

Review

Computer-Aided Diagnosis of Gastrointestinal Ulcer and Hemorrhage Using Wireless Capsule Endoscopy: Systematic Review and Diagnostic Test Accuracy Meta-analysis

Chang Seok Bang^{1,2,3,4}, MD, PhD; Jae Jun Lee^{3,4,5}, MD; Gwang Ho Baik^{1,2}, MD

¹Department of Internal Medicine, Hallym University College of Medicine, Chuncheon, Republic of Korea

²Institute for Liver and Digestive Diseases, Hallym University, Chuncheon, Republic of Korea

³Institute of New Frontier Research, Hallym University College of Medicine, Chuncheon, Republic of Korea

⁴Division of Big Data and Artificial Intelligence, Chuncheon Sacred Heart Hospital, Chuncheon, Republic of Korea

⁵Department of Anesthesiology and Pain Medicine, Hallym University College of Medicine, Chuncheon, Republic of Korea

Corresponding Author:

Chang Seok Bang, MD, PhD

Department of Internal Medicine

Hallym University College of Medicine

77 Sakju-ro

Chuncheon, 24253

Republic of Korea

Phone: 82 33 240 5821

Fax: 82 33 241 8064

Email: csbang@hallym.ac.kr

Related Article:

This is a corrected version. See correction statement in: <https://www.jmir.org/2022/1/e36170>

Abstract

Background: Interpretation of capsule endoscopy images or movies is operator-dependent and time-consuming. As a result, computer-aided diagnosis (CAD) has been applied to enhance the efficacy and accuracy of the review process. Two previous meta-analyses reported the diagnostic performance of CAD models for gastrointestinal ulcers or hemorrhage in capsule endoscopy. However, insufficient systematic reviews have been conducted, which cannot determine the real diagnostic validity of CAD models.

Objective: To evaluate the diagnostic test accuracy of CAD models for gastrointestinal ulcers or hemorrhage using wireless capsule endoscopic images.

Methods: We conducted core databases searching for studies based on CAD models for the diagnosis of ulcers or hemorrhage using capsule endoscopy and presenting data on diagnostic performance. Systematic review and diagnostic test accuracy meta-analysis were performed.

Results: Overall, 39 studies were included. The pooled area under the curve, sensitivity, specificity, and diagnostic odds ratio of CAD models for the diagnosis of ulcers (or erosions) were .97 (95% confidence interval, .95–.98), .93 (.89–.95), .92 (.89–.94), and 138 (79–243), respectively. The pooled area under the curve, sensitivity, specificity, and diagnostic odds ratio of CAD models for the diagnosis of hemorrhage (or angioectasia) were .99 (.98–.99), .96 (.94–0.97), .97 (.95–.99), and 888 (343–2303), respectively. Subgroup analyses showed robust results. Meta-regression showed that published year, number of training images, and target disease (ulcers vs erosions, hemorrhage vs angioectasia) was found to be the source of heterogeneity. No publication bias was detected.

Conclusions: CAD models showed high performance for the optical diagnosis of gastrointestinal ulcer and hemorrhage in wireless capsule endoscopy.

(*J Med Internet Res* 2021;23(12):e33267) doi: [10.2196/33267](https://doi.org/10.2196/33267)

KEYWORDS

artificial intelligence; computer-aided diagnosis; capsule endoscopy; ulcer; hemorrhage; gastrointestinal; endoscopy; review; accuracy; meta-analysis; diagnostic; performance; machine learning; prediction models

Introduction

Wireless capsule endoscopy (WCE) allows the investigation of gastrointestinal mucosal lesions in a noninvasive manner. This provides approximately 50,000 to 60,000 video frames and allows visualization of the entire gastrointestinal mucosa in a single examination without causing discomfort to patients or risk of procedure-related adverse events [1,2]. Given that the small intestine has been a blind spot for gastroenterologists, WCE has become the standard investigation modality for obscure gastrointestinal hemorrhage and a widely accepted method for the assessment of small intestinal ulcers or tumors [1]. Despite easy accessibility, safety, and patients' comfort for the examination, WCE has a limitation regarding the interpretation. A tedious reading time of approximately 30 to 120 minutes is required, and a small number of abnormal video frames can be easily mistaken for a normal mucosa by endoscopists [1-3].

Artificial intelligence technology has been adopted in gastrointestinal endoscopy, and the automatic detection or diagnosis of abnormal lesions on endoscopic images or movies has been widely investigated [4,5]. The main benefit of the application of artificial intelligence would be the reduction of the laborious reading time and miss rate of important findings in WCE. Another advantage would be the highly accurate diagnostic performance, which is comparable to that of an endoscopist [6]. These artificial intelligence models are expected to aid in the automatic detection of important lesions in WCE images, thus making it possible to perform automatic reading and interpretation of the entire examination.

Previous studies have reported the performance of computer-aided diagnosis (CAD) models using artificial intelligence in WCE [7,8]. Machine learning- or deep

learning-based artificial intelligence models with potential benefits have been reported in these studies. Based on these findings, 2 meta-analyses have been conducted for the pooled diagnostic performance of deep-learning models or convolutional neural network models for the diagnosis of gastrointestinal hemorrhage or ulcers using WCE [7,8]. However, the first meta-analysis searched only 1 database, and a substantial number of important articles were omitted. Moreover, an inaccurate crude number of true positives (TP), false positives (FP), false negatives (FN), or true negatives (TN) of CAD models in each study was reported [7]. This inaccurate pooled diagnostic performance can mislead the readers. The second meta-analysis searched multiple databases; however, it also did not include several important papers, and only a single medical librarian searched all the databases [8]. The main pitfall was the simple pooling of the sensitivity or specificity in each study without considering the distribution of abnormal lesions among the total included lesions in each study. Moreover, the diagnostic performance for the gastrointestinal ulcers and hemorrhage was not separated but combined into a single outcome, and quality assessments in each included study were also omitted. The method of exploring the reason for the heterogeneity and the assessment of publication bias also adhered to the interventional meta-analysis methodology in both meta-analyses but did not satisfy the diagnostic test accuracy (DTA) meta-analysis methodology. Given that the method of conducting interventional and DTA meta-analyses is different and that a widely accepted standard methodology exists in conducting the DTA meta-analysis, this can also mislead the readers (Table 1). Therefore, systematic reviews conducted thus far have been inadequate, and the real diagnostic validity of CAD models in WCE has not yet been determined. This study aimed to evaluate the DTA of CAD models for gastrointestinal ulcers or hemorrhage using WCE images through the standard methodology.

Table 1. Comparison of previous meta-analyses with the current study.

Parameters	This study	Soffer et al [7]	Mohan et al [8]
Number of included studies	20 studies on gastrointestinal ulcers and 19 studies on gastrointestinal hemorrhage	5 studies on gastrointestinal ulcers and 5 studies on gastrointestinal hemorrhage	9 studies for the diagnosis of gastrointestinal ulcers or hemorrhage (did not perform separate analysis between ulcers and hemorrhage)
Main outcome	Separate diagnostic performance of CAD ^a models for the gastrointestinal ulcers or hemorrhage using WCE ^b	Separate diagnostic performance of CAD models for the gastrointestinal ulcers or hemorrhage using WCE	Pooled diagnostic performance of CAD models for gastrointestinal ulcers and hemorrhage using WCE (not a meta-analysis with DTA ^c ; lack of consideration for the prevalence of ulcers or hemorrhage in each study and thus no calculation of TP ^d , FP ^e , FN ^f , or TN ^g in each study)
Search strategy	Search of MEDLINE through PubMed, Web of Science, and the Cochrane Library (2 independent authors searched the databases)	Search of MEDLINE through PubMed (2 independent authors searched the database)	Search of ClinicalTrials.gov, Ovid EBM ^h Reviews, Ovid, Embase, Ovid MEDLINE, Scopus, and Web of Science (a single medical librarian searched all the databases)
Inaccurate calculation (coding) of TP/FP/FN/TN	N/A ⁱ	Inaccurate calculation detected in the study's figures	Not a meta-analysis with DTA; lack of consideration for the prevalence of ulcers or hemorrhage in each study and thus no calculation of TP, FP, FN, or TN in each study
Determination of the heterogeneity between studies	Correlation coefficient between the logarithm of the sensitivity and specificity, beta of HSROC ^j model, visual examination of the SROC curve	I^2 statistics (DTA meta-analysis did not determine heterogeneity with I^2 statistic)	I^2 statistic (DTA meta-analysis did not determine heterogeneity with I^2 statistics)
Quality assessment	QUADAS-2 ^k	QUADAS-2	Not assessed
Publication bias	Deeks funnel plot asymmetry test	Not assessed	Not assessed

^aCAD: computer-aided diagnosis.

^bWCE: wireless capsule endoscopy.

^cDTA: diagnostic test accuracy.

^dTP: true positive.

^eFP: false positive.

^fFN: false negative.

^gTN: true negative.

^hEBM: evidence-based medicine.

ⁱN/A: not applicable.

^jHSROC: hierarchical summary receiver operating characteristic.

^kQUADAS-2: Quality Assessment of Diagnostic Accuracy Studies second version.

Methods

Adherence to the Checklist for Systematic Reviews and Meta-analyses

This study was conducted in accordance with the statement of the PRISMA (Preferred Reporting Items for a Systematic Review and Meta-analysis) of DTA Studies [9]. The study protocol was registered at the International Prospective Register of Systematic Reviews (PROSPERO) database before initiation

of the systematic review (#CRD42021253454). Approval from the institutional review board of the Chuncheon Sacred Heart Hospital was waived.

Search Strategy for Relevant Literature

The authors established searching formulae using keywords related to the performance of CAD models in the detection of ulcer or hemorrhage using WCE images. Medical Subject Headings (MeSH) terminology keywords were used for the establishment of searching formulae (Textbox 1).

Textbox 1. Literature searching strategy for the core databases.

1. CAD of gastrointestinal ulcers in WCE

- Database: MEDLINE (through PubMed)

#1. “artificial intelligence”[tiab] OR “AI”[tiab] OR “deep learning”[tiab] OR “machine learning”[tiab] OR “computer”[tiab] OR “neural network”[tiab] OR “CNN”[tiab] OR “automatic”[tiab] OR “automated”[tiab]: 532189

#2. “capsule endoscopy”[tiab] OR “capsule endoscopy”[Mesh]: 5110

#3. “ulcer”[tiab] OR “ulcer”[Mesh] OR “erosion”[tiab]: 138857

#4. #1 AND #2 AND #3: 29

#5. #4 AND English[Lang]: 29

- Database: Web of Science

#1. artificial intelligence OR AI OR deep learning OR machine learning OR computer OR neural network OR CNN OR automatic OR automated: 1236876

#2. capsule endoscopy: 3524

#3. ulcer: 33664

#4. #1 AND #2 AND #3: 49

- Database: Cochrane Library

#1. artificial intelligence: ab, ti, kw; OR AI: ab, ti, kw; OR deep learning: ab, ti, kw; OR machine learning: ab, ti, kw; OR computer: ab, ti, kw; OR neural network: ab, ti, kw; OR CNN: ab, ti, kw; OR automatic: ab, ti, kw; OR automated: ab, ti, kw: 60327

#2. MeSH descriptor: [capsule endoscopy] explode all trees: 131

#3. capsule endoscopy: ab, ti, kw: 724

#4. #2 OR #3: 724

#5. MeSH descriptor: [ulcer] explode all trees: 1413

#6. ulcer: ab, ti, kw; OR erosion: ab, ti, kw: 20844

#7. #5 or #6: 20844

#8. #1 and #4 and #7: 2 trials

2. CAD of Gastrointestinal hemorrhage in WCE

- Database: MEDLINE (through PubMed)

#1. “artificial intelligence”[tiab] OR “AI”[tiab] OR “deep learning”[tiab] OR “machine learning”[tiab] OR “computer”[tiab] OR “neural network”[tiab] OR “CNN”[tiab] OR “automatic”[tiab] OR “automated”[tiab]: 532189

#2. “capsule endoscopy”[tiab] OR “capsule endoscopy”[Mesh]: 5110

#3. “bleeding”[tiab] OR “hemorrhage”[Mesh] OR “angioectasia”[tiab]: 475519

#4. #1 AND #2 AND #3: 82

#5. #4 AND English[Lang]: 79

- Database: Web of Science

#1. artificial intelligence OR AI OR deep learning OR machine learning OR computer OR neural network OR CNN OR automatic OR automated: 1236876

#2. capsule endoscopy: 3524

#3. bleeding OR hemorrhage OR angioectasia: 146789

#4 #1 AND #2 AND #3: 87

- Database: Cochrane Library

#1. artificial intelligence: ab, ti, kw; OR AI: ab, ti, kw; OR deep learning: ab, ti, kw or machine learning: ab, ti, kw; OR computer: ab, ti, kw; OR neural network: ab, ti, kw; OR CNN: ab, ti, kw; OR automatic: ab, ti, kw; OR automated: ab, ti, kw: 60327

#2. MeSH descriptor: [capsule endoscopy] explode all trees: 131

#3. capsule endoscopy: ab, ti, kw: 724

#4. #2 or #3: 724

#5. MeSH descriptor: [hemorrhage] explode all trees: 14887

#6. bleeding: ab, ti, kw; OR angioectasia: ab, ti, kw: 46708

#7. #5 OR #6: 53831

#8. #1 AND #4 AND #7: 8 (trials)

Abbreviations

CAD, computer-aided diagnosis; WCE: wireless capsule endoscopy; tiab: searching code for title and abstract; Mesh: Medical Subject Headings; ab, ti, kw: search code for abstract, title, and keywords; Lang: search code for language; lim: searching code by limiting certain conditions.

Two authors, CSB and JJJ, independently performed a core database search of MEDLINE through PubMed, Web of Science, and Cochrane Library using pre-established search formulae from inception to May 2021. Duplicate articles were excluded. The titles and abstracts of all identified articles were reviewed, and irrelevant articles were excluded. Full-text reviews were subsequently performed to determine whether the pre-established inclusion criteria were satisfied in the identified literature. The references of relevant articles were also reviewed to identify any additional studies. Any disagreements of results in the searching process between CSB and JJJ were resolved by discussion or consultation with the other author (GHB).

Inclusion Criteria of the Literature

The literature included in this systematic review met the following inclusion criteria: designed to evaluate the diagnostic performance of CAD models for gastrointestinal ulcers or hemorrhage based on WCE images; presentation of the diagnostic performance of CAD models, including sensitivity, specificity, likelihood ratios, predictive values, or accuracy, which enabled the estimation of TP, FP, FN, and TN values of CAD models; and written in English. The exclusion criteria were as follows: narrative review articles, studies with incomplete data, systematic review or meta-analyses, comments, proceedings with only an abstract, or study protocols. A full publication with PDF files of available proceedings was considered to be a full article. Articles meeting at least 1 of the exclusion criteria were excluded from this study.

Assessment of Methodological Quality in the Selected Literature

The methodological quality of the included articles was assessed by CSB and JJJ using the second version of Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2). This tool comprises 4 domains, including “patient selection,” “index test,” “reference standard,” and “flow and timing,” with the first 3 domains possessing an “applicability” assessment. CSB and JJJ assessed each part as having either a high, low, or unclear risk of bias [10].

Data Extraction in the Selected Literature, Primary Outcomes of This Study, and Additional Analyses

CSB and JJJ independently extracted the data from each included article and cross-checked the extracted data. If data were unclear, the corresponding author of the study was contacted by email to obtain insight into the original data set. A descriptive synthesis was made by a systematic review

process, and DTA meta-analysis was conducted if the included studies were sufficiently homogenous.

The primary outcomes were the TP, FP, FN, and TN values in each study. For the CAD of gastrointestinal ulcers or hemorrhage using WCE images, the primary outcomes were defined as follows: TP, the number of patients with a positive finding by a CAD model and who had ulcers or hemorrhage as evidenced by WCE images; FP, the number of patients with a positive finding by a CAD model and who did not have ulcers or hemorrhage based on WCE images; FN, the number of patients with a negative finding by a CAD model and who had ulcers or hemorrhage as evidenced by WCE images; and TN, number of patients with a negative finding by a CAD model and who did not have ulcers or hemorrhage based on the WCE images. With these definitions, TP, FP, FN, and TN values were calculated for each included study.

For additional analysis, such as subgroup analysis or meta-regression, the authors extracted the following variables from each included study: published year, geographic origin of the data (ie, Western vs Asian vs public data or unknown), type of CAD models, type of endoscopic images, number of total images included, type of test data sets (internal test vs external test), and the target disease (ulcer vs erosions, hemorrhage, or angioectasias).

Statistics

The hierarchical summary receiver operating characteristic (HSROC) method was used for the DTA meta-analysis [11]. A forest plot of the sensitivity and specificity and a SROC curve were generated. The level of heterogeneity across the included articles was determined by the correlation coefficient between logit-transformed sensitivity and specificity by the bivariate method [12]; for this, the asymmetry parameter was β , where $\beta=0$ corresponds to a symmetric ROC curve in which the diagnostic odds ratio (DOR) does not vary along the curve according to the HSROC method [11]. A positive correlation coefficient and a β value with a significant probability ($P<.05$) indicate heterogeneity between the studies [11,13]. Visual examination of the SROC curve was also performed to identify heterogeneity. Subgroup analysis by univariate meta-regression using the modifiers identified during the systematic review was also performed to identify the reasons for heterogeneity. STATA software version 15.1 (StataCorp), including the packages “metandi” and “midas,” was used for the DTA meta-analysis. Publication bias was evaluated using the Deeks funnel plot asymmetry test. For the subgroup analyses including less than 4 studies, the Moses-Shapiro-Littenberg method [14] was

adopted using Meta-DiSc 1.4 (XI Cochrane Colloquium) because the “metandi” and “midas” packages require the inclusion of a minimum of 4 studies for DTA meta-analysis.

Results

Study Selection

A total of 254 studies (80 studies for the CAD of gastrointestinal ulcers or erosions and 174 studies for the CAD of gastrointestinal hemorrhage using WCE) were identified following a literature search of 3 databases. Fifteen studies were

additionally identified by manual screening of bibliographies. After excluding duplicate studies, additional articles were excluded after a review of titles and abstracts. Full-text versions of the remaining 54 and 118 articles were obtained and thoroughly reviewed based on the aforementioned inclusion and exclusion criteria in each topic. Among these, 133 articles were excluded from the final enrollment. Finally, 20 studies [15-34] for the CAD of gastrointestinal ulcers or erosions and 19 [17,19,24,34-49] studies for the diagnosis of gastrointestinal hemorrhage were included in the systematic review. A flowchart of the selection process is shown in Figure 1 and 2.

Figure 1. Flowchart of the search process for the diagnostic performance of computer-aided diagnosis for gastrointestinal ulcers or erosions in wireless capsule endoscopy.

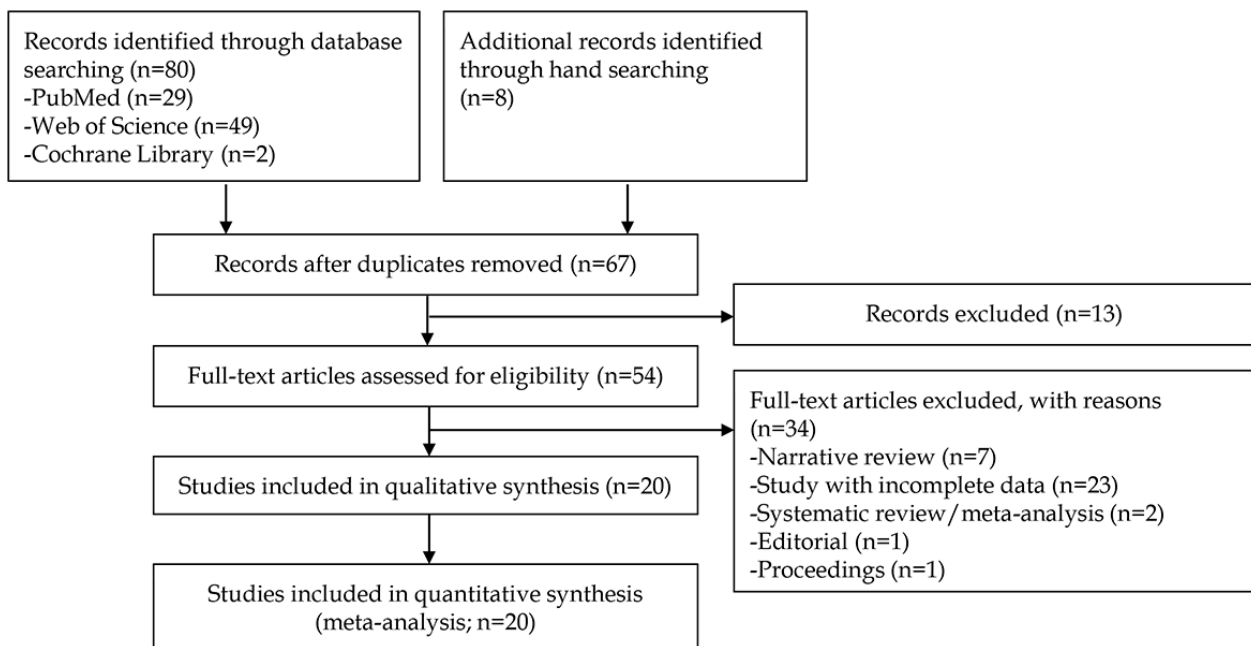
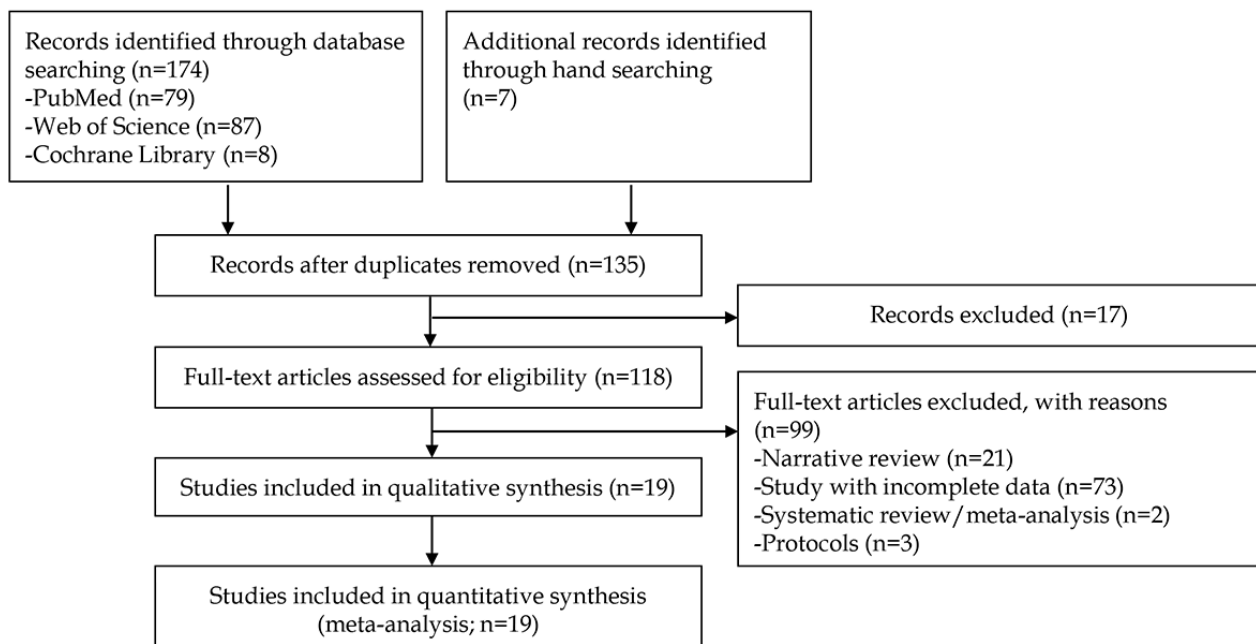


Figure 2. Flowchart of the search process for the diagnostic performance of computer-aided diagnosis for the gastrointestinal hemorrhage in wireless capsule endoscopy.



Clinical Features in the Included Studies

Among the 20 studies [15-34] for the CAD of gastrointestinal ulcers or erosions using WCE, a total of 40,809 images were identified (14,866 cases vs 25,943 controls) for the assessment of the diagnostic performance. Given that the duplicate data were identified (Karargyris et al in 2009 [15] and Karargyris et al in 2011 [20]), all the analyses used the data of 19 studies [15-19,21-34] (a study by Karargyris et al in 2011 [20] was omitted in the meta-analysis).

Ten studies [16-18,21,25-27,29,31,32] used endoscopic images from Asian populations, and 3 studies [22,23,33] used endoscopic images from Western populations. However, 6 studies [15,19,24,28,30,34] used public database images or an unknown source of images. Regarding the type of CAD model, a deep neural network or convolutional neural network was used in 9 studies [16-18,27-29,31-33], and machine learning-based models were used in 10 studies [15,19,21-26,30,34]. Most of the included studies [15-28,30-34] presented the diagnostic performance for the intestinal ulcers. However, the study by Aoki et al [29] presented an indistinguishable performance for the intestinal ulcers or erosions, and the study by Fan et al [27] presented a separate performance for the intestinal ulcers and erosions. Therefore, subgroup analyses were performed for the target lesions. Detailed clinical features of the included studies are presented in [Multimedia Appendix 1](#).

Among the 19 studies [17,19,24,34-49] for the diagnosis of gastrointestinal hemorrhage using WCE, a total of 41,323 images were identified (6952 cases vs 34,371 controls) for the assessment of the diagnostic performance.

Five studies [17], [35], [44], [48], [49] used endoscopic images from Asian populations, and 1 study [47] used endoscopic images from Western populations. However, the remaining 13 studies [19,24,34,36-43,45,46] used public database images or an unknown source of images. Regarding the type of CAD model, the deep neural network or convolutional neural network was used in 8 studies [17,35,39,43,46-49], and machine learning-based models were used in 11 studies [19,24,34,36-38,40-42,44,45]. Most of the included studies [17,19,24,34-46,48] presented the diagnostic performance for intestinal hemorrhage. However, studies by Leenhardt et al [47]

and Tsuboi et al [49] presented the performance for angiodysplasias. Therefore, subgroup analyses were performed for the target lesions. Detailed clinical features of the included studies are presented in [Multimedia Appendix 2](#).

Quality Assessment of Study Methodology

The quality and quantity of baseline training data are important because the CAD models are established using learning features of the baseline training data. A sufficient number of training images are required for the establishment of practical CAD models, and endoscopic specialists should participate in the labeling work for the accurate preparation of training data. We also could not guarantee the quality of images in public databases searched on the internet. We determined that proper learning requires at least 30 training images (quantity standard) from real clinical hospital data (quality standard) labeled by an endoscopic specialist (quality standard). If both quality and quantity standards were satisfied, there was considered to be a low risk of bias in the patient selection domain. If only 1 of these quality or quantity standards was satisfied, there was considered to be an unclear risk of bias. If both were not satisfied, there was considered to be a high risk of bias.

In terms of the CAD of gastrointestinal ulcers or erosions in WCE, only 7 studies [25-27,29,31-34] were rated as low risk of bias, 9 studies [16-18,21-24,30,34] were rated as unclear risk of bias, and 3 studies [15,19,28] were rated as high risk of bias in the “patient selection” domain. The remaining domains were rated as having a low risk of bias in all the included studies ([Figure 3](#)). Therefore, classification of methodological quality in the “patient selection” domain was adopted as a modifier in the subgroup or meta-regression analysis.

In terms of the CAD of gastrointestinal hemorrhage in WCE, only 3 studies [44,47,49] were rated as having a low risk of bias, 10 studies [17,24,34,35,37,38,40,45,46,48] were rated as having an unclear risk of bias, and 6 studies [19,36,39,41-43] were rated as having a high risk of bias in the “patient selection” domain. The remaining domains were rated as having a low risk of bias in all the included studies ([Figure 4](#)). Therefore, the classification of methodological quality in the “patient selection” domain was adopted as a modifier in the subgroup or meta-regression analysis.

Figure 3. Summary graph of quality in methodology for the computer-aided diagnosis of gastrointestinal ulcers or erosions in wireless capsule endoscopy.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
2009 Karargyris A et al.	High	Low	Low	Low	Low	Low	Low
2009 Li B et al. (1)	Unclear	Low	Low	Low	Low	Low	Low
2009 Li B et al. (2)	Unclear	Low	Low	Low	Low	Low	Low
2009 Li B et al. (3)	Unclear	Low	Low	Low	Low	Low	Low
2011 Hwang S	High	Low	Low	Low	Low	Low	Low
2012 Yu L et al.	Unclear	Low	Low	Low	Low	Low	Low
2013 Charisis VS et al.	Unclear	Low	Low	Low	Low	Low	Low
2013 Eid A et al.	Unclear	Low	Low	Low	Low	Low	Low
2014 Yeh JY et al.	Unclear	Low	Low	Low	Low	Low	Low
2015 Yuan Y et al.	Low	Low	Low	Low	Low	Low	Low
2017 Suman S et al.	Low	Low	Low	Low	Low	Low	Low
2018 Fan S et al.	Low	Low	Low	Low	Low	Low	Low
2019 Alaskar H et al.	High	Low	Low	Low	Low	Low	Low
2019 Aoki T et al.	Low	Low	Low	Low	Low	Low	Low
2019 Charfi S et al.	Unclear	Low	Low	Low	Low	Low	Low
2019 Wang S et al. (1)	Low	Low	Low	Low	Low	Low	Low
2019 Wang S et al. (2)	Low	Low	Low	Low	Low	Low	Low
2020 Klang E et al.	Low	Low	Low	Low	Low	Low	Low
2020 Kundu AK et al.	Unclear	Low	Low	Low	Low	Low	Low

● High
 ? Unclear
 + Low

Figure 4. Summary graph of quality in methodology for the computer-aided diagnosis of gastrointestinal hemorrhage in wireless capsule endoscopy.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
2009 Li B et al. (2)	?	+	+	+	+	+	+
2009 Li B et al. (4)	?	+	+	+	+	+	+
2009 Penna B et al.	-	+	+	+	+	+	+
2011 Hwang S	-	+	+	+	+	+	+
2014 Fu Y et al.	?	+	+	+	+	+	+
2014 Ghosh T et al.	?	+	+	+	+	+	+
2014 Sainju S et al.	-	+	+	+	+	+	+
2014 Yeh JY et al.	?	+	+	+	+	+	+
2015 Dilna C et al.	?	+	+	+	+	+	+
2015 Ghosh T et al.	-	+	+	+	+	+	+
2015 Mathew M et al.	-	+	+	+	+	+	+
2016 Jia X et al.	-	+	+	+	+	+	+
2016 Liu DY et al.	+	+	+	+	+	+	+
2016 Yuan Y et al.	?	+	+	+	+	+	+
2017 Jia X et al.	?	+	+	+	+	+	+
2019 Leenhardt R et al.	+	+	+	+	+	+	+
2020 Aoki T et al.	?	+	+	+	+	+	+
2020 Kundu AK et al.	?	+	+	+	+	+	+
2020 Tsuboi A et al.	+	+	+	+	+	+	+

- High
 ? Unclear
 + Low

DTA Meta-analysis for the Performance of CAD Models

Among the 20 studies [15-34] for the meta-analysis of the CAD of gastrointestinal ulcers or erosions using WCE, the area under the curve (AUC), sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, and DOR were 0.97 (95% CI 0.95-0.98), 0.93 (95% CI 0.89-0.95), 0.92 (95% CI 0.89-0.94), 11.2 (95% CI 8.6-14.7), 0.08 (95% CI 0.05-0.12), and 138 (95%

CI 79-243), respectively (Multimedia Appendix 3 and Figure 5). The SROC curve is illustrated in Figure 6. To investigate the clinical utility of the CAD models, Fagan’s nomogram [50] was generated. Positive findings indicated that gastrointestinal ulcers or erosions were detected by the CAD models, while negative findings indicated that gastrointestinal ulcers or erosions were not detected by the CAD models. Assuming a 23% prevalence of gastrointestinal ulcers or erosions, Fagan’s nomogram showed that the posterior probability of ulcers or

erosions was 76% if the finding of the CAD model was positive and that the posterior probability of ulcers or erosions was only 3% if the finding of the CAD model was negative (Figure 7).

Among the 19 studies [17,19,24,34-49] for the meta-analysis of the CAD of gastrointestinal hemorrhage in WCE, the AUC, sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, and DOR were 0.99 (95% CI 0.98-0.99), 0.96 (95% CI 0.94-0.97), 0.97 (95% CI 0.95-0.99), 38.3 (95% CI 19.6-74.8), 0.04 (95% CI 0.03-0.07), and 888 (95% CI 343-2303), respectively (Multimedia Appendix 4 and Figure

8). The SROC curve is illustrated in Figure 9. Positive findings of Fagan’s nomogram indicated that gastrointestinal hemorrhage was detected by the CAD models. Negative findings indicated that gastrointestinal hemorrhage was not detected by the CAD models. Assuming a 10% prevalence of small intestinal hemorrhage in all gastrointestinal bleeding [51], Fagan’s nomogram showed that the posterior probability of small intestinal hemorrhage was 81% if the finding of the CAD model was positive and that the posterior probability of small intestinal hemorrhage was only 0.5% if the finding of the CAD model was negative (Figure 10).

Figure 5. Coupled forest plots of sensitivity and specificity in computer-aided diagnosis models for the diagnosis of gastrointestinal ulcers or erosions in wireless capsule endoscopy images.

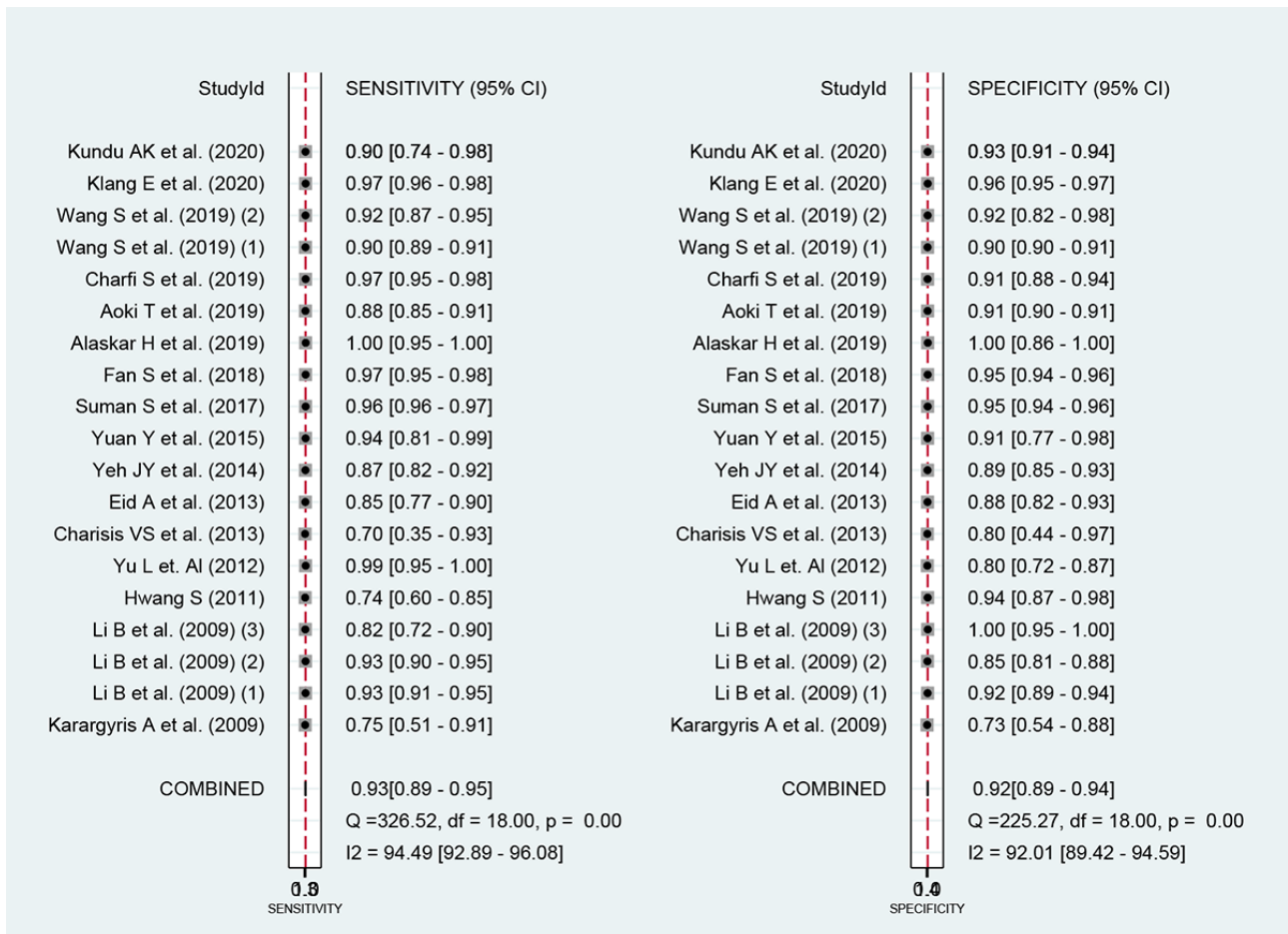


Figure 6. Coupled forest plots of sensitivity and specificity in computer-aided diagnosis models for the diagnosis of gastrointestinal ulcers or erosions in wireless capsule endoscopy images. AUC: area under the curve; SENS: sensitivity; SPEC: specificity; SROC: summary receiver operating characteristic.

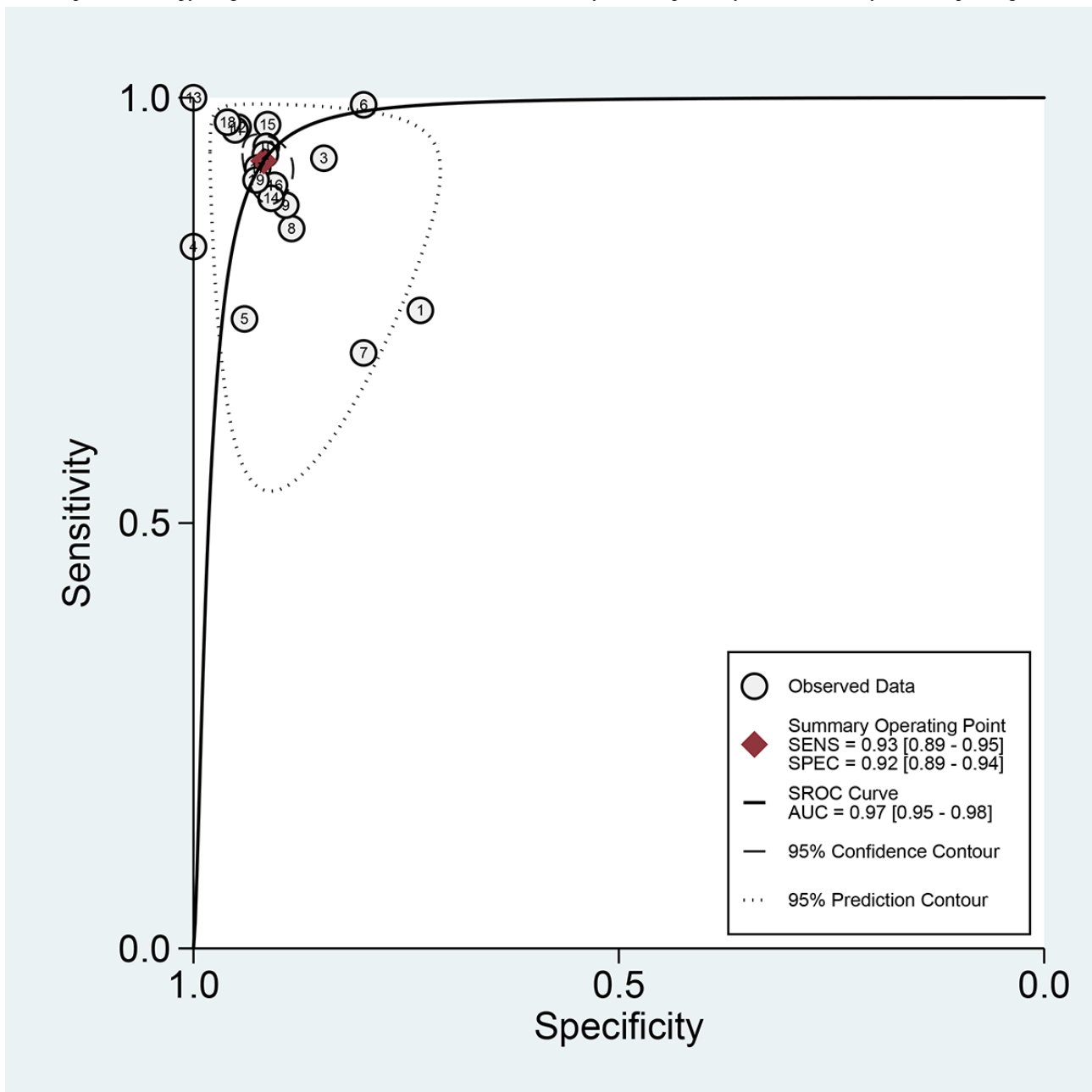


Figure 7. Fagan’s nomogram for the computer-aided diagnosis of gastrointestinal ulcers or erosions in wireless capsule endoscopy images.

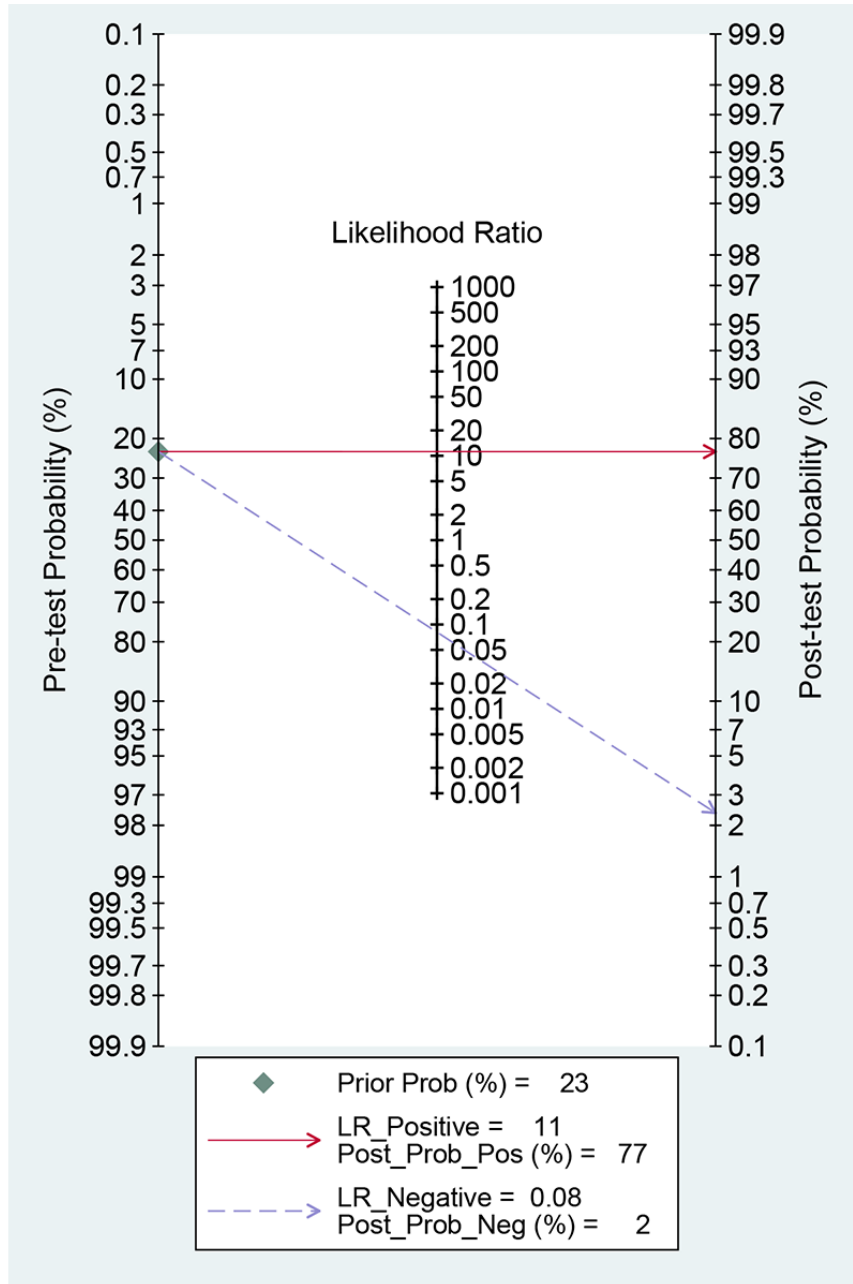


Figure 8. Coupled forest plots of sensitivity and specificity in computer-aided diagnosis models for the diagnosis of gastrointestinal hemorrhage in wireless capsule endoscopy images.

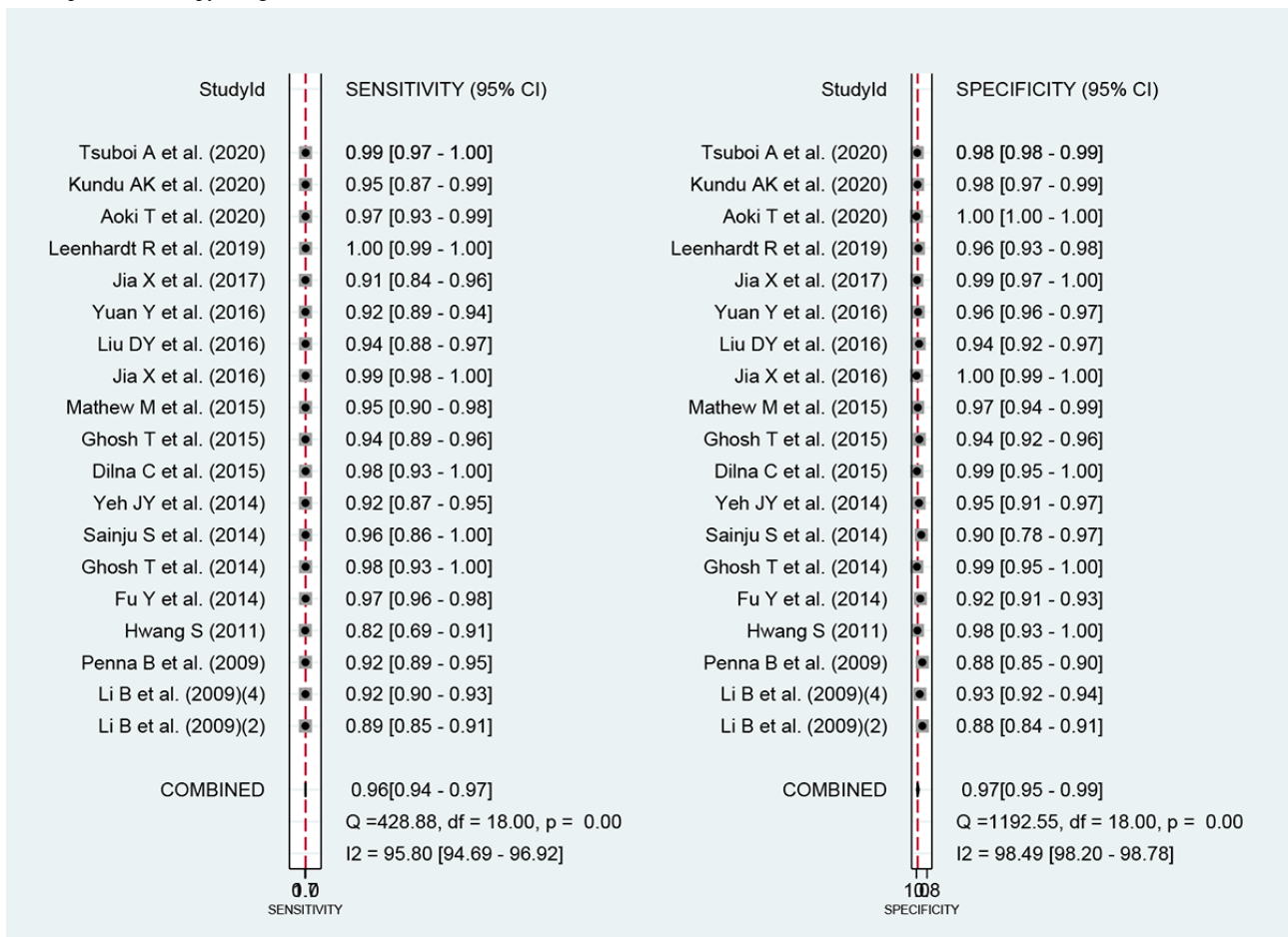


Figure 9. Summary receiver operating characteristic curve with 95% confidence region and prediction region of computer-aided diagnosis models for the diagnosis of gastrointestinal hemorrhage in wireless capsule endoscopy images. AUC: area under the curve; SENS: sensitivity; SPEC: specificity; SROC: summary receiver operating characteristic.

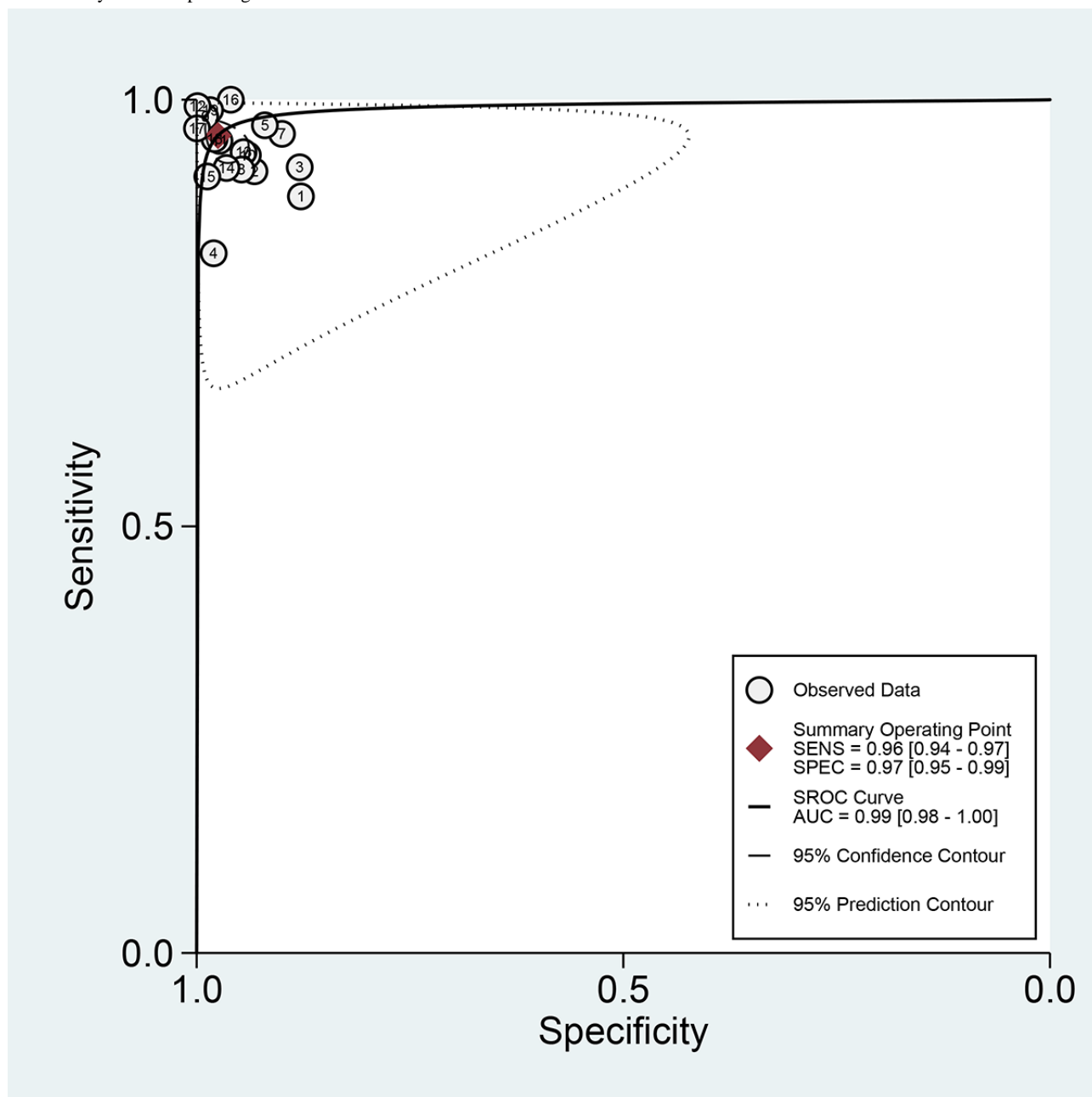
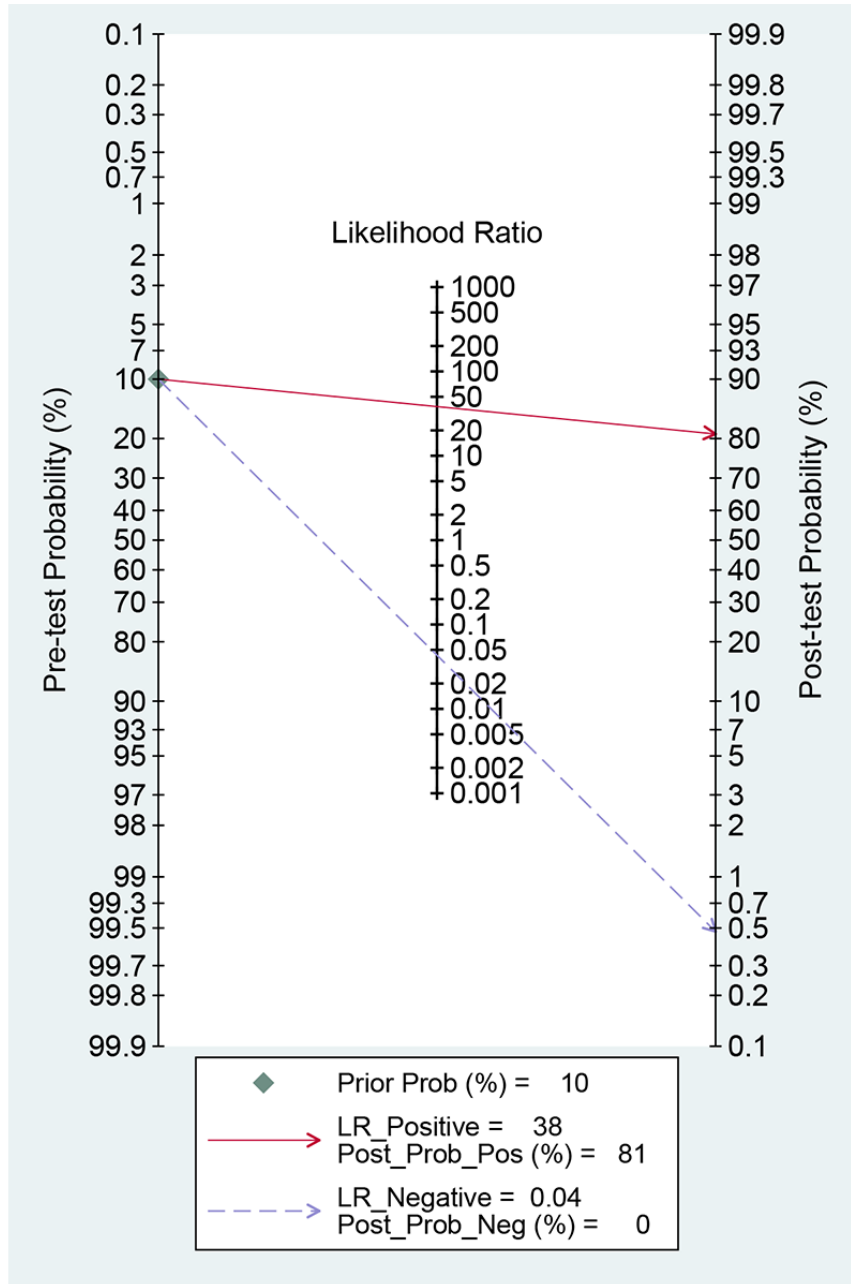


Figure 10. Fagan’s nomogram for the computer-aided diagnosis of small intestinal hemorrhage in wireless capsule endoscopy images.



Assessment of Heterogeneity With Meta-regression and Subgroup Analysis

For the CAD of gastrointestinal ulcers or erosions in WCE, we first observed a positive correlation coefficient between the logit-transformed sensitivity and specificity ($r=0.28$) in the bivariate model analysis. However, an asymmetric β parameter in the HSROC model showed a nonsignificant P value ($P=.15$), implying that heterogeneity was not present among the studies. Second, a coupled forest plot of sensitivity and specificity was observed (Figure 5). Compared with the enrolled studies, the study by Karargyris et al (2009) [15] showed lower sensitivity and specificity. This study was found to have a high risk of bias in the methodology quality assessment (Figure 3). Therefore, subgroup analysis was carried out according to the methodological quality, and the performance was robust although slightly higher values were observed in the studies of

high methodological quality (Multimedia Appendix 3). Third, the shape of the SROC curve for the gastrointestinal ulcers or erosions in WCE was symmetric, and the 95% prediction region was not wide (Figure 6). Fourth, meta-regression using modifiers identified in the systematic review was conducted, and published year, number of training images, and target disease (ulcer vs erosion) were found to be the source of heterogeneity (published year: $P=.04$; number of training images: $P=.02$; target disease ulcer vs erosion: $P=.38$; type of endoscopic image: $P=.01$). Finally, a subgroup analysis based on the potential modifiers was performed, and the overall performance of the studies published within 10 years (vs studies published more than 10 years ago) and studies with more than 100 training images (vs studies with fewer than 100 training images) showed higher values (Multimedia Appendix 3).

For the CAD of gastrointestinal hemorrhage in WCE, we first observed a positive correlation coefficient between the

logit-transformed sensitivity and specificity ($r=0.48$) in the bivariate model analysis. However, an asymmetric β parameter in the HSROC model showed a nonsignificant P value ($P=.06$), implying that heterogeneity was not present among the studies. Second, a coupled forest plot of sensitivity and specificity was observed (Figure 8), and there was no significant outlier. Third, the shape of the SROC curve for the gastrointestinal ulcers and erosions in WCE was symmetric, and the 95% prediction region was not wide (Figure 9). Fourth, a meta-regression using the modifiers identified in the systematic review was conducted, and published year, number of training images, and target disease (hemorrhage vs angioectasia) were found to be the source of heterogeneity (published year: $P<.01$; number of training images: $P=.04$; target disease hemorrhage vs angioectasia: $P<.01$). Finally, a subgroup analysis based on the

potential modifiers was performed, and the overall performance of the studies published within 10 years (vs studies published more than 10 years ago) and studies with more than 100 training images (vs studies with fewer than 100 training images) showed higher values (Multimedia Appendix 4).

Evaluation of Publication bias

The Deeks funnel plot of studies for the gastrointestinal ulcers or erosions in WCE exhibited a symmetrical shape with respect to the regression line (Figure 11), and the asymmetry test showed no evidence of publication bias ($P=.77$). The Deeks funnel plot of studies for the gastrointestinal hemorrhage in WCE exhibited a symmetrical shape with respect to the regression line (Figure 12), and the asymmetry test showed no evidence of publication bias ($P=.93$).

Figure 11. Deeks funnel plot of computer aided diagnosis models for the diagnosis of gastrointestinal ulcers or erosions in wireless capsule endoscopy images.

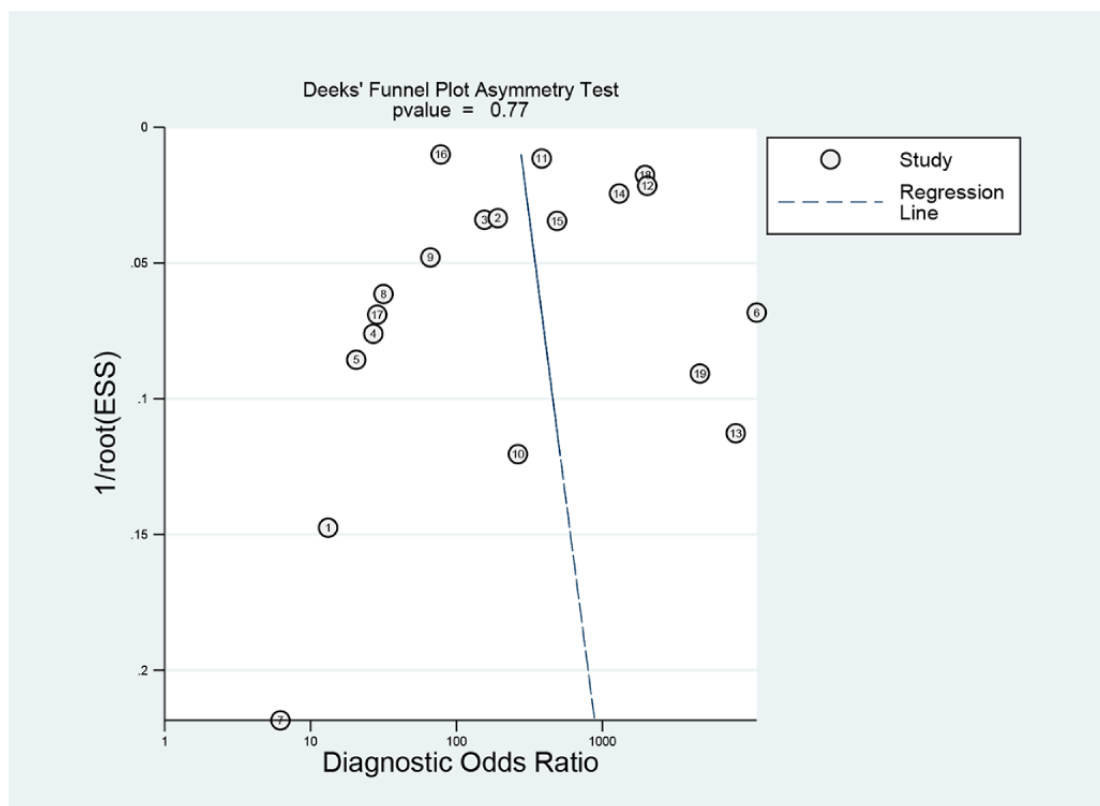
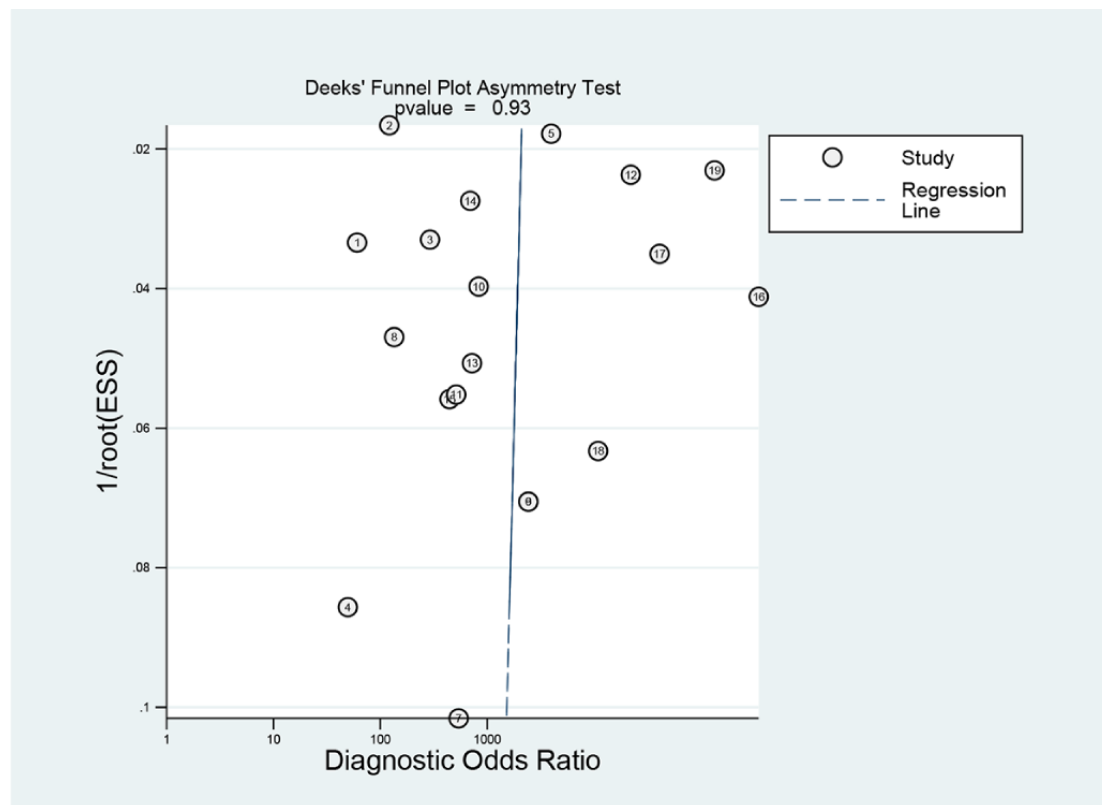


Figure 12. Deeks funnel plot of computer-aided diagnosis models for the diagnosis of gastrointestinal hemorrhage in wireless capsule endoscopy images.



Discussion

Principal Findings

In this study, CAD models showed high performance values for the diagnosis of gastrointestinal ulcers or erosions and hemorrhage in WCE images. Practical values in Fagan's nomogram indicated the potential to use CAD models in clinical practice. Although the main analyses found some heterogeneity among the included studies, the meta-regression showed the common reasons for heterogeneity (published year, number of training images, and target disease—ulcers vs erosions and hemorrhage vs angioectasia), and subgroup analyses demonstrated that recently published studies (vs studies published more than 10 years ago) with a greater amount of training data (vs studies with fewer than 100 training images) showed better performance of CAD models. Thorough subgroup analyses indicated the robust quality of the evidence.

Interpretation of WCE images is an important task for gastroenterologists. Due to the fact that WCE presents the images of the whole gastrointestinal tract, lesions that are difficult to detect with conventional endoscopy can be identified. Diminutive but important culprit lesions also can be found in the WCE examination. The noninvasive nature of this examination and patients' comfort have also promoted the use of this technique in the diagnosis of obscure gastrointestinal hemorrhage or small intestinal disorders. However, the interpretation process is tedious. At least 30 to 120 minutes of reading time is required for the endoscopists [1-3]. It is necessary to maintain concentration throughout the reading time so as not to miss important lesions. CAD models have the

potential to automate the reading process of WCE with their high diagnostic performance, especially for sensitivity and specificity. The overall performance is slightly higher for gastrointestinal hemorrhage than for ulcers or erosions. It is presumed that this is because red-colored blood is easier to distinguish than are white- or yellow-colored ulcers or erosions, which are similar to the color of the background mucosa for the pixel-based or red-green-blue spectrum-based feature learning of CAD.

In the context of the learning way of CADs, neural network-based CAD models showed a slightly higher performance than that of traditional machine learning-based CAD models (Multimedia Appendices 3 and 4). CNN is not always better than machine learning for accurate classification. However, image recognition with local feature extraction can be highly optimized with its complex layers and deep nodes calculations and dimensional reductions for neural network CAD models. Considering that the machine learning-based models in the included studies used color or textures features in the images of WCE, neural network-based models might focus on the other local features or combined features, such as the shape of the lesions or feature differences between the lesions and background mucosa. Explainable artificial intelligence analyses are on the rise, and the application of this technique would provide a method of determination in the CAD models [52].

Although meta-analyses of same topic have already been published, this study was conducted to evaluate the DTA of CAD models for gastrointestinal ulcers or hemorrhage using WCE images with a standard methodology (Table 1) [7,8].

Although previous studies also reported the high performance of CAD models, many important articles were omitted, the heterogeneity between studies was determined by I^2 statistic (which is used in interventional meta-analysis), methodological quality assessments were omitted, and the publication bias was also not assessed.

Limitations

Despite the robust evidence in this meta-analysis, several inevitable limitations were identified. First, all the performance data were only measured in an internal-test setting in each included study. Modeling is an assumption that observations follow certain statistical rules, and external validation is a method to check whether this assumption is correct or generalizable. Therefore, the confirmation of performance in the established CAD models with unused data in the training or internal testing process is essential [53]. However, no single study conducted performance verification in an external validation setting. Second, the definition of intestinal ulcers or erosions was vague. Erosion usually refers to damage that is limited to the mucosa (loss of the epithelium but with the basement membrane or lamina propria being intact). However, the definition of ulcers usually involves more extensive loss of the mucosa beyond the lamina propria. Although the discrimination between these 2 conditions is not perfect under visual inspection, there was no clear definition in the included studies. This can lead to the underestimation or overestimation

of the performance of CAD models. Third, many studies used baseline training data from a public database, and we could not guarantee the quality of images in the public databases available from the internet. The diagnostic performance of the CAD models can only be valid for the population under evaluation and depends on the prevalence of target conditions for the selected population (so-called spectrum bias or class imbalance) [2,54]. This class imbalance was not considered in the included studies. Most of the studies except for 1 [34] applied a 1:1 to 1:4 ratio (target condition:normal mucosa) of the training data set. However, Kundu et al [34] used 31 ulcer images and 1617 normal mucosal images (about a 1:52 ratio) and 65 bleeding images and 1617 normal mucosal images (about a 1:25 ratio) in the training data set. Considering that the method of establishing artificial intelligence models is changing from a model-centric (ie, change or optimize the model to improve performance) to a data-centric approach (ie, systematically change the distribution of the quality of data to improve performance), model establishment that takes into account spectrum bias is required. Overall, qualified training data with clear definitions and a focus on external validation-oriented performance CAD model establishment are required and expected for future perspectives in this topic.

In conclusion, CAD models showed high performance for the optical diagnosis of gastrointestinal ulcers and hemorrhage in WCE.

Acknowledgments

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (grant number 2020-0-01604).

Authors' Contributions

CSB was responsible for conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, writing the original draft, and reviewing and editing the final draft. JJL was responsible for data curation, formal analysis, investigation, and resources. GHB was responsible for data curation, formal analysis, investigation, and resources.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Clinical characteristics of the included studies for the diagnosis of ulcers or erosions in wireless capsule endoscopy images using computer-aided diagnosis.

[\[DOCX File , 23 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Clinical characteristics of the included studies for the diagnosis of gastrointestinal hemorrhage in wireless capsule endoscopy images using computer-aided diagnosis.

[\[DOCX File , 22 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Summary of performance and subgroup analysis of the included studies for the diagnosis of ulcers or erosions in wireless capsule endoscopy images using computer-aided diagnosis.

[\[DOCX File , 21 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Summary of performance and subgroup analysis of the included studies for the diagnosis of bleeding in wireless capsule endoscopy images using computer-aided diagnosis.

[\[DOCX File , 18 KB-Multimedia Appendix 4\]](#)

References

1. ASGE Technology Committee, Wang A, Banerjee S, Barth BA, Bhat YM, Chauhan S, et al. Wireless capsule endoscopy. *Gastrointest Endosc* 2013 Dec;78(6):805-815. [doi: [10.1016/j.gie.2013.06.026](https://doi.org/10.1016/j.gie.2013.06.026)] [Medline: [24119509](https://pubmed.ncbi.nlm.nih.gov/24119509/)]
2. Yang YJ, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019 Apr 14;25(14):1666-1683 [FREE Full text] [doi: [10.3748/wjg.v25.i14.1666](https://doi.org/10.3748/wjg.v25.i14.1666)] [Medline: [31011253](https://pubmed.ncbi.nlm.nih.gov/31011253/)]
3. McAlindon ME, Ching H, Yung D, Sidhu R, Koulaouzidis A. Capsule endoscopy of the small bowel. *Ann Transl Med* 2016 Oct;4(19):369 [FREE Full text] [doi: [10.21037/atm.2016.09.18](https://doi.org/10.21037/atm.2016.09.18)] [Medline: [27826572](https://pubmed.ncbi.nlm.nih.gov/27826572/)]
4. Bang CS. [Deep learning in upper gastrointestinal disorders: status and future perspectives]. *Korean J Gastroenterol* 2020 Mar 25;75(3):120-131 [FREE Full text] [doi: [10.4166/kjg.2020.75.3.120](https://doi.org/10.4166/kjg.2020.75.3.120)] [Medline: [32209800](https://pubmed.ncbi.nlm.nih.gov/32209800/)]
5. Bang CS, Lim H, Jeong HM, Hwang SH. Use of endoscopic images in the prediction of submucosal invasion of gastric neoplasms: automated deep learning model development and usability study. *J Med Internet Res* 2021 Apr 15;23(4):e25167 [FREE Full text] [doi: [10.2196/25167](https://doi.org/10.2196/25167)] [Medline: [33856356](https://pubmed.ncbi.nlm.nih.gov/33856356/)]
6. Bang CS, Lee JJ, Baik GH. Computer-aided diagnosis of esophageal cancer and neoplasms in endoscopic images: a systematic review and meta-analysis of diagnostic test accuracy. *Gastrointest Endosc* 2021 May;93(5):1006-1015.e13 [FREE Full text] [doi: [10.1016/j.gie.2020.11.025](https://doi.org/10.1016/j.gie.2020.11.025)] [Medline: [33290771](https://pubmed.ncbi.nlm.nih.gov/33290771/)]
7. Soffer S, Klang E, Shimon O, Nachmias N, Eliakim R, Ben-Horin S, et al. Deep learning for wireless capsule endoscopy: a systematic review and meta-analysis. *Gastrointest Endosc* 2020 Oct;92(4):831-839.e8. [doi: [10.1016/j.gie.2020.04.039](https://doi.org/10.1016/j.gie.2020.04.039)] [Medline: [32334015](https://pubmed.ncbi.nlm.nih.gov/32334015/)]
8. Mohan BP, Khan SR, Kassab LL, Ponnada S, Chandan S, Ali T, et al. High pooled performance of convolutional neural networks in computer-aided diagnosis of GI ulcers and/or hemorrhage on wireless capsule endoscopy images: a systematic review and meta-analysis. *Gastrointest Endosc* 2021 Feb;93(2):356-364.e4. [doi: [10.1016/j.gie.2020.07.038](https://doi.org/10.1016/j.gie.2020.07.038)] [Medline: [32721487](https://pubmed.ncbi.nlm.nih.gov/32721487/)]
9. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, the PRISMA-DTA Group, et al. Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA* 2018 Jan 23;319(4):388-396. [doi: [10.1001/jama.2017.19163](https://doi.org/10.1001/jama.2017.19163)] [Medline: [29362800](https://pubmed.ncbi.nlm.nih.gov/29362800/)]
10. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011 Oct 18;155(8):529-536. [doi: [10.7326/0003-4819-155-8-201110180-00009](https://doi.org/10.7326/0003-4819-155-8-201110180-00009)] [Medline: [22007046](https://pubmed.ncbi.nlm.nih.gov/22007046/)]
11. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001 Oct 15;20(19):2865-2884. [doi: [10.1002/sim.942](https://doi.org/10.1002/sim.942)] [Medline: [11568945](https://pubmed.ncbi.nlm.nih.gov/11568945/)]
12. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJP, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005 Oct;58(10):982-990. [doi: [10.1016/j.jclinepi.2005.02.022](https://doi.org/10.1016/j.jclinepi.2005.02.022)] [Medline: [16168343](https://pubmed.ncbi.nlm.nih.gov/16168343/)]
13. Harbord RM, Whiting P, Metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. *The Stata Journal* 2009 Aug 01;9(2):211-229. [doi: [10.1177/1536867x0900900203](https://doi.org/10.1177/1536867x0900900203)]
14. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making* 1993;13(4):313-321. [doi: [10.1177/0272989X9301300408](https://doi.org/10.1177/0272989X9301300408)] [Medline: [8246704](https://pubmed.ncbi.nlm.nih.gov/8246704/)]
15. Karagyris A, Bourbakis N. Identification of ulcers in wireless capsule endoscopy videos. 2009 Presented at: International Symposium on Biomedical Imaging: From Nano to Macro; June 28-July 1, 2009; Boston, MA, USA. [doi: [10.1109/isbi.2009.5193107](https://doi.org/10.1109/isbi.2009.5193107)]
16. Li B, Meng MQ. Texture analysis for ulcer detection in capsule endoscopy images. *Image and Vision Computing* 2009 Aug;27(9):1336-1342. [doi: [10.1016/j.imavis.2008.12.003](https://doi.org/10.1016/j.imavis.2008.12.003)]
17. Li B, Meng MQ. Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments. *Comput Biol Med* 2009 Mar;39(2):141-147. [doi: [10.1016/j.compbimed.2008.11.007](https://doi.org/10.1016/j.compbimed.2008.11.007)] [Medline: [19147126](https://pubmed.ncbi.nlm.nih.gov/19147126/)]
18. Li B, Qi L, Meng M, Fan Y. Using ensemble classifier for small bowel ulcer detection in wireless capsule endoscopy images. 2009 Presented at: IEEE International Conference on Robotics and Biomimetics; Dec 19-13, 2009; Guilin, Guangxi, China. [doi: [10.1109/robio.2009.5420455](https://doi.org/10.1109/robio.2009.5420455)]
19. Hwang S. Bag-of-visual-words approach to abnormal image detection in wireless capsule endoscopy videos. 2011 Presented at: International Symposium on Visual Computing: Advances in Visual Computing; Sep 26-28, 2011; Las Vegas, NV, USA. [doi: [10.1007/978-3-642-24031-7_32](https://doi.org/10.1007/978-3-642-24031-7_32)]
20. Karagyris A, Bourbakis N. Detection of small bowel polyps and ulcers in wireless capsule endoscopy videos. *IEEE Trans. Biomed. Eng* 2011 Oct;58(10):2777-2786. [doi: [10.1109/tbme.2011.2155064](https://doi.org/10.1109/tbme.2011.2155064)]

21. Yu L, Yuen PC, Lai J: Ulcer detection in wireless capsule endoscopy images, In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), IEEE, 2012 Nov;45-48. 2012 Presented at: 21st International Conference on Pattern Recognition (ICPR 2012); November 11-15, 2012; Tsukuba, Japan.
22. Charisis V, Katsimerou C, Hadjileontiadis L, Liatsos C, Sergiadis GD: Computer-aided capsule endoscopy images evaluation based on color rotation and texture features: An educational tool to physicians, In Proceedings of the 26th IEEE International Symposium On Computer-Based Medical Systems, IEEE, 2013;203-208. 2013 Jun 20 Presented at: 26th IEEE International Symposium on Computer-Based Medical Systems; June 20-22, 2013; Porto, Portugal. [doi: [10.1109/cbms.2013.6627789](https://doi.org/10.1109/cbms.2013.6627789)]
23. Eid A, Charisis VS, Hadjileontiadis LJ, Sergiadis GD. A curvelet-based lacunarity approach for ulcer detection from wireless capsule endoscopy images, In Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems, IEEE, 2013; 273-278. 2013 Presented at: 26th IEEE International Symposium on Computer-Based Medical Systems; June 20-22, 2013; Porto, Portugal. [doi: [10.1109/cbms.2013.6627801](https://doi.org/10.1109/cbms.2013.6627801)]
24. Yeh J, Wu T, Tsai W. Bleeding and ulcer detection using wireless capsule endoscopy images. JSEA 2014;07(05):422-432. [doi: [10.4236/jsea.2014.75039](https://doi.org/10.4236/jsea.2014.75039)]
25. Yuan Y, Wang J, Li B, Meng MQ. Saliency based ulcer detection for wireless capsule endoscopy diagnosis. IEEE Trans Med Imaging 2015 Oct;34(10):2046-2057. [doi: [10.1109/TMI.2015.2418534](https://doi.org/10.1109/TMI.2015.2418534)] [Medline: [25850085](https://pubmed.ncbi.nlm.nih.gov/25850085/)]
26. Suman S, Hussin F, Malik A, Ho S, Hilmi I, Leow A, et al. Feature selection and classification of ulcerated lesions using statistical analysis for WCE images. Applied Sciences 2017 Oct 24;7(10):1097. [doi: [10.3390/app7101097](https://doi.org/10.3390/app7101097)]
27. Fan S, Xu L, Fan Y, Wei K, Li L. Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images. Phys Med Biol 2018 Aug 10;63(16):165001. [doi: [10.1088/1361-6560/aad51c](https://doi.org/10.1088/1361-6560/aad51c)] [Medline: [30033931](https://pubmed.ncbi.nlm.nih.gov/30033931/)]
28. Alaskar H, Hussain A, Al-Aseem N, Liatsis P, Al-Jumeily D. Application of convolutional neural networks for automated ulcer detection in wireless capsule endoscopy images. Sensors (Basel) 2019 Mar 13;19(6):1265-1265 [FREE Full text] [doi: [10.3390/s19061265](https://doi.org/10.3390/s19061265)] [Medline: [30871162](https://pubmed.ncbi.nlm.nih.gov/30871162/)]
29. Aoki T, Yamada A, Aoyama K, Saito H, Tsuboi A, Nakada A, et al. Automatic detection of erosions and ulcerations in wireless capsule endoscopy images based on a deep convolutional neural network. Gastrointest Endosc 2019 Feb;89(2):357-363.e2. [doi: [10.1016/j.gie.2018.10.027](https://doi.org/10.1016/j.gie.2018.10.027)] [Medline: [30670179](https://pubmed.ncbi.nlm.nih.gov/30670179/)]
30. Charfi S, El Ansari M. Computer-aided diagnosis system for ulcer detection in wireless capsule endoscopy videos, In 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), IEEE, 2017;1-5. 2017 Presented at: 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP); May 22-24, 2017; Fez, Morocco. [doi: [10.1109/atsip.2017.8075590](https://doi.org/10.1109/atsip.2017.8075590)]
31. Wang S, Xing Y, Zhang L, Gao H, Zhang H. A systematic evaluation and optimization of automatic detection of ulcers in wireless capsule endoscopy on a large dataset using deep convolutional neural networks. Phys Med Biol 2019 Dec 05;64(23):235014. [doi: [10.1088/1361-6560/ab5086](https://doi.org/10.1088/1361-6560/ab5086)] [Medline: [31645019](https://pubmed.ncbi.nlm.nih.gov/31645019/)]
32. Wang S, Xing Y, Zhang L, Gao H, Zhang H. Deep convolutional neural network for ulcer recognition in wireless capsule endoscopy: experimental feasibility and optimization. Comput Math Methods Med 2019;2019:7546215 [FREE Full text] [doi: [10.1155/2019/7546215](https://doi.org/10.1155/2019/7546215)] [Medline: [31641370](https://pubmed.ncbi.nlm.nih.gov/31641370/)]
33. Klang E, Barash Y, Margalit RY, Soffer S, Shimon O, Albshesh A, et al. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. Gastrointest Endosc 2020 Mar;91(3):606-613.e2. [doi: [10.1016/j.gie.2019.11.012](https://doi.org/10.1016/j.gie.2019.11.012)] [Medline: [31743689](https://pubmed.ncbi.nlm.nih.gov/31743689/)]
34. Kundu AK, Fattah SA, Wahid KA. Multiple Linear Discriminant Models for Extracting Salient Characteristic Patterns in Capsule Endoscopy Images for Multi-Disease Detection. IEEE J Transl Eng Health Med 2020;8:3300111 [FREE Full text] [doi: [10.1109/JTEHM.2020.2964666](https://doi.org/10.1109/JTEHM.2020.2964666)] [Medline: [32190429](https://pubmed.ncbi.nlm.nih.gov/32190429/)]
35. Li B, Meng MQ. Computer-aided detection of bleeding regions for capsule endoscopy images. IEEE Trans Biomed Eng 2009 May;56(4):1032-1039. [doi: [10.1109/TBME.2008.2010526](https://doi.org/10.1109/TBME.2008.2010526)] [Medline: [19174349](https://pubmed.ncbi.nlm.nih.gov/19174349/)]
36. Penna B, Tillo T, Grangetto M, Magli E, Olmo G. A technique for blood detection in wireless capsule endoscopy images. 2009 Presented at: 17th European Signal Processing Conference; Aug 24-29, 2009; Glasgow, Scotland.
37. Fu Y, Zhang W, Mandal M, Meng MQ. Computer-aided bleeding detection in WCE video. IEEE J Biomed Health Inform 2014 Mar;18(2):636-642. [doi: [10.1109/JBHI.2013.2257819](https://doi.org/10.1109/JBHI.2013.2257819)] [Medline: [24608063](https://pubmed.ncbi.nlm.nih.gov/24608063/)]
38. Ghosh T, Bashar S, Alam M, Wahid K, Fattah SA. A statistical feature based novel method to detect bleeding in wireless capsule endoscopy images, IEEE, 2014;1-4. 2014 Presented at: The 17th International Conference on Informatics, Electronics & Vision; May 23-24, 2014; Dhaka, Bangladesh. [doi: [10.1109/iciev.2014.6850777](https://doi.org/10.1109/iciev.2014.6850777)]
39. Sainju S, Bui FM, Wahid KA. Automated bleeding detection in capsule endoscopy videos using statistical features and region growing. J Med Syst 2014 May;38(4):25. [doi: [10.1007/s10916-014-0025-1](https://doi.org/10.1007/s10916-014-0025-1)] [Medline: [24696394](https://pubmed.ncbi.nlm.nih.gov/24696394/)]
40. Dilna C, Gopi V. A novel method for bleeding detection in Wireless Capsule Endoscopic images. 2015 Presented at: International Conference on Computing and Network Communications; Dec 16-19, 2015; Trivandrum, India. [doi: [10.1109/coconet.2015.7411289](https://doi.org/10.1109/coconet.2015.7411289)]
41. Ghosh T, Fattah S, Bashar S, Shahnaz C, Wahid K, Zhu WP. An automatic bleeding detection technique in wireless capsule endoscopy from region of interest, In. 2015 Jul 21 Presented at: 2015 IEEE International Conference on Digital Signal Processing; July 21-24, 2015; Singapore. [doi: [10.1109/icdsp.2015.7252090](https://doi.org/10.1109/icdsp.2015.7252090)]

42. Mathew M, Gopi VP. Transform based bleeding detection technique for endoscopic images. 2015 Feb 26 Presented at: The 2nd International Conference on Electronics and Communication Systems (ICECS); Feb 26-27, 2015; Coimbatore, India. [doi: [10.1109/ecs.2015.7124882](https://doi.org/10.1109/ecs.2015.7124882)]
43. Xiao Jia, Meng MQ. A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images. *Annu Int Conf IEEE Eng Med Biol Soc* 2016 Aug;2016:639-642. [doi: [10.1109/EMBC.2016.7590783](https://doi.org/10.1109/EMBC.2016.7590783)] [Medline: [28268409](https://pubmed.ncbi.nlm.nih.gov/28268409/)]
44. 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/)]
45. Yuan Y, Li B, Meng MQ. Bleeding Frame and Region Detection in the Wireless Capsule Endoscopy Video. *IEEE J Biomed Health Inform* 2016 Mar;20(2):624-630. [doi: [10.1109/JBHI.2015.2399502](https://doi.org/10.1109/JBHI.2015.2399502)] [Medline: [25675468](https://pubmed.ncbi.nlm.nih.gov/25675468/)]
46. Xiao Jia, Meng MQ. Gastrointestinal bleeding detection in wireless capsule endoscopy images using handcrafted and CNN features. *Annu Int Conf IEEE Eng Med Biol Soc* 2017 Jul;2017:3154-3157. [doi: [10.1109/EMBC.2017.8037526](https://doi.org/10.1109/EMBC.2017.8037526)] [Medline: [29060567](https://pubmed.ncbi.nlm.nih.gov/29060567/)]
47. Leenhardt R, Vasseur P, Li C, Saurin JC, Rahmi G, Cholet F, CAD-CAP Database Working Group. A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. *Gastrointest Endosc* 2019 Jan;89(1):189-194. [doi: [10.1016/j.gie.2018.06.036](https://doi.org/10.1016/j.gie.2018.06.036)] [Medline: [30017868](https://pubmed.ncbi.nlm.nih.gov/30017868/)]
48. Aoki T, Yamada A, Kato Y, Saito H, Tsuboi A, Nakada A, et al. Automatic detection of blood content in capsule endoscopy images based on a deep convolutional neural network. *J Gastroenterol Hepatol* 2020 Jul;35(7):1196-1200. [doi: [10.1111/jgh.14941](https://doi.org/10.1111/jgh.14941)] [Medline: [31758717](https://pubmed.ncbi.nlm.nih.gov/31758717/)]
49. Tsuboi A, Oka S, Aoyama K, Saito H, Aoki T, Yamada A, et al. Artificial intelligence using a convolutional neural network for automatic detection of small-bowel angioectasia in capsule endoscopy images. *Dig Endosc* 2020 Mar;32(3):382-390. [doi: [10.1111/den.13507](https://doi.org/10.1111/den.13507)] [Medline: [31392767](https://pubmed.ncbi.nlm.nih.gov/31392767/)]
50. Vuik FER, Nieuwenburg SAV, Moen S, Schreuders EH, Oudkerk Pool MD, Peterse EFP, et al. Population-Based Prevalence of Gastrointestinal Abnormalities at Colon Capsule Endoscopy. *Clin Gastroenterol Hepatol* 2020 Oct 31:S1542-S3565. [doi: [10.1016/j.cgh.2020.10.048](https://doi.org/10.1016/j.cgh.2020.10.048)] [Medline: [33130189](https://pubmed.ncbi.nlm.nih.gov/33130189/)]
51. Murphy B, Winter DC, Kavanagh DO. Small Bowel Gastrointestinal Bleeding Diagnosis and Management-A Narrative Review. *Front Surg* 2019;6:25 [FREE Full text] [doi: [10.3389/fsurg.2019.00025](https://doi.org/10.3389/fsurg.2019.00025)] [Medline: [31157232](https://pubmed.ncbi.nlm.nih.gov/31157232/)]
52. Bang CS, Ahn JY, Kim J, Kim Y, Choi IJ, Shin WG. Establishing Machine Learning Models to Predict Curative Resection in Early Gastric Cancer with Undifferentiated Histology: Development and Usability Study. *J Med Internet Res* 2021 May 15;23(4):e25053 [FREE Full text] [doi: [10.2196/25053](https://doi.org/10.2196/25053)] [Medline: [33856358](https://pubmed.ncbi.nlm.nih.gov/33856358/)]
53. Bang CS, Lee JJ, Baik GH. Computer-aided diagnosis of diminutive colorectal polyps in endoscopic images: systematic review and meta-analysis of diagnostic test accuracy. *J Med Internet Res* 2021 Aug 25;23(8):e29682 [FREE Full text] [doi: [10.2196/29682](https://doi.org/10.2196/29682)] [Medline: [34432643](https://pubmed.ncbi.nlm.nih.gov/34432643/)]
54. 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/)]

Abbreviations

- AUC:** area under the curve
- CAD:** computer-aided diagnosis
- DOR:** diagnostic odds ratio
- DTA:** diagnostic test accuracy
- FP:** false positive
- FN:** false negative
- HSROC:** hierarchical summary receiver operating characteristic
- MeSH:** Medical Subject Headings
- PRISMA:** Preferred Reporting Items for a Systematic Review and Meta-analysis
- PROSPERO:** International Prospective Register of Systematic Reviews
- QUADAS-2:** the second version of Quality Assessment of Diagnostic Accuracy Studies
- TP:** true positive
- TN:** true negative
- WCE:** wireless capsule endoscopy

Edited by G Eysenbach; submitted 30.08.21; peer-reviewed by SI Seo; comments to author 09.10.21; revised version received 10.10.21; accepted 13.10.21; published 14.12.21

Please cite as:

Bang CS, Lee JJ, Baik GH

Computer-Aided Diagnosis of Gastrointestinal Ulcer and Hemorrhage Using Wireless Capsule Endoscopy: Systematic Review and Diagnostic Test Accuracy Meta-analysis

J Med Internet Res 2021;23(12):e33267

URL: <https://www.jmir.org/2021/12/e33267>

doi: [10.2196/33267](https://doi.org/10.2196/33267)

PMID: [34904949](https://pubmed.ncbi.nlm.nih.gov/34904949/)

©Chang Seok Bang, Jae Jun Lee, Gwang Ho Baik. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 14.12.2021. 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.