

Original Paper

The Accuracy of Artificial Intelligence in the Endoscopic Diagnosis of Early Gastric Cancer: Pooled Analysis Study

Pei-Chin Chen^{1,2*}, MD, MSc; Yun-Ru Lu^{2,3*}, MD; Yi-No Kang^{4,5,6,7*}, MA; Chun-Chao Chang⁸, MD

¹Department of Internal Medicine, Taipei Medical University Hospital, Taipei, Taiwan

²Department of General Medicine, Taipei Medical University Hospital, Taipei, Taiwan

³Department of Anesthesiology, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan

⁴Evidence-Based Medicine Center, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan

⁵Institute of Health Behaviors and Community Sciences, College of Public Health, National Taiwan University, Taipei, Taiwan

⁶Cochrane Taiwan, Taipei Medical University, Taipei, Taiwan

⁷Department of Health Care Management, College of Health Technology, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan

⁸Division of Gastroenterology and Hepatology, Department of Internal Medicine, Taipei Medical University Hospital, Taipei, Taiwan

*these authors contributed equally

Corresponding Author:

Chun-Chao Chang, MD

Division of Gastroenterology and Hepatology

Department of Internal Medicine

Taipei Medical University Hospital

No 252, Wuxing St

Taipei, 110

Taiwan

Phone: 886 227372181

Email: chunchao@tmu.edu.tw

Abstract

Background: Artificial intelligence (AI) for gastric cancer diagnosis has been discussed in recent years. The role of AI in early gastric cancer is more important than in advanced gastric cancer since early gastric cancer is not easily identified in clinical practice. However, to our knowledge, past syntheses appear to have limited focus on the populations with early gastric cancer.

Objective: The purpose of this study is to evaluate the diagnostic accuracy of AI in the diagnosis of early gastric cancer from endoscopic images.

Methods: We conducted a systematic review from database inception to June 2020 of all studies assessing the performance of AI in the endoscopic diagnosis of early gastric cancer. Studies not concerning early gastric cancer were excluded. The outcome of interest was the diagnostic accuracy (comprising sensitivity, specificity, and accuracy) of AI systems. Study quality was assessed on the basis of the revised Quality Assessment of Diagnostic Accuracy Studies. Meta-analysis was primarily based on a bivariate mixed-effects model. A summary receiver operating curve and a hierarchical summary receiver operating curve were constructed, and the area under the curve was computed.

Results: We analyzed 12 retrospective case control studies (n=11,685) in which AI identified early gastric cancer from endoscopic images. The pooled sensitivity and specificity of AI for early gastric cancer diagnosis were 0.86 (95% CI 0.75-0.92) and 0.90 (95% CI 0.84-0.93), respectively. The area under the curve was 0.94. Sensitivity analysis of studies using support vector machines and narrow-band imaging demonstrated more consistent results.

Conclusions: For early gastric cancer, to our knowledge, this was the first synthesis study on the use of endoscopic images in AI in diagnosis. AI may support the diagnosis of early gastric cancer. However, the collocation of imaging techniques and optimal algorithms remain unclear. Competing models of AI for the diagnosis of early gastric cancer are worthy of future investigation.

Trial Registration: PROSPERO CRD42020193223; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=193223

(*J Med Internet Res* 2022;24(5):e27694) doi: [10.2196/27694](https://doi.org/10.2196/27694)

KEYWORDS

artificial intelligence; early gastric cancer; endoscopy

Introduction

Gastric cancer is the fifth most common cancer and the third leading cause of cancer deaths worldwide, contributing to 19.1 million disability-adjusted life years in 2017 [1,2]. Its primary risk factors are *Helicobacter pylori* infection and a family history of gastric cancer [3,4]. Despite advancements in endoscopic, surgical, and systemic therapies, the global 5-year survival rate of those with gastric cancer remains low (25%-30%) [5]. Gastric cancer has an excellent prognosis at early stages, with a 5-year survival rate of approximately 95%, but it has a median survival rate of less than one year at advanced stages [6,7]. Its favorable early prognosis is reflected in the lower mortality rates of gastric cancer in East Asia, which can be ascribed to the implementation of nationwide screening [8]. This reinforces the importance of early diagnosis. However, gastrointestinal endoscopy, the standard detection method for early gastric cancer, has an unsatisfactory sensitivity of 70% and is operator dependent [9]. Despite efforts to increase the detection rate, a valid screening method has yet to be developed [10,11]. The recent advancement in artificial intelligence (AI) systems, which provides highly accurate and efficient image recognition, may indicate a solution to this problem.

Although significant increases in AI exist in many fields and in health care [12-19], AI has various definitions [20]. According to the cognitive modeling approach, AI can be seen as machines that perform or exhibit actions corresponding to intelligence such as human behavior [20,21]. Machine learning, a subset of AI, involves studying how computers learn to improve task performance through experience without being programmed. This learning is achieved through various approaches. For instance, support vector machines, widely used in data classification, are machine learning algorithms that work by calculating the best separating plane for distinguishing between different objects. Deep learning, another machine learning method, simulates the multiple hierarchical layers of neural networks to make decisions based on features extracted from massive training data. Convolutional neural networks are deep learning algorithms primarily used in image recognition [22].

Since the breakthrough of deep learning in the 2010s, the use of AI in clinical practice has increased dramatically [22,23], and many studies have applied AI for screening or diagnosis [24-27]. Several studies have provided promising results for the AI-assisted endoscopic diagnosis of gastric cancer [28]. In a multicenter case control study of 84,424 participants, a deep learning-aided system demonstrated a detection rate of upper gastrointestinal cancer comparable to that of an expert endoscopist [29]. Other studies have investigated the diagnostic accuracy of AI for gastric polyps and the invasion depth of gastric cancers [30,31]. Nevertheless, the rate of detection of early gastric cancer, which allows for prompt intervention and increased survival rates, remains low. Multiple studies on the AI-assisted diagnosis of early gastric cancer have been conducted in the past 5 years, but results have been inconsistent and highly variable. Furthermore, the role of AI in early gastric

cancer is more important than in advanced gastric cancer since early gastric cancer is not easily identified in clinical practice; however, to our knowledge, past syntheses appear to have limited focus on the population with early gastric cancer. Thus, we investigated the performance of AI-assisted endoscopic diagnosis of early gastric cancer.

Methods**Definition**

Early gastric cancer was defined as mucosal and submucosal (T1) gastric cancer irrespective of lymph node involvement. Studies involving advanced gastric cancer, precancerous lesions such as intestinal metaplasia and dysplasia, and gastric cancer without specific annotations were excluded. The accuracy of AI was defined as the area under the hierarchical summary receiver operating characteristic curve or the area under the curve (AUC).

Study Search and Selection Strategy

This meta-analysis was performed according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. We systematically searched the PubMed, Embase, Cochrane Library, and Web of Science databases for studies that assessed the diagnostic accuracy of AI in early gastric cancer from endoscopic images from database inception to June 2020. We used “gastric cancer,” “endoscopy,” and “artificial intelligence” as relevant terms with Boolean operators “OR” and “AND” (Multimedia Appendix 1). Two authors, P-CC and L-YR, independently screened the study titles and abstracts. Studies that used AI to diagnose early gastric cancer from endoscopic images were included. Studies that did not provide a 2×2 contingency table were not included in the final analysis. This study was registered in PROSPERO (registration CRD42020193223).

Study Quality Assessment and Data Extraction

The quality of the included studies was assessed independently by 2 authors (P-CC and L-YR) on the basis of the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2), and all disagreement was resolved through discussion with the third author (Y-NK). The assessment included risk of bias and applicability to the QUADAS-2 domains: patient selection, index test, reference standard, and flow and timing. From the included studies, we extracted data on the number of endoscopic images of lesions diagnosed as early gastric cancer (ie, true positive), the number of endoscopic images of benign lesions misdiagnosed as malignant (ie, false positive), the number of endoscopic images of malignant lesions misdiagnosed as benign (ie, false negative), and the number of endoscopic images of benign lesions correctly diagnosed as benign (ie, true negative). We also extracted data on the country of origin, AI methods, and image modalities used.

Study Outcomes and Statistical Analysis

The primary outcome was the accuracy of AI to diagnose early gastric cancer from endoscopic images. Secondary outcomes focused on the sensitivity analysis of (a) different AI methods, (b) endoscopic imaging modalities, (c) studies that compared AI and endoscopist performance, (d) studies that evaluated larger gastric lesions (>20 mm), (e) studies that simply differentiated abnormal and normal lesions rather than using pathological staging, and (f) studies that separated the training and testing data sets during AI training. Sensitivity analysis was conducted if a subgroup contained more than two studies. We only assessed the heterogeneity of the included studies. Following extraction, the data were primarily analyzed using STATA 14 (StataCorp LP, StataCorp) except for subgroups with fewer than four studies. The `midas` and `metandi` commands were used to determine sensitivity, specificity, and AUC and analyze the summary receiver operating characteristic (SROC) and hierarchical summary receiver operating characteristic (HSROC) curves. Basic formulas for the analyses were as follows:

$$\ln \text{DOR} = (\text{logit TPR}) - (\text{logit FPR}) \quad (1)$$

$$\text{proxy for the threshold} = (\text{logit TPR}) - (\text{logit FPR}) \quad (2)$$

$$\text{TPR of SROC} = 1 / [1 / (1 + e^{a/(1-b)}) \times (\text{FPR}/(1-\text{FPR}))^{(1+b)/(1-b)}] \quad (3)$$

In the formulas, “a” is the intercept, “b” is slope, and DOR refers to the diagnostic odds ratio. Moreover, TPR is the true positive rate, and FPR is the false-positive rate. The `modchk` tool was used to examine goodness-of-fit and bivariate normality before SROC analysis in a bivariate mixed-effects model. The `metabias` command and the `pubbias` syntax were used to perform the Egger test and Deeks funnel plot asymmetry tests, respectively. The Egger test for diagnostic meta-analysis was based on the formula proposed by Hasselblad and Hedges, and the formula is mainly to detect publication bias detection via testing standard normal deviate among the included studies [32,33].

$$\text{standard normal deviate} = a + b \times \text{SE}_{(d)}^{-1} \quad (4)$$

In the regression model, with intercept “a” and slope “b,” the standard normal deviation could be estimated by using diagnostic *d* divided by SE of the diagnostic *d*. The `metaprop` package in STATA was mainly used to synthesize the sensitivity and specificity. I^2 statistics were used to determine levels of heterogeneity via the formula as follows:

$$I^2 = ((Q - df)/Q) \times 100 \quad (5)$$

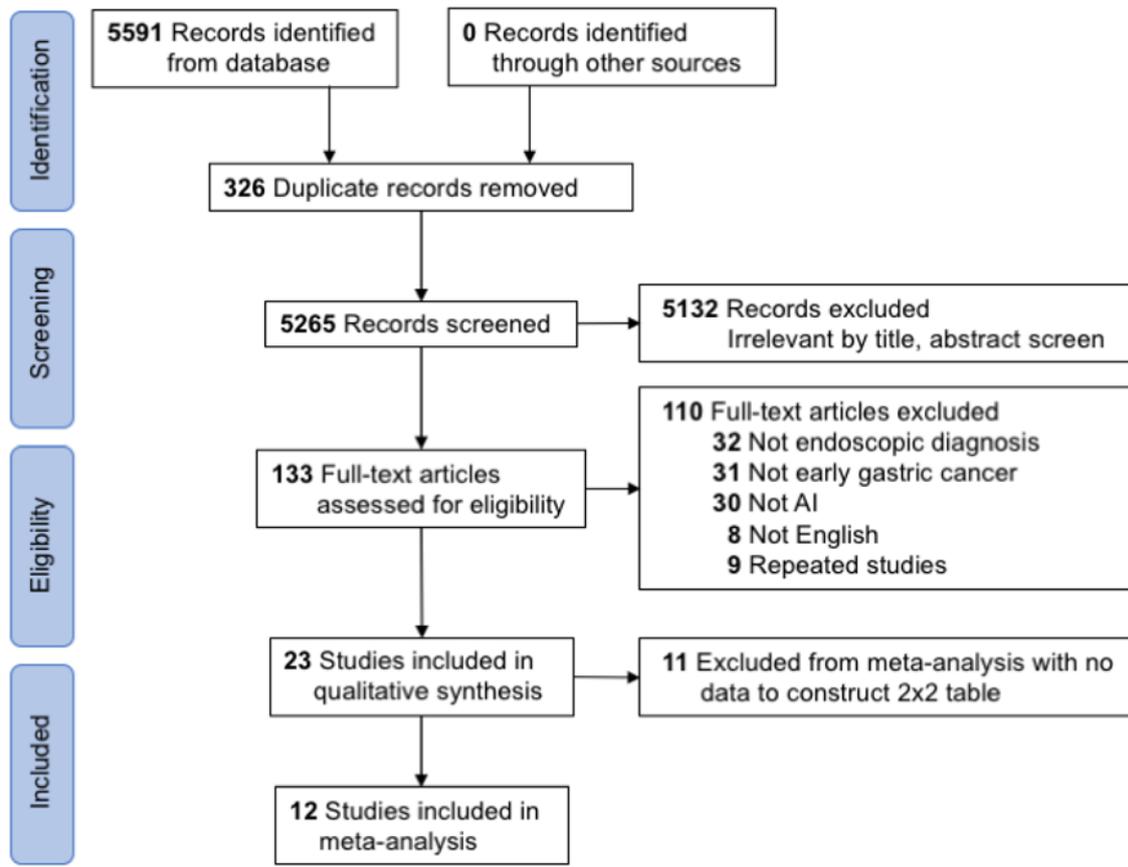
where *Q* refers to Cochran *Q*, and *df* is the degree of freedom. Because R software (The R Foundation) does not restrict the number of observations used in the meta-analysis, it was used for sensitivity analysis if subgroups consisted of fewer than four studies. Indeed, a meta-analysis in R could be carried out when more than two studies report the same outcome by pooling data with logit transformation and Clopper-Pearson interval method (also called exact binomial interval) based on inverse variance. Function `metaprop` in package `meta` for R was applied to carry out sensitivity analysis, and the `mada` package in R was used to calculate the pooled accuracy. Besides, the `metagen` package in R was used to synthesize endoscopist performance because of the lack of detailed data on each endoscopist.

Results

Literature Search and Review

Of the 5591 studies identified in the literature review, 5265 underwent title and abstract screening after duplication removal. The flowchart of the literature review process was constructed according to the PRISMA flowchart format (Figure 1). We excluded 5132 irrelevant studies and assessed the eligibility of the remaining 133 studies through full-text reading. Studies evaluating nonearly gastric cancer (eg, advanced gastric cancer and metaplasia) were excluded. Overall, 23 studies investigated the performance of AI on early gastric cancer diagnosis from endoscopic images. Finally, 12 studies comprising a total of 11,685 cases were included in the meta-analysis [34-45].

Figure 1. Flowchart of the study selection process according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) format. AI: artificial intelligence.



Study Description and Bias Assessment

Detailed information on the 12 studies is listed in [Table 1](#). All studies were conducted in Asia, including Japan (k=8), China (k=2), and Korea (k=2), in or after 2012. All were case control studies with testing data sets containing 81 to 3390 images. Patients in 10 studies had pathological proof of early gastric cancer, whereas in the other 2 studies, the endoscopic images were collected through description. White light imaging (WLI), narrow-band imaging (NBI), flexible spectral imaging color enhancement, and mixed imaging modalities were used in 4 (33%), 2 (17%), 1 (8%), and 2 (17%) studies, respectively. Moreover, 8 (67%) studies used deep learning methods (eg, convolutional neural networks) as their AI backbone, and 3 (25%) studies employed nondeep learning methods (support

vector machines and discriminant analysis of principal components). Comparisons of the diagnostic performance of AI and endoscopists were conducted in 3 (25%) studies, and 2 (17%) studies included endoscopic images of small lesions (<20 mm) in early gastric cancer. In 3 (25%) studies, the training and testing data sets were not separated for AI training. [Table 1](#) presents a detailed description of the 12 studies.

We also assessed the quality of the studies along with the risk of bias according to the revised QUADAS-2 tool ([Multimedia Appendix 2](#)). All studies, including the 3 that failed to separate the training and testing data sets, had high bias risks for patient selection because of their retrospective design. Moreover, 2 (17%) studies assessed early gastric cancer but did not mention pathological staging. Thus, they were classified as having a high risk of bias for the index test.

Table 1. Characteristics of the included studies.

Study ID	Country of origin	Testing image number	Reference standard	Image modality	AI ^a method	AI training and testing data set	Standard reference	Endoscopic comparison	Other information
Kubota et al, 2012 [43]	Japan	902	Pathology	Not mentioned ^b	Multilayer neural network	Not separated	Unclear	No	Detected with pathological grading prediction
Miyaki et al, 2013 [44]	Japan	92	Pathology	FICE ^c	SVM ^d (scale-invariant feature transform)	Separated	Pathology	No	Differentiated early gastric cancer from noncancerous tissues
Liu et al, 2016 [41]	China	400	Pathology	Not mentioned ^b	Principal component discriminant analysis (YCbCr color space)	Separated	Pathology	No	Differentiated early gastric cancer from normal tissues
Kanesaka et al, 2018 [37]	Japan	81	Pathology	NBI ^e	SVM (grey-level co-occurrence feature)	Separated	Pathology	No	Included only depressed type early gastric cancers that were <10 mm in size
Sakai et al, 2018 [36]	Japan	926	Pathology	WLI ^f	CNN ^g (GoogLeNet)	Not separated	Pathology	No	— ^h
Yamakawa et al, 2018 [45]	Japan	817	Unclear ⁱ	Not mentioned ^j	Not mentioned	Separated	Unclear	No	Differentiated early gastric cancer from nonneoplastic tissues
Cho et al, 2019 [35]	Korea	200	Pathology	WLI	CNN (Inception-Resnet-v2)	Separated	Pathology	Yes	Detected early gastric cancer with pathological grading prediction
Namikawa et al, 2019 [34]	Japan	1479 ^j	Unclear ⁱ	WLI, NBI, Chromo ^k	CNN	Separated	Pathology	No	Differentiated early gastric cancer from gastric ulcers
Wu et al, 2019 [39]	China	200	Pathology	WLI, NBI, BLI ^l	CNN (VGG16 + Resnet-50)	Separated	Pathology	Yes	Differentiated early gastric cancer from gastritis and normal tissues
Yoon et al 2019 [42]	Korea	3390	Pathology	WLI	CNN (VGG16)	Not separated	Pathology	No	—
Horiuchi et al, 2020 [38]	Japan	258	Pathology	NBI	CNN (GoogLeNet)	Separated	Pathology	No	Differentiated early gastric cancer from <i>Helicobacter pylori</i> -related gastritis
Ikenoyama et al, 2020 [40]	Japan	2940	Pathology	WLI	CNN (Single-shot multiBox Detector)	Separated	Pathology	Yes	Included only early gastric lesions that were <20 mm

^aAI: artificial intelligence.

^bStudies that failed to mention imaging modalities.

^cFICE: flexible spectral imaging color enhancement.

^dSVM: support vector machine.

^eNBI: narrow-band imaging.

^fWLI: white light imaging.

^gCNN: convolutional neural network.

^hNot available.

ⁱStudies that mentioned early gastric cancer but without reference to pathological staging.

^jStudies were reported in meeting abstracts.

^kChromo: chromoendoscopy.

^lBLI: blue laser imaging.

Diagnostic Performance of AI for Early Gastric Cancer

To assess the diagnostic ability of AI to detect early gastric cancer from endoscopic images, we performed a meta-analysis on the selected 12 studies. Goodness-of-fit (Figure 2A) and bivariate normality (Figure 2B) demonstrated that the included data were appropriate for further analysis. The pooled sensitivity and specificity of AI were 0.86 (95% CI 0.75-0.92) and 0.90 (95% CI 0.84-0.93), respectively (Figures 2C and 2D). Empirical Bayesian predictions were consistent with the observed sensitivity and specificity (Multimedia Appendix 3). Highly heterogeneous estimates ($I^2 > 90%$) necessitated subgroup analysis and sensitivity analysis. Laminated figures of the SROC and HSROC plots indicate an AUC of 0.94 (95% CI 0.92-0.96) with a confidence region (Figure 3A). However, the scatter matrix (Multimedia Appendix 4) suggests that in clinical practice, diagnosis of early gastric cancer may not substantially benefit from AI assistance. The Deeks funnel plot asymmetry test (Figure 3B) and Egger test (Multimedia Appendix 5) did not detect significant publication bias in the pooled results of AI-assisted diagnosis of early gastric cancer.

We assessed the diagnostic performance of various AI methods and endoscopic imaging modalities for early gastric cancer (Table 2). The pooled sensitivity and specificity in studies using deep learning methods were 0.84 (95% CI 0.69-0.93) and 0.88 (95% CI 0.80-0.93), respectively. Studies using nondeep learning methods had a pooled sensitivity and specificity of 0.91 (95% CI 0.86-0.95) and 0.90 (95% CI 0.87-0.93), respectively. The accuracy of the nondeep learning group (AUC=0.96) was higher than that of the deep learning group (AUC=0.93; Multimedia Appendices 6 and 7).

For endoscopic imaging modalities, studies using WLI had a sensitivity and specificity of 0.73 (95% CI 0.42-0.91) and 0.89 (95% CI 0.76-0.96), respectively. Studies using NBI reported a sensitivity and specificity of 0.96 (95% CI 0.92-0.98) and 0.83 (95% CI 0.54-0.95), respectively. The accuracy of the NBI group (AUC=0.96) was higher than that of the WLI group (AUC=0.90; Multimedia Appendices 8 and 9). Table S1 (Multimedia Appendix 10) shows a comparison of the diagnostic performance of AI and endoscopists for early gastric cancer from the three studies (n=91).

Figure 2. Overall sensitivity and specificity of artificial intelligence–assisted diagnosis of early gastric cancer. (A) Goodness-of-fit; (B) bivariate normality; (C) forest plot of overall sensitivity; and (D) forest plot of overall specificity. FP: false positive; TN: true negative.

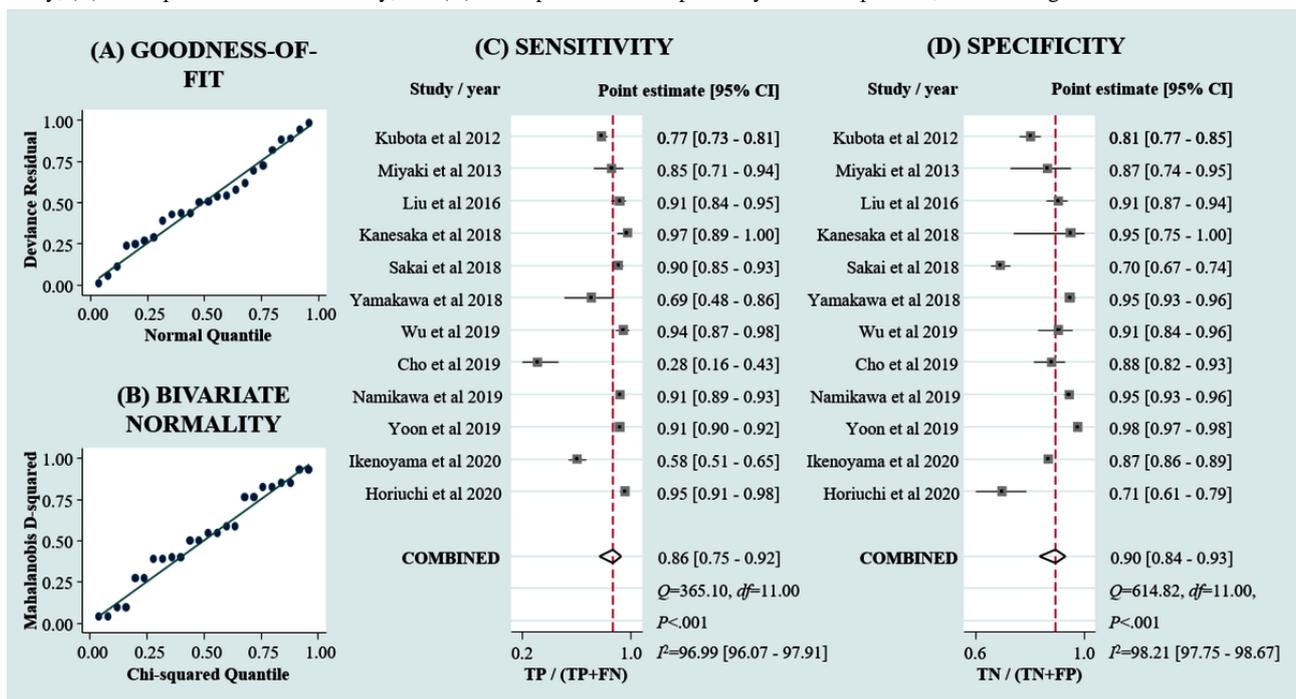
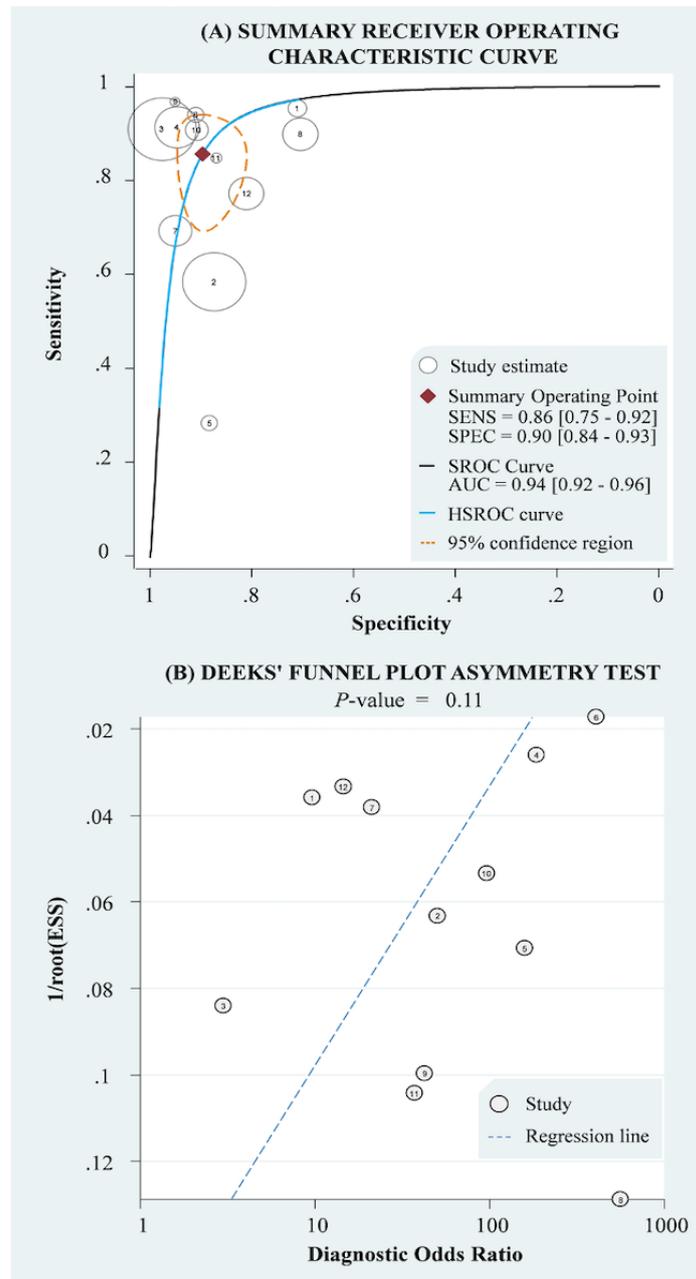


Figure 3. Summary receiver operating characteristic curve, HSROC, AUC, and the Deeks funnel plot asymmetry test of artificial intelligence–assisted diagnosis of early gastric cancer. AUC: area under the curve; ESS: effective sample sizes; HSROC: hierarchical summary receiver operating characteristic; SENS: sensitivity; SPEC: specificity; SROC: summary receiver operator characteristic.



Additional Analysis

We excluded some studies with a high risk of bias and performed sensitivity analysis on the remaining studies (Tables S2-S5 [Multimedia Appendices 11-14](#)). Furthermore, we also examined how the results were affected by studies with unknown AI methods. Sensitivity analyses indicated that pooled estimates were not seriously affected by the factors (Table 2). Lower

heterogeneity and specificity were observed in endoscopist performance when we excluded studies that only evaluated small lesions and studies that predicted pathological staging (Tables S2 and S3 in [Multimedia Appendices 11 and 12](#)). Lower heterogeneity was also noted in WLI subgroups if the training and testing data sets were separated for AI training (Table S4 in [Multimedia Appendix 13](#)). No other additional analyses provided credible evidence.

Table 2. Pooled sensitivity, specificity, and accuracy of the studies included in the meta-analysis and sensitivity analysis.

Group (studies and number of patients)	Sensitivity (95% CI)	I^2 , %	Specificity (95% CI)	I^2 , %	AUC ^a
Overall (12 studies, n=11,685)	0.86 (0.75-0.92)	97	0.90 (0.84-0.93)	97	0.94
Subgroup analysis on different AI^b methods					
Deep learning (8 studies, n=10,295)	0.84 (0.69-0.93)	98	0.88 (0.80-0.93)	98	0.93
Nondeep learning (3 studies, n=573)	0.91 (0.86-0.95)	18	0.90 (0.87-0.93)	0	0.96
Subgroup analysis on various imaging modalities					
WLI ^c (4 studies, n=7456)	0.73 (0.42-0.91)	99	0.89 (0.76-0.96)	99	0.902
NBI ^d (2 studies, n=339)	0.96 (0.92-0.98)	0	0.83 (0.54-0.95)	51	0.959
Sensitivity analysis					
Excluding studies with unknown method (11 studies, n=10,868)	0.87 (0.76-0.93)	97	0.89 (0.83-0.93)	97	0.936
Excluding studies with sample size <100 (10 studies, n=11,512)	0.84 (0.71-0.92)	97	0.89 (0.83-0.94)	98	0.932
Excluding studies without separation of testing data (9 studies, n=6467)	0.85 (0.70-0.93)	96	0.90 (0.86-0.93)	91	0.934
Excluding studies with any situation abovementioned (6 studies, n=5477)	0.84 (0.62-0.94)	98	0.89 (0.83-0.93)	92	0.923

^aAUC: area under the curve.

^bAI: artificial intelligence.

^cWLI: white light imaging.

^dNBI: narrow-band imaging.

Discussion

Principal Findings

To our knowledge, this was the first systematic review and meta-analysis of AI-assisted endoscopic diagnosis of early gastric cancer. The accuracy, sensitivity, and specificity were 0.94, 0.86, and 0.90, respectively. High heterogeneity was noted. Sensitivity analysis revealed less heterogeneity in studies using nondeep learning AI methods and endoscopic NBI.

Our results indicate good sensitivity and specificity of AI-assisted detection of early gastric cancer. However, high heterogeneity was also noted among the included studies, which may be attributed to between-study differences in machine learning methods and imaging modalities [46]. In a meta-analysis of AI prediction of colonic polyp histology, AI performance was better when deep learning was used as a backbone and when NBI was used to identify the lesions [46]. In this study, we also investigated the roles of various machine learning methods and imaging modalities. Unfortunately, only 2 studies in the deep learning subgroup used the same deep learning algorithm, and no two studies in the nondeep learning subgroup classified the lesions according to the same features. Only 6 studies specified their endoscopic imaging modalities. Less heterogeneity was observed in the nondeep learning and NBI groups, possibly because of the compliance of early gastric cancer diagnosis to the vessel plus surface classification system under NBI. This indicates that nondeep learning methods and NBI may provide more consistent results and can be applied in clinical practice earlier than deep learning methods and WLI. Further investigations are warranted.

We assessed the diagnostic performance of AI and endoscopists (n=91) for early gastric cancer detection, which was compared in 3 studies. The endoscopists were assigned to only 1 subgroup because of the inconsistent definitions of expert and nonexpert endoscopists between studies. The sensitivity and specificity of AI were 0.67 and 0.87, respectively, and those of the endoscopists were 0.68 and 0.92, respectively. In both groups, diagnostic performance varied widely with high heterogeneity. The diagnostic performance of AI was better than that of WLI compared with other studies; a meta-analysis reported a pooled sensitivity and specificity of 48% and 67% between endoscopists and WLI, whereas those between endoscopists and NBI were 83% and 97%, respectively [47]. In this study, AI and endoscopist performance were comparable in individual studies, but this effect diminished when studies were pooled. Further research comparing AI and endoscopist performance for early gastric cancer diagnosis is required.

Only 2 of the included studies evaluated only small lesions [37,40]. Smaller lesions and mucosal lesions were less accurately detected by AI [42]. Kanekawa et al [37] included only depressed and small (<10 mm) lesions, and the AI system of nondeep learning methods was trained using a small data set of 126 images from NBI. In another study, early gastric cancer lesions less than 20 mm in diameter were included in the WLI testing data set, and the deep learning AI system was trained using a data set of 13,584 images of early and advanced gastric cancer [40]. Because these 2 studies used distinct materials and methods, their findings may not be representative. The accuracy of AI-assisted detection of small gastric cancer lesions warrants further investigation.

Some studies have explored the application of AI to other aspects of gastroendoscopy. For example, Wu et al [39] used AI to monitor endoscopic blind spots and identify regions indicative of early gastric cancer. A randomized controlled trial in China reported that AI reduced the rate of endoscopic blind spots [48]. Other studies have tested the accuracy of AI in predicting the invasion depth of gastric cancer—conventionally assessed through endoscopic ultrasound—from endoscopic images. In their study of AI-assisted simultaneous detection of gastric cancer and invasion depth, Yoon et al [42] reported a sensitivity and specificity of invasion depth of 79.2% and 77.8%, respectively. In a study by Zhu et al [31], the predicted sensitivity and specificity from the T1 to the T4 stage were 76% and 96%, respectively. Nevertheless, relevant evidence is limited, and further investigation is required.

The considerable advancement of AI in precise image recognition challenges the roles of physicians in disease diagnosis. AI systems offer certain advantages over physician diagnosis, the foremost of which are faster image processing rates and continuous work. In all included studies that specified image processing time, that of AI systems was shorter than that of endoscopists. AI assistance may reduce the risk of human error that arises from performing numerous endoscopic examinations. Moreover, the training of AI systems is considerably faster and less complicated than that of endoscopists. Well-trained AI systems learn from analyzing numerous images, whereas endoscopists rely on their individual skills and clinical experience. Training endoscopists is expensive and time-consuming because of the steep learning curve for the various image-enhancing techniques. In addition, AI may work as a double-check system during or after endoscopy, given its high sensitivity and specificity. AI allows for a second opinion, which is particularly valuable now that gastroendoscopy has been popularized and nationwide screening for gastric cancer has been implemented.

Limitations

Our study had several limitations. First, all the included studies were retrospective case control studies performed in Asia, some of which compared early gastric cancer and normal gastric tissues, and some compared benign gastric lesions such as ulcers and gastritis. The possibility of selection bias cannot be ruled out. A randomized controlled trial comparing the diagnostic performance of AI and endoscopists for early and advanced gastric cancer (NCT04040374) is currently underway. Second, all the studies identified gastric lesions from still, clear, endoscopic images; images with blood or mucus were excluded. In daily practice, however, gastroendoscopy is recorded in video format, and still images are only captured for suspicious lesions. Blood, food debris, mucus, and foam, which reduce the accuracy of AI, are commonly encountered during examination [39]. Several studies have reported excellent accuracy of AI systems in recognizing gastric cancer from endoscopic video [39,49]. However, further studies and faster image processing rates are necessary. Third, our pooled estimates were highly heterogeneous, and the subgroup and sensitivity analyses did not substantially reduce heterogeneity. The statistical heterogeneity may be ascribed to differences in the AI methods and endoscopic imaging techniques. These potential sources of heterogeneity should be discussed in future research. At present, AI may assist endoscopists in double-checking suspicious lesions.

Conclusions

To our knowledge, this is the first meta-analysis of the performance of AI in detecting early gastric cancer using endoscopic images. The available evidence suggests that AI can support the diagnosis of early gastric cancer; however, the collocation of imaging techniques and optimal algorithm remains unclear. Larger prospective cohort studies should be conducted to further validate the diagnostic accuracy of AI. Moreover, competing models of AI for the detection of early gastric cancer are worthy of future investigation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary File 1. Search strategy (primary search strategy).
[\[PDF File \(Adobe PDF File\), 402 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Supplementary File 2. Study quality assessment according to the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies [revised]).
[\[PDF File \(Adobe PDF File\), 437 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Supplementary File 3. Forest plot of empirical Bayes predicted and observed findings.
[\[PDF File \(Adobe PDF File\), 564 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Supplementary File 4. Scatter matrix.

[\[PDF File \(Adobe PDF File\), 463 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Supplementary File 5. Egger's test.

[\[PDF File \(Adobe PDF File\), 452 KB-Multimedia Appendix 5\]](#)

Multimedia Appendix 6

Supplementary File 6. Subgroup analysis for studies that used deep learning.

[\[PDF File \(Adobe PDF File\), 511 KB-Multimedia Appendix 6\]](#)

Multimedia Appendix 7

Supplementary File 7. Subgroup analysis for studies without deep learning.

[\[PDF File \(Adobe PDF File\), 424 KB-Multimedia Appendix 7\]](#)

Multimedia Appendix 8

Supplementary File 8. Subgroup analysis for studies that used white light image.

[\[PDF File \(Adobe PDF File\), 506 KB-Multimedia Appendix 8\]](#)

Multimedia Appendix 9

Supplementary File 9. Subgroup analysis for studies that used narrow band imaging techniques.

[\[PDF File \(Adobe PDF File\), 424 KB-Multimedia Appendix 9\]](#)

Multimedia Appendix 10

Supplementary Table 1. Characteristics of the studies that compared diagnostic performance of artificial intelligence to endoscopists and its sensitivity analysis.

[\[PDF File \(Adobe PDF File\), 350 KB-Multimedia Appendix 10\]](#)

Multimedia Appendix 11

Supplementary Table 2. Sensitivity analysis of the studies that included gastric lesions other than small gastric cancer lesions.

[\[PDF File \(Adobe PDF File\), 404 KB-Multimedia Appendix 11\]](#)

Multimedia Appendix 12

Supplementary Table 3. Sensitivity analysis of the studies that do not detect early gastric cancer lesions based on pathological grading.

[\[PDF File \(Adobe PDF File\), 403 KB-Multimedia Appendix 12\]](#)

Multimedia Appendix 13

Supplementary Table 4. Sensitivity analysis of the studies that separated training and testing data set during artificial intelligence training.

[\[PDF File \(Adobe PDF File\), 403 KB-Multimedia Appendix 13\]](#)

Multimedia Appendix 14

Supplementary Table 5. Sensitivity analysis of the studies with low risk on index test.

[\[PDF File \(Adobe PDF File\), 577 KB-Multimedia Appendix 14\]](#)

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018 Nov;68(6):394-424 [\[FREE Full text\]](#) [doi: [10.3322/caac.21492](https://doi.org/10.3322/caac.21492)] [Medline: [30207593](https://pubmed.ncbi.nlm.nih.gov/30207593/)]
2. Ouyang G, Pan G, Liu Q, Wu Y, Liu Z, Lu W, et al. The global, regional, and national burden of pancreatitis in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *BMC Med* 2020 Dec 10;18(1):388 [\[FREE Full text\]](#) [doi: [10.1186/s12916-020-01859-5](https://doi.org/10.1186/s12916-020-01859-5)] [Medline: [33298026](https://pubmed.ncbi.nlm.nih.gov/33298026/)]

3. Ford AC, Forman D, Hunt RH, Yuan Y, Moayyedi P. Helicobacter pylori eradication therapy to prevent gastric cancer in healthy asymptomatic infected individuals: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2014 May 20;348(may20 1):g3174-g3174 [FREE Full text] [doi: [10.1136/bmj.g3174](https://doi.org/10.1136/bmj.g3174)] [Medline: [24846275](https://pubmed.ncbi.nlm.nih.gov/24846275/)]
4. Yaghoobi M, McNabb-Baltar J, Bijarchi R, Hunt RH. What is the quantitative risk of gastric cancer in the first-degree relatives of patients? A meta-analysis. *World J Gastroenterol* 2017 Apr 07;23(13):2435-2442 [FREE Full text] [doi: [10.3748/wjg.v23.i13.2435](https://doi.org/10.3748/wjg.v23.i13.2435)] [Medline: [28428723](https://pubmed.ncbi.nlm.nih.gov/28428723/)]
5. Rawla P, Barsouk A. Epidemiology of gastric cancer: global trends, risk factors and prevention. *Prz Gastroenterol* 2019;14(1):26-38 [FREE Full text] [doi: [10.5114/pg.2018.80001](https://doi.org/10.5114/pg.2018.80001)] [Medline: [30944675](https://pubmed.ncbi.nlm.nih.gov/30944675/)]
6. Ajani JA, Lee J, Sano T, Janjigian YY, Fan D, Song S. Gastric adenocarcinoma. *Nat Rev Dis Primers* 2017 Jun 01;3(1):17036. [doi: [10.1038/nrdp.2017.36](https://doi.org/10.1038/nrdp.2017.36)] [Medline: [28569272](https://pubmed.ncbi.nlm.nih.gov/28569272/)]
7. Katai H, Ishikawa T, Akazawa K, Isobe Y, Miyashiro I, Oda I, Registration Committee of the Japanese Gastric Cancer Association. Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001-2007). *Gastric Cancer* 2018 Jan 17;21(1):144-154. [doi: [10.1007/s10120-017-0716-7](https://doi.org/10.1007/s10120-017-0716-7)] [Medline: [28417260](https://pubmed.ncbi.nlm.nih.gov/28417260/)]
8. Fock KM, Talley N, Moayyedi P, Hunt R, Azuma T, Sugano K, Asia-Pacific Gastric Cancer Consensus Conference. Asia-Pacific consensus guidelines on gastric cancer prevention. *J Gastroenterol Hepatol* 2008 Mar;23(3):351-365. [doi: [10.1111/j.1440-1746.2008.05314.x](https://doi.org/10.1111/j.1440-1746.2008.05314.x)] [Medline: [18318820](https://pubmed.ncbi.nlm.nih.gov/18318820/)]
9. Choi KS, Jun JK, Park E, Park S, Jung KW, Han MA, et al. Performance of different gastric cancer screening methods in Korea: a population-based study. *PLoS One* 2012 Nov 29;7(11):e50041 [FREE Full text] [doi: [10.1371/journal.pone.0050041](https://doi.org/10.1371/journal.pone.0050041)] [Medline: [23209638](https://pubmed.ncbi.nlm.nih.gov/23209638/)]
10. Yoshida N, Doyama H, Yano T, Horimatsu T, Uedo N, Yamamoto Y, et al. Early gastric cancer detection in high-risk patients: a multicentre randomised controlled trial on the effect of second-generation narrow band imaging. *Gut* 2021 Jan 02;70(1):67-75 [FREE Full text] [doi: [10.1136/gutjnl-2019-319631](https://doi.org/10.1136/gutjnl-2019-319631)] [Medline: [32241898](https://pubmed.ncbi.nlm.nih.gov/32241898/)]
11. Pasechnikov V, Chukov S, Fedorov E, Kikuste I, Leja M. Gastric cancer: prevention, screening and early diagnosis. *World J Gastroenterol* 2014 Oct 14;20(38):13842-13862 [FREE Full text] [doi: [10.3748/wjg.v20.i38.13842](https://doi.org/10.3748/wjg.v20.i38.13842)] [Medline: [25320521](https://pubmed.ncbi.nlm.nih.gov/25320521/)]
12. Abbasi J. Artificial Intelligence Tools for Sepsis and Cancer. *JAMA* 2018 Dec 11;320(22):2303. [doi: [10.1001/jama.2018.19383](https://doi.org/10.1001/jama.2018.19383)] [Medline: [30535201](https://pubmed.ncbi.nlm.nih.gov/30535201/)]
13. Abbasi J. Artificial Intelligence-Based Skin Cancer Phone Apps Unreliable. *JAMA* 2020 Apr 14;323(14):1336. [doi: [10.1001/jama.2020.4543](https://doi.org/10.1001/jama.2020.4543)] [Medline: [32286633](https://pubmed.ncbi.nlm.nih.gov/32286633/)]
14. Abbasi J. Artificial Intelligence Improves Breast Cancer Screening in Study. *JAMA* 2020 Feb 11;323(6):499. [doi: [10.1001/jama.2020.0370](https://doi.org/10.1001/jama.2020.0370)] [Medline: [32044919](https://pubmed.ncbi.nlm.nih.gov/32044919/)]
15. Hwang TJ, Kesselheim AS, Vokinger KN. Lifecycle Regulation of Artificial Intelligence- and Machine Learning-Based Software Devices in Medicine. *JAMA* 2019 Dec 17;322(23):2285-2286. [doi: [10.1001/jama.2019.16842](https://doi.org/10.1001/jama.2019.16842)] [Medline: [31755907](https://pubmed.ncbi.nlm.nih.gov/31755907/)]
16. Matheny ME, Whicher D, Thadaney Israni S. Artificial Intelligence in Health Care: A Report From the National Academy of Medicine. *JAMA* 2020 Feb 11;323(6):509-510. [doi: [10.1001/jama.2019.21579](https://doi.org/10.1001/jama.2019.21579)] [Medline: [31845963](https://pubmed.ncbi.nlm.nih.gov/31845963/)]
17. Rubin R. Artificial Intelligence for Cervical Precancer Screening. *JAMA* 2019 Feb 26;321(8):734. [doi: [10.1001/jama.2019.0888](https://doi.org/10.1001/jama.2019.0888)] [Medline: [30806677](https://pubmed.ncbi.nlm.nih.gov/30806677/)]
18. Shortliffe EH, Sepúlveda MJ. Clinical Decision Support in the Era of Artificial Intelligence. *JAMA* 2018 Dec 04;320(21):2199-2200. [doi: [10.1001/jama.2018.17163](https://doi.org/10.1001/jama.2018.17163)] [Medline: [30398550](https://pubmed.ncbi.nlm.nih.gov/30398550/)]
19. Voelker R. Cardiac Ultrasound Uses Artificial Intelligence to Produce Images. *JAMA* 2020 Mar 17;323(11):1034. [doi: [10.1001/jama.2020.2547](https://doi.org/10.1001/jama.2020.2547)] [Medline: [32181835](https://pubmed.ncbi.nlm.nih.gov/32181835/)]
20. Samoli S, Lopez Cobo M, Gomez Gutierrez E, De Prato G, Martinez-Plumed F, Delipetrev B. Defining Artificial Intelligence: Towards an operational definition and taxonomy of artificial intelligence. Luxembourg: Publications Office of the European Union; 2020.
21. Russell S, Norvig P. Artificial Intelligence: A Modern Approach. *Artificial Intelligence: A Modern Approach*. Pearson Education, Inc; 2010:A.
22. Greenspan H, van Ginneken B, Summers RM. Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Trans Med Imaging* 2016 May;35(5):1153-1159. [doi: [10.1109/tmi.2016.2553401](https://doi.org/10.1109/tmi.2016.2553401)]
23. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017 Dec;2(4):230-243 [FREE Full text] [doi: [10.1136/svn-2017-000101](https://doi.org/10.1136/svn-2017-000101)] [Medline: [29507784](https://pubmed.ncbi.nlm.nih.gov/29507784/)]
24. Bang CS, Lee JJ, Baik GH. Artificial Intelligence for the Prediction of Helicobacter Pylori Infection in Endoscopic Images: Systematic Review and Meta-Analysis Of Diagnostic Test Accuracy. *J Med Internet Res* 2020 Sep 16;22(9):e21983 [FREE Full text] [doi: [10.2196/21983](https://doi.org/10.2196/21983)] [Medline: [32936088](https://pubmed.ncbi.nlm.nih.gov/32936088/)]
25. Ćirković A. Evaluation of Four Artificial Intelligence-Assisted Self-Diagnosis Apps on Three Diagnoses: Two-Year Follow-Up Study. *J Med Internet Res* 2020 Dec 04;22(12):e18097 [FREE Full text] [doi: [10.2196/18097](https://doi.org/10.2196/18097)] [Medline: [33275113](https://pubmed.ncbi.nlm.nih.gov/33275113/)]

26. Liu T, Tsang W, Huang F, Lau OY, Chen Y, Sheng J, et al. Patients' Preferences for Artificial Intelligence Applications Versus Clinicians in Disease Diagnosis During the SARS-CoV-2 Pandemic in China: Discrete Choice Experiment. *J Med Internet Res* 2021 Feb 23;23(2):e22841 [FREE Full text] [doi: [10.2196/22841](https://doi.org/10.2196/22841)] [Medline: [33493130](https://pubmed.ncbi.nlm.nih.gov/33493130/)]
27. Shen J, Chen J, Zheng Z, Zheng J, Liu Z, Song J, et al. An Innovative Artificial Intelligence-Based App for the Diagnosis of Gestational Diabetes Mellitus (GDM-AI): Development Study. *J Med Internet Res* 2020 Sep 15;22(9):e21573 [FREE Full text] [doi: [10.2196/21573](https://doi.org/10.2196/21573)] [Medline: [32930674](https://pubmed.ncbi.nlm.nih.gov/32930674/)]
28. Gonçalves WG, Dos Santos MHP, Lobato FMF, Ribeiro-Dos-Santos Â, de Araújo GS. Deep learning in gastric tissue diseases: a systematic review. *BMJ Open Gastroenterol* 2020 Mar 26;7(1):e000371 [FREE Full text] [doi: [10.1136/bmjgast-2019-000371](https://doi.org/10.1136/bmjgast-2019-000371)] [Medline: [32337060](https://pubmed.ncbi.nlm.nih.gov/32337060/)]
29. Luo H, Xu G, Li C, He L, Luo L, Wang Z, et al. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. *The Lancet Oncology* 2019 Dec;20(12):1645-1654. [doi: [10.1016/S1470-2045\(19\)30637-0](https://doi.org/10.1016/S1470-2045(19)30637-0)] [Medline: [31591062](https://pubmed.ncbi.nlm.nih.gov/31591062/)]
30. Zhang X, Chen F, Yu T, An J, Huang Z, Liu J, et al. Real-time gastric polyp detection using convolutional neural networks. *PLoS One* 2019 Mar 25;14(3):e0214133 [FREE Full text] [doi: [10.1371/journal.pone.0214133](https://doi.org/10.1371/journal.pone.0214133)] [Medline: [30908513](https://pubmed.ncbi.nlm.nih.gov/30908513/)]
31. Zhu Y, Wang Q, Xu M, Zhang Z, Cheng J, Zhong Y, et al. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. *Gastrointest Endosc* 2019 Apr;89(4):806-815.e1. [doi: [10.1016/j.gie.2018.11.011](https://doi.org/10.1016/j.gie.2018.11.011)] [Medline: [30452913](https://pubmed.ncbi.nlm.nih.gov/30452913/)]
32. Hasselblad V, Hedges LV. Meta-analysis of screening and diagnostic tests. *Psychol Bull* 1995 Jan;117(1):167-178. [doi: [10.1037/0033-2909.117.1.167](https://doi.org/10.1037/0033-2909.117.1.167)] [Medline: [7870860](https://pubmed.ncbi.nlm.nih.gov/7870860/)]
33. Song F, Khan KS, Dinnes J, Sutton AJ. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002 Feb;31(1):88-95. [doi: [10.1093/ije/31.1.88](https://doi.org/10.1093/ije/31.1.88)] [Medline: [11914301](https://pubmed.ncbi.nlm.nih.gov/11914301/)]
34. Namikawa K, Hirasawa T, Ikenoyama Y, Ishioka M, Tamashiro A, Shiroma S, et al. 343 Can Artificial Intelligence-Based Diagnostic System Perform Differential Diagnosis Of Gastric Cancer And Gastric Ulcer? In: *Gastrointestinal Endoscopy. 2019 Jun Presented at: Digestive Disease Week 2019 American Society for Gastrointestinal Endoscopy Program; May 18-21, 2019; San Diego, CA p. AB74.* [doi: [10.1016/j.gie.2019.04.047](https://doi.org/10.1016/j.gie.2019.04.047)]
35. Cho B, Bang CS, Park SW, Yang YJ, Seo SI, Lim H, et al. Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* 2019 Dec;51(12):1121-1129. [doi: [10.1055/a-0981-6133](https://doi.org/10.1055/a-0981-6133)] [Medline: [31443108](https://pubmed.ncbi.nlm.nih.gov/31443108/)]
36. Hori K, Takemoto S, Sakai Y. Automatic detection of early gastric cancer in endoscopic images using a transferring convolutional neural network. *Conference Abstract. In: United European Gastroenterology Journal. 2018 Presented at: United European Gastroenterology Week 2018; October 21, 2018; Vienna, Austria p. A97.* [doi: [10.1109/embc.2018.8513274](https://doi.org/10.1109/embc.2018.8513274)]
37. Kanesaka T, Lee T, Uedo N, Lin K, Chen H, Lee J, et al. Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging. *Gastrointest Endosc* 2018 May;87(5):1339-1344. [doi: [10.1016/j.gie.2017.11.029](https://doi.org/10.1016/j.gie.2017.11.029)] [Medline: [29225083](https://pubmed.ncbi.nlm.nih.gov/29225083/)]
38. Horiuchi Y, Aoyama K, Tokai Y, Hirasawa T, Yoshimizu S, Ishiyama A, et al. Convolutional Neural Network for Differentiating Gastric Cancer from Gastritis Using Magnified Endoscopy with Narrow Band Imaging. *Dig Dis Sci* 2020 May 04;65(5):1355-1363. [doi: [10.1007/s10620-019-05862-6](https://doi.org/10.1007/s10620-019-05862-6)] [Medline: [31584138](https://pubmed.ncbi.nlm.nih.gov/31584138/)]
39. Wu L, Zhou W, Wan X, Zhang J, Shen L, Hu S, et al. A deep neural network improves endoscopic detection of early gastric cancer without blind spots. *Endoscopy* 2019 Jun 12;51(6):522-531. [doi: [10.1055/a-0855-3532](https://doi.org/10.1055/a-0855-3532)] [Medline: [30861533](https://pubmed.ncbi.nlm.nih.gov/30861533/)]
40. Ikenoyama Y, Hirasawa T, Ishioka M, Namikawa K, Yoshimizu S, Horiuchi Y, et al. Detecting early gastric cancer: Comparison between the diagnostic ability of convolutional neural networks and endoscopists. *Dig Endosc* 2021 Jan 02;33(1):141-150 [FREE Full text] [doi: [10.1111/den.13688](https://doi.org/10.1111/den.13688)] [Medline: [32282110](https://pubmed.ncbi.nlm.nih.gov/32282110/)]
41. Liu D, Gan T, Rao N, Xing Y, Zheng J, Li S, et al. Identification of lesion images from gastrointestinal endoscope based on feature extraction of combinational methods with and without learning process. *Med Image Anal* 2016 Aug;32:281-294. [doi: [10.1016/j.media.2016.04.007](https://doi.org/10.1016/j.media.2016.04.007)] [Medline: [27236223](https://pubmed.ncbi.nlm.nih.gov/27236223/)]
42. Yoon HJ, Kim S, Kim J, Keum J, Oh S, Jo J, et al. A Lesion-Based Convolutional Neural Network Improves Endoscopic Detection and Depth Prediction of Early Gastric Cancer. *J Clin Med* 2019 Aug 26;8(9):1310 [FREE Full text] [doi: [10.3390/jcm8091310](https://doi.org/10.3390/jcm8091310)] [Medline: [31454949](https://pubmed.ncbi.nlm.nih.gov/31454949/)]
43. Kubota K, Kuroda J, Yoshida M, Ohta K, Kitajima M. Medical image analysis: computer-aided diagnosis of gastric cancer invasion on endoscopic images. *Surg Endosc* 2012 May;26(5):1485-1489. [doi: [10.1007/s00464-011-2036-z](https://doi.org/10.1007/s00464-011-2036-z)] [Medline: [22083334](https://pubmed.ncbi.nlm.nih.gov/22083334/)]
44. Miyaki R, Yoshida S, Tanaka S, Kominami Y, Sanomura Y, Matsuo T, et al. Quantitative identification of mucosal gastric cancer under magnifying endoscopy with flexible spectral imaging color enhancement. *J Gastroenterol Hepatol* 2013 May 25;28(5):841-847. [doi: [10.1111/jgh.12149](https://doi.org/10.1111/jgh.12149)] [Medline: [23424994](https://pubmed.ncbi.nlm.nih.gov/23424994/)]
45. Yamakawa R, Harada M, Kawauchi K, Nyuzuki S, Masaya I. UEG Week 2018 Poster Presentations. 2018 Oct Presented at: *United European Gastroenterology 2018; October 21, 2018; Vienna, Austria.* [doi: [10.1177/2050640618792819](https://doi.org/10.1177/2050640618792819)]
46. Lui TKL, Guo C, Leung WK. Accuracy of artificial intelligence on histology prediction and detection of colorectal polyps: a systematic review and meta-analysis. *Gastrointest Endosc* 2020 Jul;92(1):11-22.e6. [doi: [10.1016/j.gie.2020.02.033](https://doi.org/10.1016/j.gie.2020.02.033)] [Medline: [32119938](https://pubmed.ncbi.nlm.nih.gov/32119938/)]

47. Zhang Q, Wang F, Chen Z, Wang Z, Zhi F, Liu S, et al. Comparison of the diagnostic efficacy of white light endoscopy and magnifying endoscopy with narrow band imaging for early gastric cancer: a meta-analysis. *Gastric Cancer* 2016 Apr 29;19(2):543-552. [doi: [10.1007/s10120-015-0500-5](https://doi.org/10.1007/s10120-015-0500-5)] [Medline: [25920526](https://pubmed.ncbi.nlm.nih.gov/25920526/)]
48. Yu H, Wu L. 378 Randomized controlled trial of wisense, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy. *Gastrointestinal Endoscopy* 2019 Jun;89(6):AB74-AB75. [doi: [10.1016/j.gie.2019.04.048](https://doi.org/10.1016/j.gie.2019.04.048)]
49. Horiuchi Y, Hirasawa T, Ishizuka N, Tokai Y, Namikawa K, Yoshimizu S, et al. Performance of a computer-aided diagnosis system in diagnosing early gastric cancer using magnifying endoscopy videos with narrow-band imaging (with videos). *Gastrointest Endosc* 2020 Oct;92(4):856-865.e1. [doi: [10.1016/j.gie.2020.04.079](https://doi.org/10.1016/j.gie.2020.04.079)] [Medline: [32422155](https://pubmed.ncbi.nlm.nih.gov/32422155/)]

Abbreviations

AI: artificial intelligence

AUC: area under the curve

HSROC: hierarchical summary receiver operating characteristic

NBI: narrow-band imaging

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies (revised)

SROC: summary receiver operating characteristic

WLI: white light imaging

Edited by R Kukafka; submitted 02.02.21; peer-reviewed by M Feng, S Pang; comments to author 11.05.21; revised version received 23.10.21; accepted 15.11.21; published 16.05.22

Please cite as:

Chen PC, Lu YR, Kang YN, Chang CC

The Accuracy of Artificial Intelligence in the Endoscopic Diagnosis of Early Gastric Cancer: Pooled Analysis Study

J Med Internet Res 2022;24(5):e27694

URL: <https://www.jmir.org/2022/5/e27694>

doi: [10.2196/27694](https://doi.org/10.2196/27694)

PMID:

©Pei-Chin Chen, Yun-Ru Lu, Yi-No Kang, Chun-Chao Chang. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 16.05.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.