-GIL with a GIST prevalence of 68% and leiomyoma being 30% in the present study. The Asian continent has the largest number of publications, a total of 7 articles (4 Japanese, 2 Korean, and 1 Chinese), and Europe has one (Turkey). As for the AI model used, all 8 studies were developed with a CNN algorithm.

Risk of bias and quality

The quality of the included studies was evaluated according to the QUADAS-2 tool. The risk of bias of the 8 studies is shown in Figure 2, where 7 were categorized as high risk or uncertain risk for one or more fundamental elements due to their retrospective designs (Figure 2).

Diagnostic accuracy of AI-assisted EUS for GIST

The diagnostic accuracy of GIST for AI-assisted EUS presented summary sensitivity values of 92% (95%CI: 0.89-0.95; $P < 0.01$), specificity of 80% (95%CI: 0.75-0.75; $P < 0.01$) (Figure 3), with substantial heterogeneity for both ($I^2 = 75.2\%$ and $I^2 = 71\%$, respectively). A positive likelihood ratio of 4.26 (95%CI: 2.7-6.7; $P = 0.01$), negative likelihood ratio of 0.09 (95%CI: 0.04-0.18; $P < 0.01$), and diagnostic odds ratio of 71.74 (95%CI: 22.43-229.46; $P < 0.01$) was achieved. Figure 4 shows the SROC curve, with an AUC of 0.949 ($P = 0.03$) indicating high diagnostic accuracy (Table 2). The positive post-test probability was 90% (95%CI: 0.88-0.92), and the negative post-test probability was 16% (95%CI: 0.11-0.22), as shown in Fagan's nomogram (Figure 5A).

Subgroup analysis

Diagnostic accuracy of AI-assisted EUS for GIST vs GIL: For differentiation between GIST and GIL, the AI-assisted EUS presented a combined sensitivity of 93% (95%CI: 0.88-0.97; $P = 0.08$) and combined specificity of 90% (95%CI: 0.88-0.95; $P = 0.02$), positive likelihood ratio of 6.48 (95%CI: 2.14-19.6; $P = 0.01$), negative likelihood ratio of 0.06 (95%CI: 0.02-0.21; $P = 0.05$) and diagnostic odds ratio of 128.18 (95%CI: 18.6-883.25; $P = 0.03$). The heterogeneity was $I^2 = 55\%$ for sensibility and $I^2 = 68.6\%$ for specificity. The area under SROC curve expressed high diagnostic accuracy, with values of 0.966 (AUC).

Diagnostic performance of experts for GIST: Seven studies included in this meta-analysis evaluated the diagnostic performance of experienced echo-endoscopists. The combined general values of sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, and the area under the summary ROC curve were, respectively, 72% (95%CI: 0.67-0.76; $P < 0.01$), 70% (95%CI: 0.64-0.76; $P < 0.01$), 2.51 (95%CI: 1.75-3.61; $P = 0.08$), 0.42 (95%CI: 0.33-0.52; $P = 0.16$), 6.88 (95%CI: 3.95-11.99; $P = 0.14$) and 0.777 (AUC). The heterogeneity was $I^2 = 77.3\%$ for sensibility and $I^2 = 73.8\%$ for specificity.

Diagnostic performance of experienced endoscopists' for GIST vs GIL: Considering only the differentiation of GIST and GIL, the combined sensitivity was 73% (95%CI: 0.65-0.80; $P < 0.01$), specificity 75% (95%CI: 0.65-0.84; $P = 0.91$), positive likelihood ratio 2.61 (95%CI: 1.75-3.88; $P = 0.70$), negative likelihood ratio 0.37 (95%CI: 0.22-0.64; $P = 0.02$), diagnostic odds ratio of 7.21 (95%CI: 2.95-17.59; $P = 0.16$) and area under the SROC curve of 0.819 (AUC). The heterogeneity was $I^2 = 77.3\%$ for sensibility and $I^2 = 0.0\%$ for specificity. The post- and pre-test probability were, respectively, 84% (95%CI: 0.80-0.86) and 46% (95%CI: 0.42-0.50), as shown in Figure 5B.

DISCUSSION

AI has the world's attention for the impacts generated after its implementation in the most diverse fields, especially in diagnostic medicine. The utilization of AI technology in the field of medical imaging enhances the diagnostic process, leading to improved accuracy and the early detection of diseases, thus ensuring enhanced disease management and clinical outcomes[31]. In the present systematic review and meta-analysis, we analyze the application of appropriately

Table 1 Summary of all studies investigating the development of machine learning algorithms for the endoscopic ultrasound diagnosis of subepithelial lesions

Ref.	Geographical setting	Study type	Patients	Sex (male/female)	GIST	GIL	Other SELs	Images	Tumor location	AI model	External validation	Endoscopists comparison	Histopathology
Minoda <i>et al</i> [29], 2022 ¹	Eastern	Retrospective	52	33/19	36	14	2	2718	Esophagus, duodenum, and colon	CNN	Yes	Yes	Yes
Hirai <i>et al</i> [33], 2022	Eastern	Retrospective	664	231/188	435	97	100	16110	Esophagus, stomach, and duodenum	CNN	Yes	Yes	Yes
Tanaka <i>et al</i> [34], 2022	Eastern	Retrospective	53	28/25	42	11	-	10600	Stomach	CNN	No	Yes	Yes
Yang <i>et al</i> [32], 2022	Eastern	Retrospective	752	337/415	348	404	-	10439	Esophagus, stomach, duodenum, colon, and rectum	CNN	Yes	Yes	Yes
Oh <i>et al</i> [30], 2021	Eastern	Retrospective	168	NI	125	43	-	546	Stomach	CNN	Yes	Yes	Yes
Seven <i>et al</i> [31], 2022	Eastern	Retrospective	145	72/73	109	36	-	1362	Esophagus, stomach, and duodenum	CNN	Yes	Yes	Yes
Kim <i>et al</i> [27], 2020	Eastern	Retrospective	248	111/137	157	55	35	1117	Stomach	CNN	Yes	Yes	Yes
Minoda <i>et al</i> [28], 2020	Eastern	Retrospective	273	138/135	184	65	24	3980	Stomach	CNN	Yes	Yes	Yes

¹The author utilized the same software developed in the previous study (2020), however evaluating non-gastric subepithelial lesions.

NI: Non-information; GIST: Gastrointestinal stromal tumor; AI: Artificial intelligence; CNN: Convolutional neural network; GIL: Gastrointestinal leiomyoma.

trained software on EUS AI-assisted diagnosis of fourth-layer SELs, mainly for GIST, representing the largest pooled data including eight studies with more than 2300 patients and 44154 images. Through the evaluation of AI performance in this review, we achieved a combined sensitivity of 92%, a combined specificity of 80%, a positive post-test probability of 90%, and a negative post-test probability of 16% when distinguishing between GIST and non-GIST SELs based on EUS images.

The imaging modality frequently used to evaluate SELs is the EUS because of its ability to characterize the size, echogenicity, originating layer, shape, vascularity, and location[1,3,35]. Regarding conventional EUS findings for GIST, just stronger echogenicity in comparison with the surrounding muscle echo is an associated independent diagnostic factor[31]. A previous study reported that echogenicity, the presence of hyperechogenic spots and anechoic spaces, tumor shape, and marginal regularity in the EUS were not helpful in differentiating GIST from non-GIST tumors, being homogeneity was the only predictive factor[35]. Thus, the differentiation between GIST from other SELs without histological evaluation is difficult using EUS images only because the interpretation of the features is subjective and dependent on the experience of the endoscopist with a heterogeneous inter-observer agreement[28,31]. Although the gold standard diagnostic is histological evaluation, someone's SELs can only be monitored with follow-up exams in the absence of risk stigmata and resection indications[9,10].

Table 2 Artificial intelligence and experienced endoscopists' diagnostic performance for gastrointestinal stromal tumor and differentiation of leiomyoma

	Prevalence (mean)	Sensitivity	Specificity	LR+	LR-	AUC	PTP+	PTP-
AI								
GIST	68%	92%	80%	4.26	0.09	0.949	90%	16%
GIST <i>vs</i> GIL	70% ¹	93%	90%	6.48	0.06	0.966	94%	12%
Endoscopists								
GIST	67%	72%	70%	2.51	0.42	0.777	84%	46%
GIST <i>vs</i> GIL	67% ¹	73%	75%	2.61	0.37	0.819	84%	43%

¹Gastrointestinal stromal tumor prevalence.

GIST: Gastrointestinal stromal tumor; GIL: Gastrointestinal leiomyoma; LR+: Likelihood ratio positive; LR-: Likelihood ratio negative; PTP+: Post-test probability positive; PTP-: Post-test probability negative; AUC: Area under the curve; AI: Artificial intelligence.

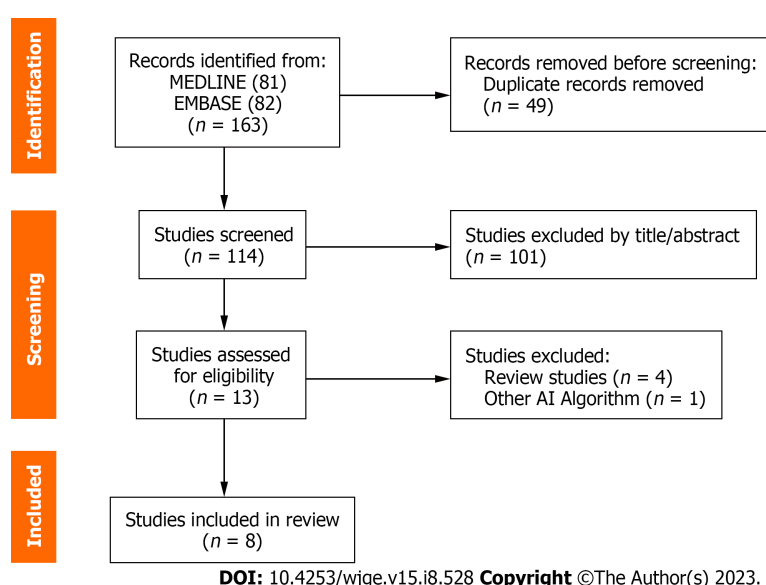


Figure 1 Flow diagram showing the study selection process for meta-analysis. AI: Artificial intelligence.

The implementation of AI technology to improve the EUS diagnostic performance compensates limitations and disagreements discussed previously. Our results for overall GIST diagnosis by AI were superior to the EUS doctors' performance earlier reported[36], showing the ability to improve differential diagnosis with efficiency, quick evaluation, and reduce unnecessary procedures and surgical interventions. AI can evaluate specific patterns in the pixel-level characteristics of a tumor, making it more accurate than the naked eye in its analysis, and was observed that the size of the lesion increases, diagnostic accuracy increases in parallel, being more expressive from > 20 mm[28,34]. Considering the risks of invasive procedures, the misdiagnosis rate of GIST, the requirement of pathologic specimens for determining malignancy potential, and eligibility for neo-adjuvant therapy, the application of AI has significant improvements in the safety clinical management of SELs[28]. Recently, software developed and trained with gastric GIST images for EUS AI-assisted diagnosis has been used in other gastrointestinal sites with excellent results[29]. Furthermore, EUS assisted by AI developed to diagnose and prediction of the malignancy risk of GIST showed excellent performance[18], assisting in the decision for resection with or without neoadjuvant therapy. Using AI as an auxiliary tool in the diagnosis in endoscopy aims primarily to increase diagnostic accuracy and reduce the number of false negatives and positives.

In the evaluation of SELs originating from the muscular layer, the differentiation between GIST and leiomyoma is one of the most challenging, since they have very similar sonographic characteristics: Both are hypoechoic, usually homogeneous, have well-defined limits, and may originate from the second or fourth-layer. Diagnosis is a challenge, even in contrast to enhanced exams[34] yet it is extremely important since it determines the need for surgical resection. When evaluating the differentiation between GIST *vs* GIL by the AI-assisted EUS, combined sensitivity and specificity of 93% and 90% respectively were obtained with an AUC of 0.966, indicating excellent performance. The AI software increases the possibility of correct diagnosis avoiding unnecessary invasive procedures and excess risk with biopsies, without size limitation even for lesions < 20 mm[32]. The AI-assisted EUS performed even better for lesions larger than 20 mm, which is highly relevant considering, for example, that GIST > 2 cm has an indication for resection while asymptomatic

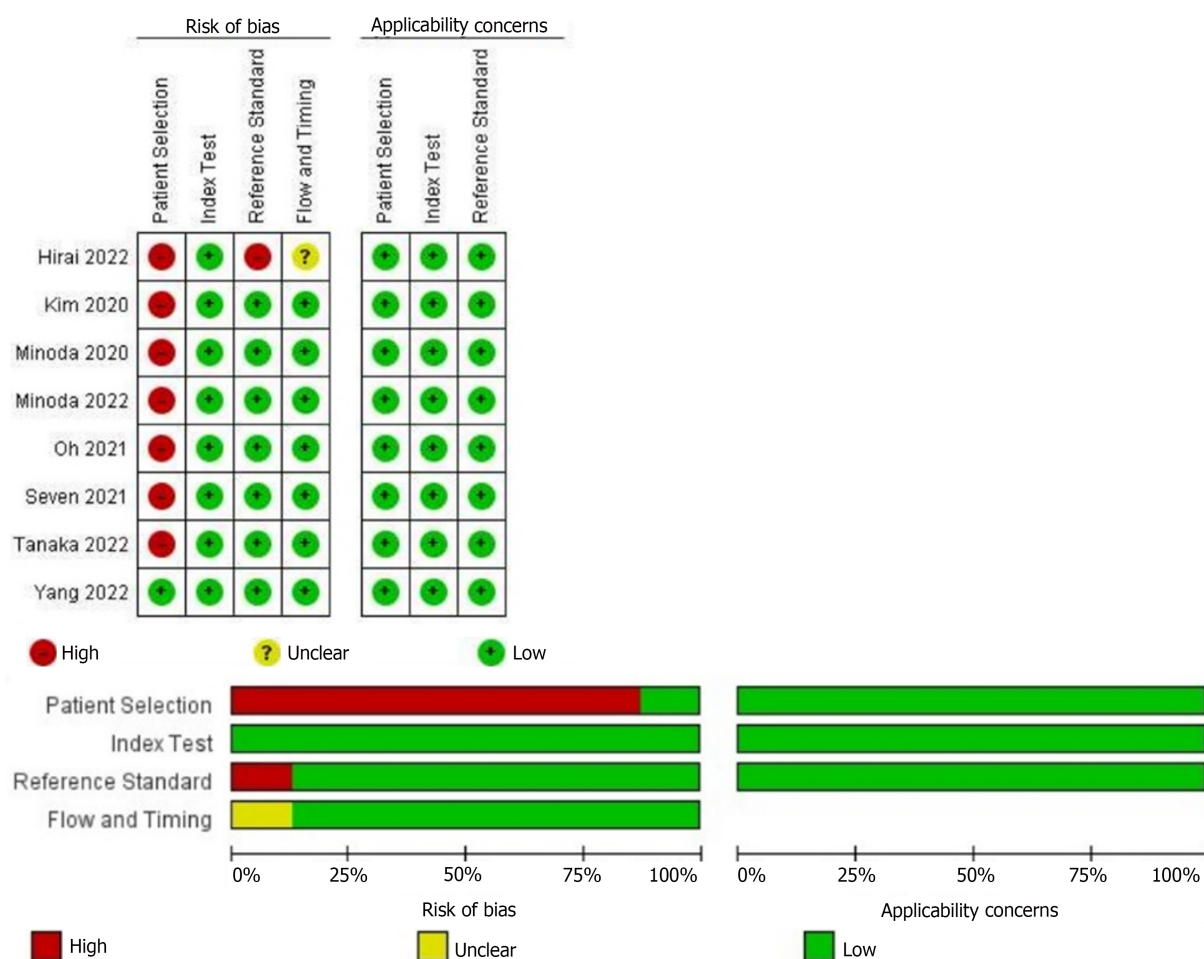


Figure 2 Risk of bias and applicability concerns of 8 included records using the Quality Assessment of Diagnostic Accuracy Studies tool.

leiomyomas should not be removed.

The performance of AI was superior to EUS experts, even in cases where the expert was informed of the source layer and the location of the lesion[31], a fact that favors the expert because with some information certain diagnoses become obvious. For example, a soft, hyperechoic lesion of the submucosa, in the antrum, is highly suspicious for lipoma. With a sensitivity of 72% (*vs* 93% AI-assisted), experts can expect to have a miss rate of approximately 3 in every 10 cases of GIST. The diagnostic accuracy of the experts for GIST, although considered good, was much lower than that of the AI system (AUC 0.819 *vs* 0.966). A previous study reported an increase in accuracy, specificity, and positive predictive value after joint diagnosis of endoscopists with AI assistance to distinguish GIST from leiomyoma, demonstrating that AI has the potential to help enhance correct diagnosis even for experienced endoscopists. In addition, the rapid development of AI systems capable of performing fast and more specific analyses without increasing operating costs or equipment updates[32] makes it possible and attractively apply in diagnostic centers of lower volume and invariably assists inexperienced endoscopists' diagnoses.

Although these results are exciting, they should be evaluated with caution due to the dynamic nature of diagnostic examinations. In these studies, experts evaluated images of SEL without being able to perform their usual maneuvers. Moreover, most studies did not provide essential information commonly known in clinical practice, such as the patient's medical history, color at white light endoscopy, or lesion consistency[37]. Additionally, the expected performance of AI can be influenced by factors such as disease prevalence and severity, the expertise and training of endoscopists, and the interaction between AI and endoscopists[19]. It is crucial to emphasize that the use of machine learning models in the medical field should not be seen as a direct competition to endoscopists. Instead, they should be regarded as auxiliary diagnostic tools and even training aids for less experienced endoscopists.

Despite the number of studies, patients, and reviewed images, this meta-analysis has certain limitations. Firstly, although we included several studies with a large number of patients, all of them were retrospective, thereby reducing the quality of evidence. However, given the scarcity of data, this meta-analysis is crucial in improving our understanding of the current level of evidence for AI-assisted EUS. Secondly, this meta-analysis exhibited significant heterogeneity, likely due to the variability in the populations included in the studies. For instance, two studies evaluated the AI's performance by categorizing the SELs based on a 20 mm cut-off for size[28,34], which limits the performance evaluation to SELs < 20 mm. Thirdly, the varying quality of EUS devices and images used in the trials limits their applicability in

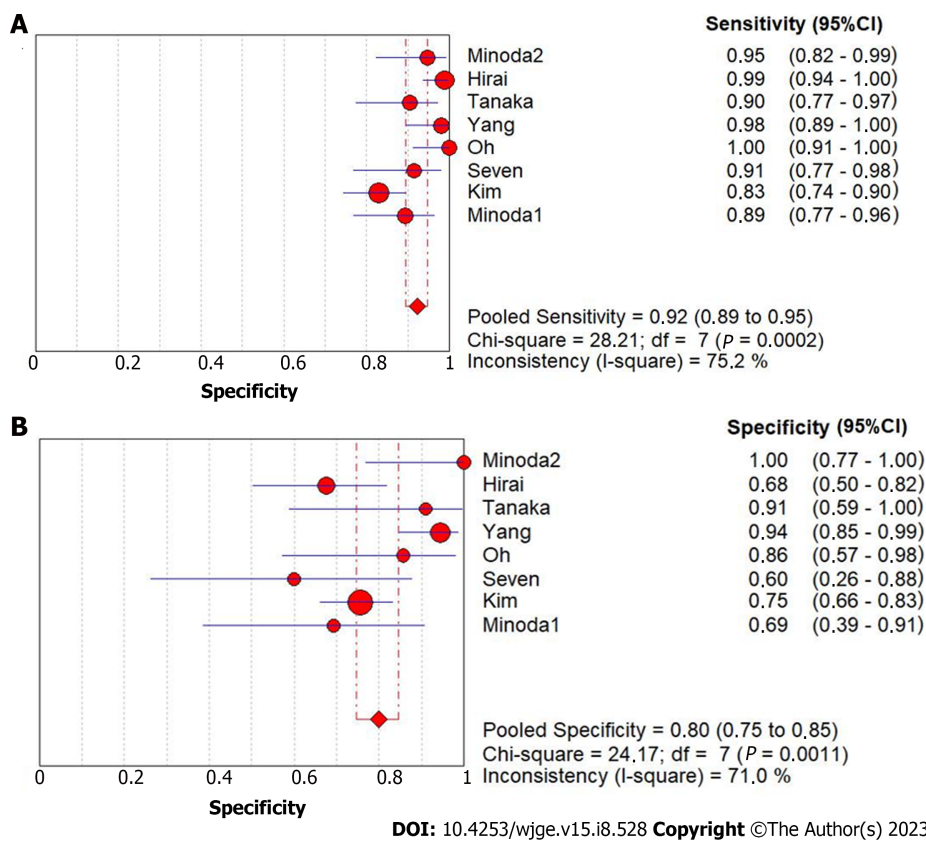


Figure 3 Sensibility and specificity of endoscopic ultrasound with Artificial Intelligence for overall diagnosis of gastrointestinal stromal tumor. A: Sensibility; B: Specificity. CI: Confidence interval.

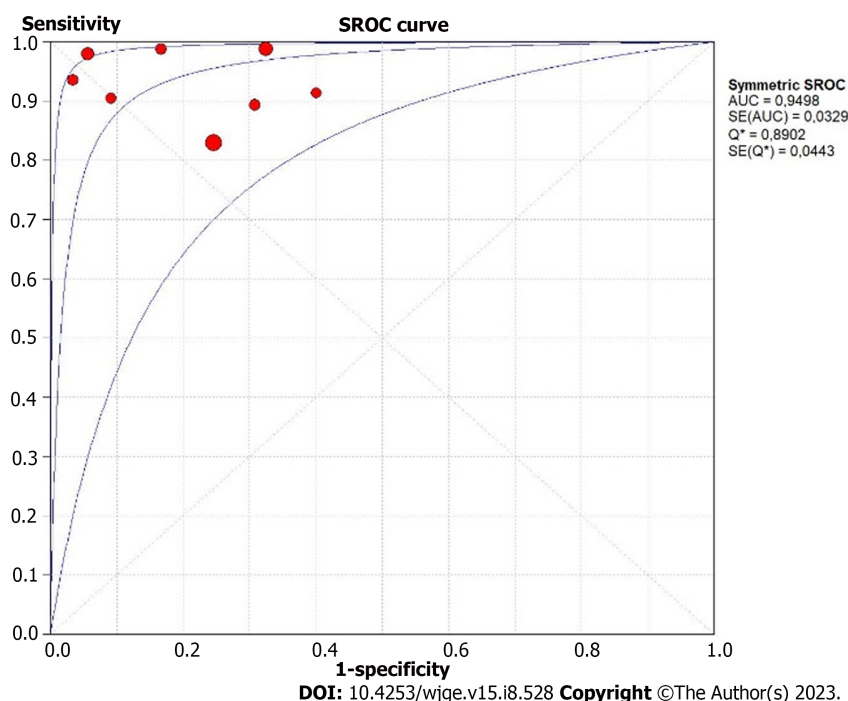
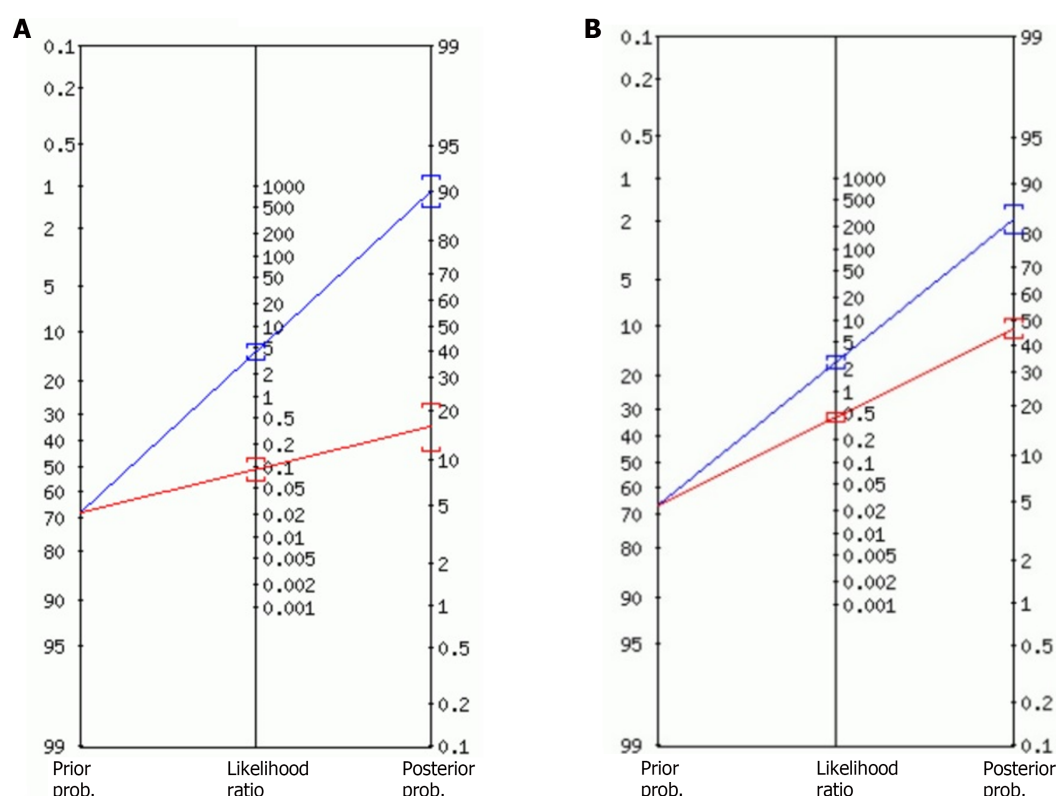


Figure 4 Summary receiver operating characteristic curve of endoscopic ultrasound with artificial intelligence for gastrointestinal stromal tumor diagnostic. Each circle represent an individual study. SROC: Summary receiver operating characteristic; AUC: Area under curve.

real-world scenarios. Many of these studies only employed internal validation datasets for training the algorithms, which may potentially result in an overestimation of the AI models' performance. This situation is indicative of overfitting[37], where a machine learning model becomes overly specialized to the training data and performs poorly on new.



DOI: 10.4253/wjge.v15.i8.528 Copyright ©The Author(s) 2023.

Figure 5 Fagan plot depicting the impact of a positive or negative result on pretest probabilities. A: Fagan plot depicting the impact of a positive or negative result of artificial intelligence; B: Fagan plot depicting the impact of a positive or negative result of experienced endoscopists' (i.e., the pooled prevalence of subepithelial lesions).

In summary, AI-supported EUS demonstrates notable diagnostic precision in retrospective investigations related to the detection of GIST. Furthermore, AI has shown superior accuracy compared to experienced endoscopists, indicating its potential as a significant diagnostic adjunct in this field. The advancement of AI algorithms and EUS devices, along with the increased accessibility of EUS and the availability of high-quality EUS images, creates a favorable environment for robust studies aiming to achieve enhanced diagnostic performance and develop valuable, clinically applicable tools. Consequently, AI technology has the potential to profoundly influence all aspects of healthcare, as indicated by current research findings.

CONCLUSION

In conclusion, our systematic review and meta-analysis demonstrated the high diagnostic accuracy of EUS AI-assisted for the differentiation of SELs, especially GIST from other fourth-layer subepithelial tumors. AI revealed the potential to become help enhance endoscopists' diagnostic performance in the EUS evaluation of SELs and avoid unnecessary invasive procedures.

ARTICLE HIGHLIGHTS

Research background

Endoscopic ultrasonography (EUS) with artificial intelligence (AI) has shown high diagnostic accuracy for subepithelial lesions (SELs), particularly gastrointestinal stromal tumors (GISTs). The performance of AI systems has demonstrated superiority over experienced endoscopists and the ability to improve diagnostic power through collaborative diagnosis.

Research motivation

This paper aims to investigate the diagnostic capabilities of AI-assisted EUS for SELs by analyzing images and comparing them with the expertise of experienced endoscopists.

Research objectives

The research aims to assess the accuracy of AI-assisted EUS in diagnosing SELs, particularly those originating from the fourth layer. Additionally, the study analyzes the diagnostic performance of experienced endoscopists and compares it with AI systems.

Research methods

Retrospective studies were selected of AI-assisted EUS for the diagnosis of SELs, using histopathology as the standard method. The included studies utilized EUS with AI for SELs diagnosis through image analysis. The risk of bias and quality of evidence were assessed, and the analysis was performed using Meta-Disc software.

Research results

This meta-analysis included eight retrospective studies with a total of 2355 patients and 44154 images. The AI-assisted EUS for GIST diagnosis showed a sensitivity of 92% [95% confidence interval (CI): 0.89-0.95; $P < 0.01$], specificity of 80% (95%CI: 0.75-0.85; $P < 0.01$), and an AUC of 0.949. For the diagnosis of GIST *vs* gastrointestinal leiomyoma (GIL) by AI-assisted EUS, specificity was 90% (95%CI: 0.88-0.95; $P = 0.02$) and AUC 0.966. The experienced endoscopists achieved a sensitivity of 72% (95%CI: 0.67-0.76; $P < 0.01$), specificity of 70% (95%CI: 0.64-0.76; $P < 0.01$), and an AUC of 0.777 for GIST. Evaluating GIST *vs* GIL, the experts achieved a sensitivity of 73% (95%CI: 0.65-0.80; $P < 0.01$) and an AUC of 0.819.

Research conclusions

This systematic review and meta-analysis demonstrate the high diagnostic accuracy of AI-assisted EUS in differentiating SELs, particularly GIST, from other fourth-layer subepithelial tumors.

Research perspectives

This study demonstrated that by integrating machine learning techniques with EUS images, AI can aid in distinguishing benign from malignant lesions and guiding treatment decisions, with high accuracy. Additionally, through AI assistance image recognition can enhance real-time diagnosis during EUS evaluations, increasing the performance of even experienced endoscopists.

FOOTNOTES

Author contributions: Gomes RSA contributed to the acquisition of data; Gomes RSA, de Oliveira GHP, Hirsch BS, Ribeiro Jordão Sasso JG, Matsubayashi CO, Kotinda APST, Veras MO, Moura DTH, Bernardo WM, and de Moura EGH contributed to the analysis of data; Gomes RSA, de Oliveira GHP, Hirsch BS, Ribeiro Jordão Sasso JG, Matsubayashi CO, Kotinda APST, Veras MO, Moura DTH, Bernardo WM, Trasolini RP, and de Moura EGH contributed to the interpretation of data; Gomes RSA, de Moura DTH, Trasolini RP, Bernardo WM, and de Moura EGH drafted the article; Gomes RSA, de Oliveira GHP, Hirsch BS, Ribeiro Jordão Sasso JG, Matsubayashi CO, Kotinda APST, Veras MO, de Moura DTH, Trasolini RP, Bernardo WM, and de Moura EGH revised the manuscript; Trasolini RP revised the English language; and all author approved the final manuscript.

Conflict-of-interest statement: All the authors report no relevant conflicts of interest for this article.

PRISMA 2009 Checklist statement: The authors have read the PRISMA 2009 Checklist, and the manuscript was prepared and revised according to the PRISMA 2009 Checklist.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

Country/Territory of origin: Brazil

ORCID number: Rômulo Sérgio Araújo Gomes 0009-0002-3092-7766; Guilherme Henrique Peixoto de Oliveira 0000-0002-1057-2390; Diogo Turiani Hourneaux de Moura 0000-0002-7446-0355; Ana Paula Samy Tanaka Kotinda 0000-0002-2915-8974; Carolina Ogawa Matsubayashi 0000-0002-2175-0439; Bruno Salomão Hirsch 0000-0002-0777-0150; Matheus de Oliveira Veras 0000-0001-9881-9855; João Guilherme Ribeiro Jordão Sasso 0000-0002-4283-4295; Roberto Paolo Trasolini 0000-0001-8059-9807; Wanderley Marques Bernardo 0000-0002-8597-5207; Eduardo Guimarães Hourneaux de Moura 0000-0003-1215-5731.

S-Editor: Wang JJ

L-Editor: A

P-Editor: Cai YX

REFERENCES

- 1 Deprez PH, Moons LMG, O'Toole D, Gincul R, Seicean A, Pimentel-Nunes P, Fernández-Esparrach G, Polkowski M, Vieth M, Borbath I,

- Moreels TG, Nieveen van Dijkum E, Blay JY, van Hooft JE. Endoscopic management of subepithelial lesions including neuroendocrine neoplasms: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy* 2022; **54**: 412-429 [PMID: [35180797](#) DOI: [10.1055/a-1751-5742](#)]
- 2 **Park CH**, Kim EH, Jung DH, Chung H, Park JC, Shin SK, Lee YC, Kim H, Lee SK. Impact of periodic endoscopy on incidentally diagnosed gastric gastrointestinal stromal tumors: findings in surgically resected and confirmed lesions. *Ann Surg Oncol* 2015; **22**: 2933-2939 [PMID: [25808096](#) DOI: [10.1245/s10434-015-4517-0](#)]
- 3 **Standards of Practice Committee**, Faulx AL, Kothari S, Acosta RD, Agrawal D, Bruining DH, Chandrasekhara V, Eloubeidi MA, Fanelli RD, Gurudu SR, Khashab MA, Lightdale JR, Muthusamy VR, Shaikat A, Qumseya BJ, Wang A, Wani SB, Yang J, DeWitt JM. The role of endoscopy in subepithelial lesions of the GI tract. *Gastrointest Endosc* 2017; **85**: 1117-1132 [PMID: [28385194](#) DOI: [10.1016/j.gie.2017.02.022](#)]
- 4 **Akahoshi K**, Oya M, Koga T, Shiratsuchi Y. Current clinical management of gastrointestinal stromal tumor. *World J Gastroenterol* 2018; **24**: 2806-2817 [PMID: [30018476](#) DOI: [10.3748/wjg.v24.i26.2806](#)]
- 5 **Nishida T**, Blay JY, Hirota S, Kitagawa Y, Kang YK. The standard diagnosis, treatment, and follow-up of gastrointestinal stromal tumors based on guidelines. *Gastric Cancer* 2016; **19**: 3-14 [PMID: [26276366](#) DOI: [10.1007/s10120-015-0526-8](#)]
- 6 **Park EY**, Kim GH. Diagnosis of Gastric Subepithelial Tumors Using Endoscopic Ultrasonography or Abdominopelvic Computed Tomography: Which is Better? *Clin Endosc* 2019; **52**: 519-520 [PMID: [31722518](#) DOI: [10.5946/ce.2019.188](#)]
- 7 **Alkhatib AA**, Faigel DO. Endoscopic ultrasonography-guided diagnosis of subepithelial tumors. *Gastrointest Endosc Clin N Am* 2012; **22**: 187-205, vii [PMID: [22632943](#) DOI: [10.1016/j.giec.2012.04.006](#)]
- 8 **de Moura DTH**, McCarty TR, Jirapinyo P, Ribeiro IB, Flumignan VK, Najdawai F, Ryou M, Lee LS, Thompson CC. EUS-guided fine-needle biopsy sampling versus FNA in the diagnosis of subepithelial lesions: a large multicenter study. *Gastrointest Endosc* 2020; **92**: 108-119.e3 [PMID: [32105712](#) DOI: [10.1016/j.gie.2020.02.021](#)]
- 9 **Cheng S**, Matuguma SE, de Oliveira GHP, Silva GLR, Cheng H, Sánchez-Luna SA, Minata MK. Colonoscopic Ultrasound-Guided Fine-Needle Aspiration Using a Curvilinear Array Transducer: A Single-Center Retrospective Cohort Study. *Dis Colon Rectum* 2022; **65**: e80-e84 [PMID: [34840298](#) DOI: [10.1097/DCR.0000000000002333](#)]
- 10 **Kamata K**, Takenaka M, Kitano M, Omoto S, Miyata T, Minaga K, Yamao K, Imai H, Sakurai T, Watanabe T, Nishida N, Chikugo T, Chiba Y, Imamoto H, Yasuda T, Lisotti A, Fusaroli P, Kudo M. Contrast-enhanced harmonic endoscopic ultrasonography for differential diagnosis of submucosal tumors of the upper gastrointestinal tract. *J Gastroenterol Hepatol* 2017; **32**: 1686-1692 [PMID: [28220958](#) DOI: [10.1111/jgh.13766](#)]
- 11 **Pesenti C**, Bories E, Caillol F, Ratone JP, Godat S, Monges G, Poizat F, Raoul JL, Ries P, Giovannini M. Characterization of subepithelial lesions of the stomach and esophagus by contrast-enhanced EUS: A retrospective study. *Endosc Ultrasound* 2019; **8**: 43-49 [PMID: [30264741](#) DOI: [10.4103/eus.eus_89_17](#)]
- 12 **Oka A**, Ishimura N, Ishihara S. A New Dawn for the Use of Artificial Intelligence in Gastroenterology, Hepatology and Pancreatology. *Diagnostics (Basel)* 2021; **11** [PMID: [34574060](#) DOI: [10.3390/diagnostics11091719](#)]
- 13 **Topol EJ**. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019; **25**: 44-56 [PMID: [30617339](#) DOI: [10.1038/s41591-018-0300-7](#)]
- 14 **Gao J**, Jiang Q, Zhou B, Chen D. Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: An overview. *Math Biosci Eng* 2019; **16**: 6536-6561 [PMID: [31698575](#) DOI: [10.3934/mbe.2019326](#)]
- 15 **Mijwil MM**, Abttan RA, Alkhazraji A. Artificial intelligence for COVID-19: A Short Article. *AJPNMS* 2022; **10** [DOI: [10.24203/ajpnms.v10i1.6961](#)]
- 16 **Norton ID**, Zheng Y, Wiersema MS, Greenleaf J, Clain JE, Dimagno EP. Neural network analysis of EUS images to differentiate between pancreatic malignancy and pancreatitis. *Gastrointest Endosc* 2001; **54**: 625-629 [PMID: [11677484](#) DOI: [10.1067/mge.2001.118644](#)]
- 17 **Chen T**, Xu L, Dong X, Li Y, Yu J, Xiong W, Li G. The roles of CT and EUS in the preoperative evaluation of gastric gastrointestinal stromal tumors larger than 2 cm. *Eur Radiol* 2019; **29**: 2481-2489 [PMID: [30617491](#) DOI: [10.1007/s00330-018-5945-6](#)]
- 18 **Seven G**, Silahatoglu G, Kochan K, Ince AT, Arici DS, Senturk H. Use of Artificial Intelligence in the Prediction of Malignant Potential of Gastric Gastrointestinal Stromal Tumors. *Dig Dis Sci* 2022; **67**: 273-281 [PMID: [33547537](#) DOI: [10.1007/s10620-021-06830-9](#)]
- 19 **Messmann H**, Bisschops R, Antonelli G, Libânio D, Sinouque P, Abdelrahim M, Ahmad OF, Areia M, Bergman JJGHM, Bhandari P, Boskoski I, Dekker E, Domagk D, Ebigbo A, Eelbode T, Eliakim R, Häfner M, Haidry RJ, Jover R, Kaminski MF, Kuvaev R, Mori Y, Palazzo M, Repici A, Rondonotti E, Rutter MD, Saito Y, Sharma P, Spada C, Spadaccini M, Veitch A, Gralnek IM, Hassan C, Dinis-Ribeiro M. Expected value of artificial intelligence in gastrointestinal endoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. *Endoscopy* 2022; **54**: 1211-1231 [PMID: [36270318](#) DOI: [10.1055/a-1950-5694](#)]
- 20 **Liberati A**, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; **339**: b2700 [PMID: [19622552](#) DOI: [10.1136/bmj.b2700](#)]
- 21 **Leefflang MM**, Deeks JJ, Takwoingi Y, Macaskill P. Cochrane diagnostic test accuracy reviews. *Syst Rev* 2013; **2**: 82 [PMID: [24099098](#) DOI: [10.1186/2046-4053-2-82](#)]
- 22 **Yang B**, Mallett S, Takwoingi Y, Davenport CF, Hyde CJ, Whiting PF, Deeks JJ, Leeflang MMG, QUADAS-C Group†, Bossuyt PMM, Brazzelli MG, Dinnies J, Gurusamy KS, Jones HE, Lange S, Langendam MW, Macaskill P, McInnes MDF, Reitsma JB, Rutjes AWS, Sinclair A, de Vet HCW, Virgili G, Wade R, Westwood ME. QUADAS-C: A Tool for Assessing Risk of Bias in Comparative Diagnostic Accuracy Studies. *Ann Intern Med* 2021; **174**: 1592-1599 [PMID: [34698503](#) DOI: [10.7326/M21-2234](#)]
- 23 **Cox DR**, Snell EJ. The Analysis of Binary Data. In: The Concise Encyclopedia of Statistics. New York: Springer New York, 1989: 4-5
- 24 **Zamora J**, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2006; **6**: 31 [PMID: [16836745](#) DOI: [10.1186/1471-2288-6-31](#)]
- 25 **Fagan TJ**. Letter: Nomogram for Bayes theorem. *N Engl J Med* 1975; **293**: 257 [PMID: [1143310](#) DOI: [10.1056/NEJM197507312930513](#)]
- 26 **Nguyen VX**, Nguyen CC, Li B, Das A. Digital image analysis is a useful adjunct to endoscopic ultrasonographic diagnosis of subepithelial lesions of the gastrointestinal tract. *J Ultrasound Med* 2010; **29**: 1345-1351 [PMID: [20733191](#) DOI: [10.7863/jum.2010.29.9.1345](#)]
- 27 **Kim YH**, Kim GH, Kim KB, Lee MW, Lee BE, Baek DH, Kim DH, Park JC. Application of A Convolutional Neural Network in The Diagnosis of Gastric Mesenchymal Tumors on Endoscopic Ultrasonography Images. *J Clin Med* 2020; **9** [PMID: [33003602](#) DOI: [10.3390/jcm9103162](#)]
- 28 **Minoda Y**, Ihara E, Komori K, Ogino H, Otsuka Y, Chinen T, Tsuda Y, Ando K, Yamamoto H, Ogawa Y. Efficacy of endoscopic ultrasound

- with artificial intelligence for the diagnosis of gastrointestinal stromal tumors. *J Gastroenterol* 2020; **55**: 1119-1126 [PMID: [32918102](#) DOI: [10.1007/s00535-020-01725-4](#)]
- 29 **Minoda Y**, Ihara E, Fujimori N, Nagatomo S, Esaki M, Hata Y, Bai X, Tanaka Y, Ogino H, Chinen T, Hu Q, Oki E, Yamamoto H, Ogawa Y. Efficacy of ultrasound endoscopy with artificial intelligence for the differential diagnosis of non-gastric gastrointestinal stromal tumors. *Sci Rep* 2022; **12**: 16640 [PMID: [36198726](#) DOI: [10.1038/s41598-022-20863-8](#)]
- 30 **Oh CK**, Kim T, Cho YK, Cheung DY, Lee BI, Cho YS, Kim JI, Choi MG, Lee HH, Lee S. Convolutional neural network-based object detection model to identify gastrointestinal stromal tumors in endoscopic ultrasound images. *J Gastroenterol Hepatol* 2021; **36**: 3387-3394 [PMID: [34369001](#) DOI: [10.1111/jgh.15653](#)]
- 31 **Seven G**, Silahtaroglu G, Seven OO, Senturk H. Differentiating Gastrointestinal Stromal Tumors from Leiomyomas Using a Neural Network Trained on Endoscopic Ultrasonography Images. *Dig Dis* 2022; **40**: 427-435 [PMID: [34619683](#) DOI: [10.1159/000520032](#)]
- 32 **Yang X**, Wang H, Dong Q, Xu Y, Liu H, Ma X, Yan J, Li Q, Yang C, Li X. An artificial intelligence system for distinguishing between gastrointestinal stromal tumors and leiomyomas using endoscopic ultrasonography. *Endoscopy* 2022; **54**: 251-261 [PMID: [33827140](#) DOI: [10.1055/a-1476-8931](#)]
- 33 **Hirai K**, Kuwahara T, Furukawa K, Kakushima N, Furune S, Yamamoto H, Marukawa T, Asai H, Matsui K, Sasaki Y, Sakai D, Yamada K, Nishikawa T, Hayashi D, Obayashi T, Komiyama T, Ishikawa E, Sawada T, Maeda K, Yamamura T, Ishikawa T, Ohno E, Nakamura M, Kawashima H, Ishigami M, Fujishiro M. Artificial intelligence-based diagnosis of upper gastrointestinal subepithelial lesions on endoscopic ultrasonography images. *Gastric Cancer* 2022; **25**: 382-391 [PMID: [34783924](#) DOI: [10.1007/s10120-021-01261-x](#)]
- 34 **Tanaka H**, Kamata K, Ishihara R, Handa H, Otsuka Y, Yoshida A, Yoshikawa T, Ishikawa R, Okamoto A, Yamazaki T, Nakai A, Omoto S, Minaga K, Yamao K, Takenaka M, Watanabe T, Nishida N, Kudo M. Value of artificial intelligence with novel tumor tracking technology in the diagnosis of gastric submucosal tumors by contrast-enhanced harmonic endoscopic ultrasonography. *J Gastroenterol Hepatol* 2022; **37**: 841-846 [PMID: [35043456](#) DOI: [10.1111/jgh.15780](#)]
- 35 **Kim SM**, Kim EY, Cho JW, Jeon SW, Kim JH, Kim TH, Moon JS, Kim JO; Research Group for Endoscopic Ultrasound of the Korean Society of Gastrointestinal Endoscopy. Predictive Factors for Differentiating Gastrointestinal Stromal Tumors from Leiomyomas Based on Endoscopic Ultrasonography Findings in Patients with Gastric Subepithelial Tumors: A Multicenter Retrospective Study. *Clin Endosc* 2021; **54**: 872-880 [PMID: [34788936](#) DOI: [10.5946/ce.2021.251](#)]
- 36 **Seo SW**, Hong SJ, Han JP, Choi MH, Song JY, Kim HK, Lee TH, Ko BM, Cho JY, Lee JS, Lee MS. Accuracy of a scoring system for the differential diagnosis of common gastric subepithelial tumors based on endoscopic ultrasonography. *J Dig Dis* 2013; **14**: 647-653 [PMID: [23992089](#) DOI: [10.1111/1751-2980.12099](#)]
- 37 **Park SH**, Han K. Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. *Radiology* 2018; **286**: 800-809 [PMID: [29309734](#) DOI: [10.1148/radiol.2017171920](#)]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA

Telephone: +1-925-3991568

E-mail: bpgoffice@wjgnet.com

Help Desk: <https://www.f6publishing.com/helpdesk>

<https://www.wjgnet.com>

