

Observational Study

Comparative study on artificial intelligence systems for detecting early esophageal squamous cell carcinoma between narrow-band and white-light imaging

Bing Li, Shi-Lun Cai, Wei-Min Tan, Ji-Chun Li, Ayimukedisi Yalikong, Xiao-Shuang Feng, Hon-Ho Yu, Pin-Xiang Lu, Zhen Feng, Li-Qing Yao, Ping-Hong Zhou, Bo Yan, Yun-Shi Zhong

ORCID number: Bing Li 0000-0001-9802-5860; Shi-Lun Cai 0000-0002-5000-9658; Wei-Min Tan 0000-0001-7677-4772; Ji-Chun Li 0000-0003-4906-8244; Ayimukedisi Yalikong 0000-0002-7328-8354; Xiao-Shuang Feng 0000-0002-7933-3996; Hon-Ho Yu 0000-0002-9580-345X; Pin-Xiang Lu 0000-0001-5941-9584; Zhen Feng 0000-0003-0424-4726; Li-Qing Yao 0000-0001-6900-6791; Ping-Hong Zhou 0000-0002-5434-0540; Bo Yan 0000-0003-0256-9682; Yun-Shi Zhong 0000-0002-3128-3168.

Author contributions: Zhong YS and Yan B conceived the study design; Li B, Cai SL, Tan WM, Li JC, Yalikong A, Yu HH, Lu PX, and Feng Z acquired the data; Li B, Cai SL, Yalikong A, Tan WM, Feng XS, Yao LQ, Zhou PH, and Zhong YS analyzed and interpreted the data; Li B, Tan WM, and Cai SL drafted the manuscript; Zhong YS critically revised the manuscript for important intellectual content; Li B, Cai SL, Tan WM, Li JC, Yalikong A, Feng XS, Yu HH, Lu PX, Feng Z, Yao LQ, Zhou PH, Yan B, and Zhong YS approved the final version of the manuscript; Li B, Cai SL and Tan WM contributed equally to this article.

Supported by National Key R&D

Bing Li, Shi-Lun Cai, Ayimukedisi Yalikong, Li-Qing Yao, Ping-Hong Zhou, Yun-Shi Zhong, Department of Endoscopy Center, Zhongshan Hospital of Fudan University, Shanghai 200032, China

Wei-Min Tan, Ji-Chun Li, Bo Yan, School of Computer Science, Fudan University, Shanghai 200433, China

Xiao-Shuang Feng, Clinical Statistical Center, Shanghai Cancer Center of Fudan University, Shanghai 200032, China

Hon-Ho Yu, Department of Gastroenterology, Kiang Wu Hospital, Macau SAR 999078, China

Pin-Xiang Lu, Zhen Feng, Department of Endoscopy Center, Xuhui Hospital, Zhongshan Hospital of Fudan University, Shanghai 200031, China

Corresponding author: Yun-Shi Zhong, MD, PhD, Professor, Department of Endoscopy Center, Zhongshan Hospital of Fudan University, No. 180 Fenglin Road, Shanghai 200032, China. zhongyunshi@yahoo.com

Abstract

BACKGROUND

Non-magnifying endoscopy with narrow-band imaging (NM-NBI) has been frequently used in routine screening of esophagus squamous cell carcinoma (ESCC). The performance of NBI for screening of early ESCC is, however, significantly affected by operator experience. Artificial intelligence may be a unique approach to compensate for the lack of operator experience.

AIM

To construct a computer-aided detection (CAD) system for application in NM-NBI to identify early ESCC and to compare it with our previously reported CAD system with endoscopic white-light imaging (WLI).

METHODS

A total of 2167 abnormal NM-NBI images of early ESCC and 2568 normal images were collected from three institutions (Zhongshan Hospital of Fudan University,

Program of China, No. 2018YFC1315000, No. 2018YFC1315005, No. 2019YFC1315800, and No. 2019YFC1315802; National Natural Science Foundation of China, No. 81861168036 and No. 81702305; Science and Technology Commission Foundation of Shanghai Municipality, No. 19411951600, and No. 19411951601; Macao SAR Science and Technology Development Foundation, No. 0023/2018/AFJ; and Dawn Program of Shanghai Education Commission, No. 18SG08.

Institutional review board

statement: This study was reviewed and approved by the Ethics Committee of Zhongshan Hospital, Fudan University.

Informed consent statement:

Patients were not required to give informed consent to the study because the analysis used anonymous clinical data that were obtained after each patient agreed to undergo treatment by written consent.

Conflict-of-interest statement: The authors declare that they have no competing interests.

Data sharing statement: The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

STROBE statement: The authors have read the STROBE Statement checklist of items, and the manuscript was prepared and revised according to the STROBE Statement-checklist of items.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works

Xuhui Hospital, and Kiang Wu Hospital) as the training dataset, and 316 pairs of images, each pair including images obtained by WLI and NBI (same part), were collected for validation. Twenty endoscopists participated in this study to review the validation images with or without the assistance of the CAD systems. The diagnostic results of the two CAD systems and improvement in diagnostic efficacy of endoscopists were compared in terms of sensitivity, specificity, accuracy, positive predictive value, and negative predictive value.

RESULTS

The area under receiver operating characteristic curve for CAD-NBI was 0.9761. For the validation dataset, the sensitivity, specificity, accuracy, positive predictive value, and negative predictive value of CAD-NBI were 91.0%, 96.7%, 94.3%, 95.3%, and 93.6%, respectively, while those of CAD-WLI were 98.5%, 83.1%, 89.5%, 80.8%, and 98.7%, respectively. CAD-NBI showed superior accuracy and specificity than CAD-WLI ($P = 0.028$ and $P \leq 0.001$, respectively), while CAD-WLI had higher sensitivity than CAD-NBI ($P = 0.006$). By using both CAD-WLI and CAD-NBI, the endoscopists could improve their diagnostic efficacy to the highest level, with accuracy, sensitivity, and specificity of 94.9%, 92.4%, and 96.7%, respectively.

CONCLUSION

The CAD-NBI system for screening early ESCC has higher accuracy and specificity than CAD-WLI. Endoscopists can achieve the best diagnostic efficacy using both CAD-WLI and CAD-NBI.

Key Words: Computer-aided detection; Esophageal squamous cell carcinoma; Endoscopy; Screening; Narrow-band imaging; White-light imaging

©The Author(s) 2021. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: The computer-assisted diagnosis (CAD) system under conventional endoscopic white-light imaging (WLI) for screening of early esophagus squamous cell carcinoma (ESCC) has high accuracy. However, few studies have examined different characteristics of CAD application in WLI and narrow-band imaging (NBI) models. In this study, the CAD system we constructed under the NBI model for screening of early ESCC had higher accuracy and specificity than the CAD-WLI system. Endoscopists could achieve the best diagnostic efficacy by using both CAD-WLI and CAD-NBI. The two CAD systems have different advantages in avoiding missed diagnosis and excessive biopsy, which could help endoscopists, especially those with less experience, in more efficient screening of early ESCC.

Citation: Li B, Cai SL, Tan WM, Li JC, Yalikong A, Feng XS, Yu HH, Lu PX, Feng Z, Yao LQ, Zhou PH, Yan B, Zhong YS. Comparative study on artificial intelligence systems for detecting early esophageal squamous cell carcinoma between narrow-band and white-light imaging. *World J Gastroenterol* 2021; 27(3): 281-293

URL: <https://www.wjgnet.com/1007-9327/full/v27/i3/281.htm>

DOI: <https://dx.doi.org/10.3748/wjg.v27.i3.281>

INTRODUCTION

Upper gastrointestinal endoscopy combined with biopsy is the method of choice to diagnose esophagus squamous cell carcinoma (ESCC), and it has been widely adopted in population screening for ESCC^[1,2]. However, it is not always easy to identify early-stage ESCC, especially for unexperienced doctors, during examination with white-light imaging (WLI)^[3,4]. The narrow-band imaging (NBI) system improves the visualization of microvasculature and mucosal patterns in the alimentary tract^[5]. Non-magnifying endoscopy with NBI (NM-NBI) has been used frequently in routine screening examinations with higher accuracy and specificity^[6,7]. However, the sensitivity of NBI for screening of mucosal high-grade neoplasia is significantly

on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

Manuscript source: Unsolicited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: China

Peer-review report's scientific quality classification

Grade A (Excellent): A

Grade B (Very good): 0

Grade C (Good): C

Grade D (Fair): 0

Grade E (Poor): 0

Received: October 9, 2020

Peer-review started: October 9, 2020

First decision: November 23, 2020

Revised: December 5, 2020

Accepted: December 22, 2020

Article in press: December 22, 2020

Published online: January 21, 2021

P-Reviewer: Kishida Y, Zavras N

S-Editor: Fan JR

L-Editor: Filipodia

P-Editor: Liu JH



different between experienced and less experienced endoscopists^[8,9]. Early lesions missed at screening may not be identified until they become more advanced and less amenable to treatment. Thus, the experience of the operator plays a critical role in the screening result for early ESCC.

Artificial intelligence (AI) may be uniquely poised to compensate for the lack of operator experience. Studies have demonstrated the ability of AI to meet or exceed the performance of human experts as a triage or screening tool for gastrointestinal diseases^[10,11]. In our previous research, we reported a novel system of computer-aided detection (CAD) to localize and identify early ESCC under conventional endoscopic WLI with sensitivity above 97%^[12]. Here, another system of CAD for application in NM-NBI for screening of early ESCC was constructed and validated. More importantly, we compared the effectiveness of the two systems based on WLI or NM-NBI in helping endoscopists to detect early ESCC. On the basis of the results, we can determine which technique is most effective to help endoscopists, *i.e.* whether to use CAD-NBI or CAD-WLI alone or both.

MATERIALS AND METHODS

Study design

This study was performed at the Endoscopy Center of three general hospitals (Zhongshan Hospital of Fudan University, Xuhui Hospital, and Kiang Wu Hospital) in partnership with the School of Computer Science of Fudan University. Patient data were anonymized, and any personal identifying information was excluded. This study was approved by the Institutional Review Board of Zhongshan Hospital, Fudan University (approval No. B2019-141R). All authors had access to the study data and reviewed and approved the final manuscript.

Datasets used for training and validation of the CAD-NBI system

First, we retrospectively obtained esophagoscopy NM-NBI images for the development of the CAD system for NBI images (CAD-NBI system). In the training dataset, a total of 2167 abnormal NM-NBI images of early ESCCs in 235 cases and 2568 normal NM-NBI images in 412 cases were collected between January 2016 and April 2018 from three institutions. Then, we collected 316 pairs of images (133 abnormal and 183 normal), each pair including WLI and NBI at the same location and at the same angle, from 112 consecutive cases. The purpose of establishing this paired image dataset includes: (1) All NBI images are used to test our newly established CAD-NBI system; (2) White light images paired with NBI in the same situation are used to test the CAD system for WLI (CAD-WLI system)^[12] that has been reported previously to compare the differences between the two CAD systems; and (3) Endoscopists are asked to review all the images from this validation dataset to evaluate their diagnostic ability with or without the help of these two CAD systems.

The criteria for normal and abnormal images refer to the previous study on CAD for screening of early ESCC from our team^[12]. The criteria for choosing normal image data were as follows: (1) The initial endoscopic inspection results for the esophagus were negative; (2) The abovementioned patients had no newly detected lesions until September 2019; and (3) The normal images were confirmed by endoscopists with ≥ 15 years of experience in endoscopy, and all the endoscopists believed that the image is normal. All patients with abnormal endoscopic images underwent endoscopic submucosal dissection procedures, and three gastrointestinal pathologists (two with > 10 years of experience and one with > 15 years of experience) conducted histological assessments in the pathology departments of both centers. Early ESCC includes low-grade and high-grade intraepithelial neoplasia and esophageal cancer (EC) that has invaded mucosal or superficial submucosal layer.

Design and development of the CAD-NBI system

In this study, we consider the esophagus lesions in endoscopic images to be semantic objects. We demonstrate the development and validation of an endoscopist-level CAD-NBI system based on deep learning algorithm for screening esophagus lesions. We propose to use fully convolutional neural network based on Visual Geometry Group model for semantic segmentation, where semantics denote the esophagus lesions. Therefore, the CAD-NBI system is used to predict the location and irregular shape of esophagus lesions, which is helpful for endoscopists to judge the size, area, and location of lesions more effectively.

To obtain an accurate predictor based on limited esophagoscopy images, some

preprocessing is conducted on the esophagoscopic images before training. First, some irrelevant regions for lesion detection, such as black background, are cropped automatically using a simple image processing algorithm developed by us. Second, we randomly flip the esophagoscopic images horizontally and vertically to augment data diversity. Third, for further data augmentation, esophagoscopic images and the corresponding lesion masks are resized to 300×300 and randomly cropped to 224×224 . During training, the network parameters are updated with the initial learning rate of 0.0001 and decayed every 2000 iterations with a decay rate of 0.9 in the staircase mode. During inference, given an esophagoscopic image that the CAD-NBI system has never seen previously, the system outputs the segmentation result of esophagoscopic lesion directly.

Comparison of the improvement of diagnostic capability of endoscopists under CAD-WLI and CAD-NBI

The accuracy of the CAD-NBI system was evaluated by the validation dataset established previously. We invited 20 endoscopists with varying experience from three centers to participate in this study in order to compare diagnostic performance between the CAD system and endoscopists. Moreover, we wanted to know the effectiveness of the two systems, CAD-WLI and CAD-NBI, for the improvement of diagnostic capabilities of endoscopists. Among the endoscopists, four were classified as highly experienced endoscopists who had performed more than 10000 conventional endoscopy examinations and 5000 NBI endoscopy examinations, eight were classified as mid-level endoscopists who had performed more than 5000 conventional endoscopy examinations and 2500 NBI endoscopy examinations, and eight were classified as junior endoscopists who had performed more than 2000 conventional endoscopy examinations and 1000 NBI endoscopy examinations.

To test the effectiveness of the two CAD systems, we designed a four-phase trial. In the first phase, the 20 endoscopists were asked to review every pair image of the validation dataset in digital format on a laptop. All of them were blinded to the histological data and asked to review the esophagoscopic images independently. The CAD-NBI system scanned and analyzed each NM-NBI image of every pair saved in JPEG/PNG format on the hard drive, and the CAD-WLI system did the same for NM-WLI images. In the second phase, after the WLI images had been marked using the CAD-WLI system (NBI images were not marked), we invited these endoscopists to again screen every pair of images in the validation dataset. This action was noteworthy as the review sequence of images was altered randomly by computer to minimize the impact of impression from the last performance in each phase. In the third phase, endoscopists were asked to review every pair of images in the validation dataset once again, after all NBI images were marked by the CAD-NBI system while WLI images were not. In the last phase, the endoscopists completed their final diagnosis for the validation paired images by referring to the results from both CAD-WLI and CAD-NBI systems. Between each two continuous phases, we set the wash-out phase as 1-1-2 mo, respectively. Moreover, these 20 endoscopists were unaware of the performance of the CAD-WLI and CAD-NBI systems. A flowchart depicting the processes used during the study is shown in [Figure 1](#).

Outcome measures

The ability of the CAD-NBI system to identify early ESCC was mathematically assessed by the area under the curve of the receiver operating characteristic curve, and the sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) were determined. The accuracy, sensitivity, specificity, PPV, and NPV were also compared between CAD-NBI, CAD-WLI, and the endoscopists.

Statistical analysis

The chi-square test and t test were used wherever applicable. A *P* value of < 0.05 was considered to be statistically significant. A two-sided McNemar test with a significance level of 0.05 was used to compare differences in accuracy, sensitivity, specificity, PPV, and NPV. All statistical analyses were performed using SPSS version 18.0 (SPSS Inc., Chicago, IL, United States).

RESULTS

The detailed characteristics of the patients and lesions in the validation image dataset

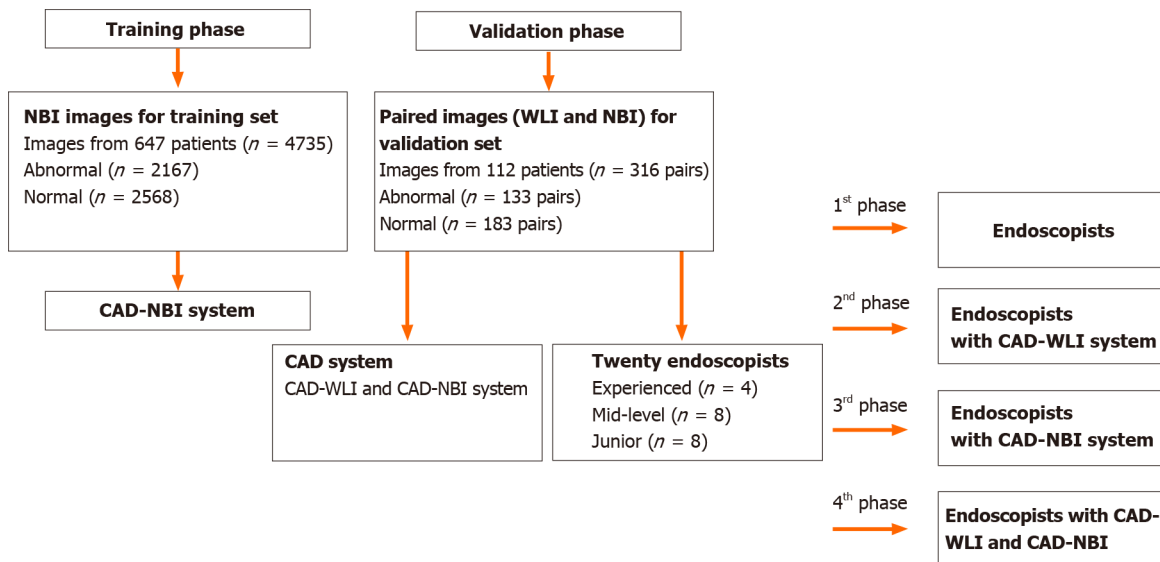


Figure 1 Flowchart of the procedures. CAD: Computer-assisted detection; NBI: Narrow-band imaging; WLI: White -light imaging.

are listed in [Table 1](#).

Performance of the CAD-NBI and CAD-WLI systems

The receiver operating characteristic curve of the CAD-NBI system is shown in [Figure 2](#), and the area under the curve was 0.9761. For 316 NBI images in the validation dataset, the sensitivity, specificity, accuracy, PPV, and NPV of CAD-NBI were 91.0%, 96.7%, 94.3%, 95.3%, and 93.6%, respectively. For 316 WLI images in the validation dataset, CAD-WLI correctly identified 131 of the 133 early ESCC lesions, with sensitivity, specificity, and accuracy of 98.5%, 83.1%, and 89.5%, respectively. The PPV and NPV of the CAD-WLI system were 80.8% and 98.7%, respectively.

Comparison of the two CAD systems ([Figure 3](#)) revealed that the accuracy and specificity of CAD-NBI were superior to those of CAD-WLI ($P = 0.028$ and $P \leq 0.001$). However, the sensitivity of the CAD-WLI system was higher than that of the CAD-NBI system ($P = 0.006$). The CAD-WLI and CAD-NBI system recognized and marked the lesion with a blue square in paired images ([Figure 4A](#) and [B](#)). In addition, when CAD-WLI mistakenly identified the normal mucosa as lesion (blue square) in WLI, CAD-NBI with high specificity could correct it in NBI ([Figure 4C](#) and [D](#)).

Comparison between CAD systems and the endoscopists

[Table 2](#) compares the performance of the two CAD systems and the endoscopists for diagnosing early ESCC. The overall accuracy, sensitivity, specificity, PPV, and NPV of the 20 endoscopists were 73.9%, 87.7%, 81.9%, 81.7%, and 82.7%, respectively. Apparently, the experienced endoscopists achieved significantly better diagnostic results than the less experienced ones, including mid-level and junior endoscopists. The average accuracy value of the experienced endoscopists for early ESCC was 93.6%, which was similar to that of the CAD-NBI system and higher than that of the CAD-WLI system. CAD-WLI achieved the highest sensitivity (98.5%), whereas its specificity was lower than that of CAD-NBI with the highest value of 96.7% and the average value of all the endoscopists (87.7%).

Improvement after referring to the results from CAD-WLI and CAD-NBI

With the assistance of either CAD-WLI or CAD-NBI, all the three groups of endoscopists showed improvement in accurately diagnosing early ESCC ([Figure 5](#)). [Table 3](#) shows the average diagnostic performance of endoscopists in the detection of early ESCC after referring to the results from the CAD-WLI system in the second phase and from the CAD-NBI system in the third phase. Next, we compared the advantages of the two systems in different aspects. The CAD-NBI system helped the endoscopists to achieve higher value than that achieved with the assistance of CAD-WLI system, especially in the mid-level group with a significant difference (85.3% *vs* 88.4%, $P = 0.012$). Experienced and mid-level endoscopists showed no significant differences in their sensitivity for lesions in the two phases, while the CAD-WLI system helped junior endoscopists to achieve higher sensitivity than that achieved

Table 1 Patient and lesion characteristics in the validation image set

Patient characteristics, <i>n</i> = 112	Value
Median age, yr (range)	59 (19-86)
Sex, male/female	67/45
Lesion characteristics, <i>n</i> = 42	
Median size, mm (range)	23 (9-42)
Location, Ce/Ut/Mt/Lt/Ae	0/3/24/15/0
Pathological diagnosis	
LGIN/HGIN	6/18
Cancer, M/SM1	15/3

Ae: Abdominal esophagus; Ce: Cervical esophagus; HGIN: High-grade intraepithelial neoplasia; LGIN: Low-grade intraepithelial neoplasia; Lt: Lower thoracic esophagus; M: Mucosal; Mt: Middle thoracic esophagus; SM1: Submucosal; Ut: Upper thoracic esophagus.

Table 2 Diagnostic performance of computer-assisted detection systems vs endoscopists

	CAD systems		Endoscopists			
	CAD-NBI	CAD-WLI	All, <i>n</i> = 20	Experienced, <i>n</i> = 4	Mid-level, <i>n</i> = 8	Junior, <i>n</i> = 8
Sensitivity, % (95%CI)	91.0	98.5	73.9 (68.1-79.7)	94.7 (85.0-100)	73.9 (71.9-75.8)	63.5 (59.7-67.3)
Specificity, % (95%CI)	96.7	83.1	87.7 (84.9-90.5)	92.8 (81.2-100)	83.4 (81.1-85.7)	89.4 (85.0-93.8)
Accuracy, % (95%CI)	94.3	89.5	81.9 (78.9-84.8)	93.6 (89.1-98.1)	79.4 (78.3-80.5)	78.5 (76.6-80.4)
PPV, % (95%CI)	95.3	80.8	81.7 (78.0-85.5)	91.4 (78.0-100)	76.5 (74.2-78.8)	82.2 (76.1-88.2)
NPV, % (95%CI)	93.6	98.7	82.7 (79.2-86.3)	96.4 (90.1-100)	81.5 (80.5-82.4)	77.2 (75.8-78.6)

CAD: Computer-assisted detection; CI: Confidence interval; NBI: Narrow-band imaging; NPV: Negative predictive value; PPV: Positive predictive value; WLI: White-light imaging.

using the CAD-NBI system (83.0% *vs* 77.3%, $P = 0.008$). In addition, there were no significant differences in the specificity of experienced endoscopists when using CAD-WLI or CAD-NBI. However, a significant improvement in diagnostic specificity was shown by mid-level (85.9% *vs* 92.6%, $P = 0.000$) and junior endoscopists (88.6% *vs* 94.9%, $P = 0.003$).

Table 4 shows the average diagnostic performance of endoscopists after referring to the results from both the CAD-WLI and CAD-NBI systems. In the fourth phase, the diagnostic capability of all the endoscopists improved to the highest level, with the accuracy, sensitivity, and specificity of 94.9%, 92.4%, and 96.7%, respectively. The accuracy of junior endoscopists was 92.9%, and it was significantly higher than that in the first (*vs* 78.5%, $P = 0.000$), second (*vs* 86.2%, $P = 0.000$), and third (*vs* 87.5%, $P = 0.000$) phases. The accuracy of mid-level endoscopists was 94.8%, and it was significantly higher than that in the first (*vs* 79.4%, $P = 0.000$), second (*vs* 85.3%, $P = 0.000$), and third (*vs* 88.4%, $P = 0.000$) phases. In the experienced endoscopist group, the average accuracy value was 98.8%, and it was also significantly higher than that in the first (*vs* 93.6%, $P = 0.011$), second (*vs* 95.5%, $P = 0.015$), and third (*vs* 96.5%, $P = 0.049$) phases (Figure 6A). In terms of sensitivity and specificity for lesions, the average values of the mid-level and junior groups also increased to the highest value after using CAD-WLI and CAD-NBI (Figure 6B and C).

DISCUSSION

CAD has been developed to overcome the limitation of less experience of diagnosis in young doctors. A recent study presented an AI system that can surpass human experts in breast cancer prediction^[13]. In the field of gastroenterology, several CAD systems have shown excellent diagnostic potential compared with human endoscopists in

Table 3 Comparison of the improvement of endoscopists under computer-assisted detection-white-light imaging and computer-assisted detection-narrow-band imaging

	2 nd phase CAD-WLI assistance	3 rd phase CAD-NBI assistance	P value
Sensitivity, % (95%CI)			
All, n = 20	86.4 (83.1-89.6)	83.1 (79.6-86.6)	0.162
Experienced, n = 4	96.8 (94.4-99.2)	95.7 (91.8-99.5)	0.454
Mid-level, n = 8	84.5 (80.0-89.0)	82.6 (79.4-85.9)	0.435
Junior, n = 8	83.0 (79.3-86.6)	77.3 (74.9-79.7)	0.008
Specificity, % (95%CI)			
All, n = 20	88.7 (86.5-90.8)	94.4 (93.0-95.8)	0.000
Experienced, n = 4	94.5 (87.9-100)	97.2 (93.6-100)	0.310
Mid-level, n = 8	85.9 (84.3-87.4)	92.6 (90.7-94.5)	0.000
Junior, n = 8	88.6 (85.1-92.1)	94.9 (92.5-97.2)	0.003
Accuracy, % (95%CI)			
All, n = 20	87.7 (85.5-89.9)	89.7 (87.8-91.5)	0.156
Experienced, n = 4	95.5 (92.4-98.6)	96.5 (93.6-99.4)	0.469
Mid-level, n = 8	85.3 (83.1-87.5)	88.4 (87.1-89.7)	0.012
Junior, n = 8	86.2 (84.2-88.3)	87.5 (86.0-89.0)	0.261
PPV, % (95%CI)			
All, n = 20	84.9 (82.2-87.5)	91.6 (89.7-93.6)	0.000
Experienced, n = 4	93.0 (85.1-100)	96.1 (91.4-100)	0.325
Mid-level, n = 8	81.3 (79.4-83.2)	89.2 (86.7-91.7)	0.000
Junior, n = 8	84.4 (80.5-88.2)	91.8 (88.4-95.2)	0.004
NPV, % (95%CI)			
All, n = 20	90.1 (87.9-92.3)	88.7 (86.5-90.9)	0.361
Experienced, n = 4	97.7 (96.0-99.3)	96.9 (94.2-99.6)	0.465
Mid-level, n = 8	88.5 (85.6-91.4)	88.1 (86.2-89.9)	0.755
Junior, n = 8	87.8 (85.6-90.1)	85.2 (83.9-86.5)	0.031

CAD: Computer-assisted detection; CI: Confidence interval; NBI: Narrow-band imaging; NPV: Negative predictive value; PPV: Positive predictive value; WLI: Whit-light imaging.

colorectal polyp classification, determination of the invasion depth of gastric cancer, and identification of small bowel diseases^[14-16]. The application of AI in automatically detecting and classifying lesions, especially in the context of medical imaging, is expected to help physicians provide more accurate diagnoses^[17].

Squamous cell carcinoma is the predominant histologic subtype of EC in Asia, where the rate of EC is quite high^[18]. Several research studies on CAD for improving the screening efficiency for ESCC have been reported. Horie *et al*^[19] first evaluated the ability of the CNN to detect EC in endoscopic images of superficial and advanced cancer and achieved a sensitivity of 98%. Zhao *et al*^[20] developed a deep learning model based on magnifying NBI images to investigate the automated classification of intrapapillary capillary loops and assist endoscopic diagnosis of early ESCC. In a recent study by Guo *et al*^[21], the authors developed a CAD system for real-time automated diagnosis of precancerous lesions and ESCC. In 2019, our team reported a novel system using deep neural network (DNN) to localize and identify early ESCC under endoscopic WLI with high accuracy and sensitivity^[12]. Moreover, after referring to the results of DNN-CAD, the average diagnostic ability of the endoscopists improved significantly. However, this CAD system could only identify early ESCC in WLI, and the specificity was only 85.4%, which may lead to unnecessary biopsies. As NBI is also an accurate diagnostic tool for early ESCC, a better CAD system that can be

Table 4 Diagnostic performance of endoscopists in screening of esophagus squamous cell carcinoma after referring to the results from computer-assisted detection-white-light imaging and computer-assisted detection-narrow-band imaging

	Sensitivity, % (95%CI)	Specificity, % (95%CI)	Accuracy, % (95%CI)	PPV, % (95%CI)	NPV, % (95%CI)
All, <i>n</i> = 20	92.4 (90.3-94.5)	96.7 (95.7-97.7)	94.9 (93.6-96.1)	95.3 (93.9-96.7)	94.7 (93.2-96.1)
Experienced, <i>n</i> = 4	98.5 (96.3-100)	99.1 (97.5-100)	98.8 (98.2-99.4)	98.7 (96.7-100)	98.9 (97.4-100)
Mid-level, <i>n</i> = 8	93.2 (91.3-95.1)	96.0 (94.5-97.6)	94.8 (93.7-95.9)	94.5 (92.5-96.6)	95.2 (93.8-96.5)
Junior, <i>n</i> = 8	88.5 (86.2-90.8)	96.1 (94.3-97.9)	92.9 (91.2-94.6)	94.3 (91.8-96.8)	92.0 (90.5-93.6)

CI: Confidence interval; NPV: Negative predictive value; PPV: Positive predictive value.

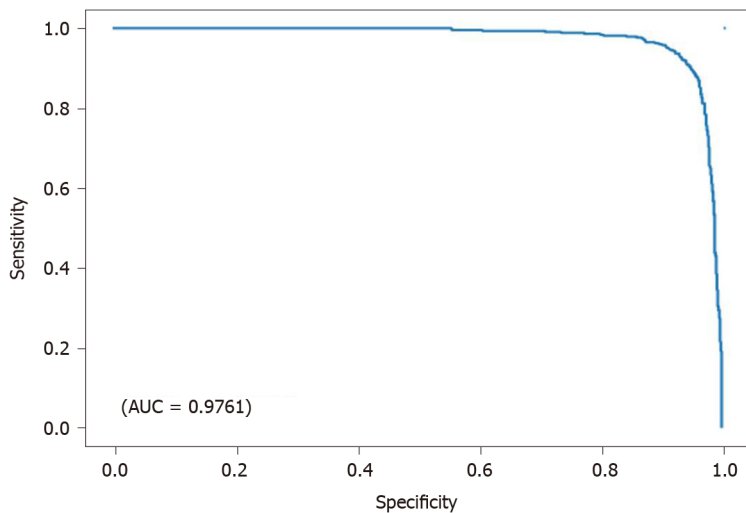


Figure 2 Receiver operating characteristic curve for the test dataset. The area under the curve (AUC) was above 97%.

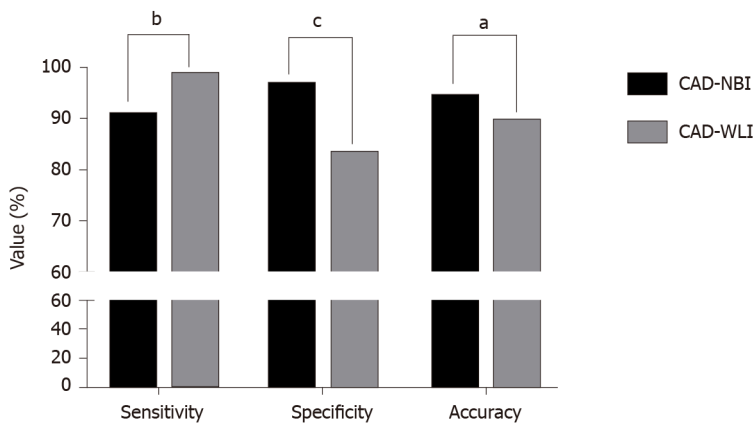


Figure 3 A comparison between the computer-assisted detection-narrow-band imaging and computer-assisted detection-white-light imaging systems in detecting early esophageal squamous cell carcinoma. ^a*P* < 0.05; ^b*P* < 0.01; ^c*P* < 0.001. CAD: Computer-assisted detection; NBI: Narrow-band imaging; WLI: White-light imaging.

used in the NBI model of endoscopy needs to be developed. In addition, a comparative study of AI application in WLI and NBI models is lacking.

In the present study, we developed a CAD system to detect early ESCC under the NBI model of endoscopy. Considering our previously developed CAD system for WLI, we wanted to compare the different characteristics of CAD-NBI and CAD-WLI systems to validate the usefulness of CAD-NBI. Therefore, 316 pairs of images, each pair including WLI and NBI at the same location and at the same angle, were collected from three institutions. The results showed that CAD for NM-NBI images of the esophagus had a good ability to diagnose early ESCC, with an accuracy of 94.3%. For

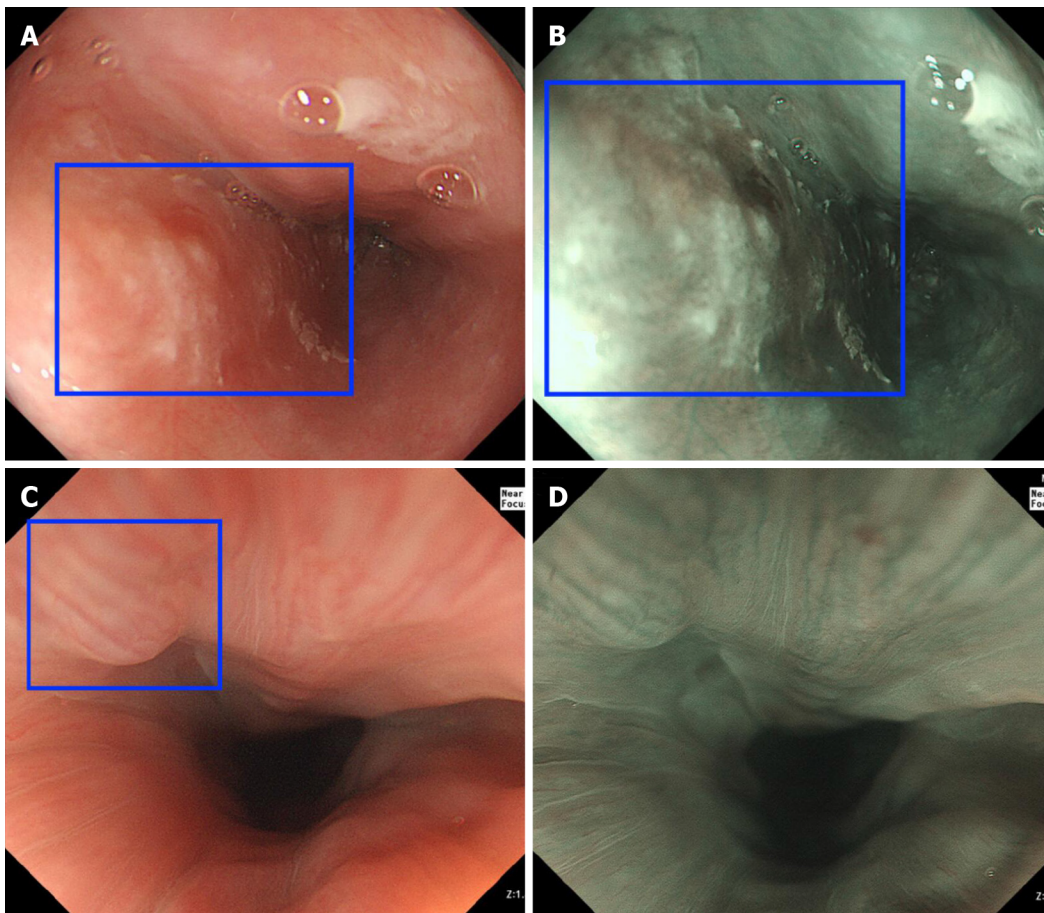


Figure 4 Examples of computer-assisted detection system-diagnosed images. A and B: under white-light imaging (WLI) and narrow-band imaging (NBI), computer-assisted detection (CAD)-WLI and CAD-NBI recognized the esophageal cancer lesion (blue square); C and D: CAD-WLI mistakenly identified the normal mucosa as a lesion (blue square) in WLI, while CAD-NBI corrected it in NBI.

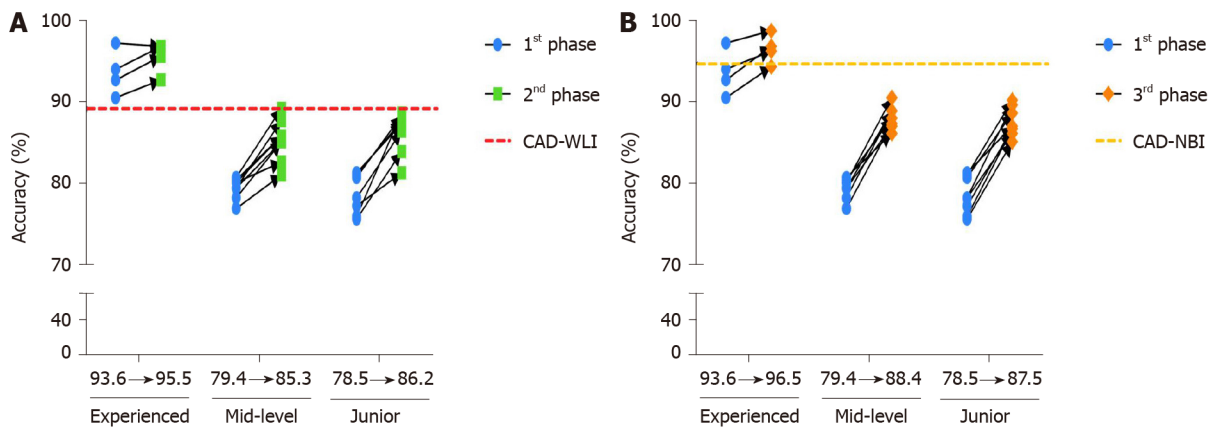


Figure 5 Improved accuracy of diagnosis with the assistance of the two computer-assisted detection systems according to the groups. A: Improvement of endoscopists' accuracy in the second phase with the assistance of computer-assisted detection (CAD)-white-light imaging; B: Improvement of endoscopists' accuracy in the third phase with the assistance of CAD-narrow-band imaging. WLI: White-light imaging.

WLI images in the validation dataset, CAD-WLI correctly identified 131 of the 133 lesions. The diagnostic ability of CAD-WLI for this validation image dataset was similar to that reported in our previous study^[12]. Comparison of the two systems showed that CAD-NBI had superior accuracy and specificity compared to CAD-WLI, while the CAD-WLI system had higher sensitivity than the CAD-NBI system. The results showed that CAD-NBI may compensate for the feature of low specificity of CAD-WLI and make the overall accuracy higher, but its own sensitivity still needs to be improved further.

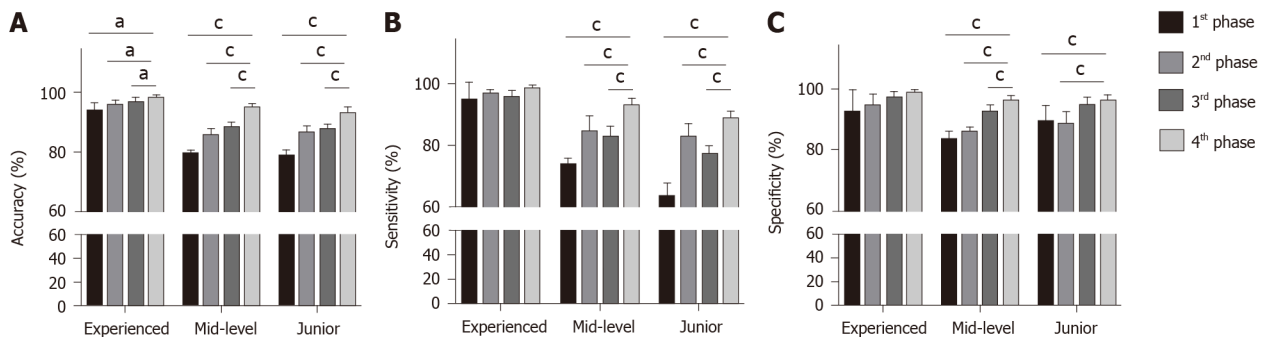


Figure 6 Average diagnostic performance of the three groups of endoscopists in four phases. A: Accuracy; B: Sensitivity; C: Specificity. ^a $P < 0.05$; ^c $P < 0.001$.

Endoscopists were asked to review the images from this validation dataset to evaluate their diagnostic ability. In the first phase, the average accuracy value of the experienced endoscopists for early ESCC was 93.6%, which was similar to that of the CAD-NBI system and higher than that of the CAD-WLI system. However, the average diagnostic accuracy of the less experienced endoscopists, including the mid-level (79.4%) and junior groups (78.5%), was lower. Subsequently, the less experienced endoscopists in particular showed improvement in their diagnostic ability with the help of both CAD-WLI in the second phase and CAD-NBI in the third phase. Junior endoscopists made a greater improvement in terms of sensitivity with the CAD-WLI system when their diagnostic specificity was further improved after referring to the results from CAD-NBI. In addition, mid-level endoscopists showed higher values of specificity and accuracy with the assistance of the CAD-NBI system than with the CAD-WLI system. In the fourth phase, by using both CAD-WLI and CAD-NBI, the average values of the three groups increased to the highest value. After simulating the clinical use of the two CAD systems through four different situations, we found that the two systems had different advantages in terms of avoiding missed diagnosis and performing excessive biopsy, and thus, endoscopists could achieve the best diagnostic efficacy by using both CAD-WLI and CAD-NBI.

The sensitivity of a previous system reported by Horie *et al*^[19] for the diagnosis of ESCC on WLI and NBI was 72% and 86%, respectively, and our CAD-WLI and CAD-NBI systems provided higher sensitivity of 98.5% and 91.0%, respectively. We have continuously developed the two CAD systems for WLI and NBI based on DNN and conducted for the first time a comparative study to reveal the respective advantages of both systems. The ultimate goal is to integrate the two systems into one to meet the needs of different equipment in a hospital at all levels from different regions. Although Guo *et al*^[21] reported that their system for the automated diagnosis of early ESCC under NBI had sensitivity and specificity of 98.04% and 95.03%, respectively, the sensitivity of CAD-WLI and specificity of CAD-NBI were similar to their reported values. In addition, our study invited 20 endoscopists to review the images of the validation dataset with or without the help of the two CAD systems; 10 of the 20 endoscopists (2 mid-level and 8 junior) are from Xuhui Hospital, which is a secondary hospital where the number of patients is much less than that in the tertiary hospital. As secondary hospitals have lower capability to diagnose and treat ESCC than tertiary hospitals, we wanted to assess fully the effectiveness of our CAD system in helping less experienced endoscopists to detect early ESCC, especially doctors in basic hospitals. Our results confirmed that the improvement of diagnosis capability was most pronounced in less experienced endoscopists.

Next, we will explain the reasons for the differences in the diagnostic characteristics of the two CAD systems. CAD-WLI and CAD-NBI are based on different concepts. CAD-WLI uses a bounding box method to detect the location of esophagus lesions in endoscopic images. The detection result of the CAD-WLI system unavoidably includes unnecessary areas such as background regions and parts of other lesions. In contrast, CAD-NBI uses the object semantic segmentation method based on the FCN model, which only outputs the regions of esophagus lesions. Therefore, the accuracy of the CAD-NBI system is higher than that of the CAD-WLI system. In addition, the CAD-NBI system is developed by an end-to-end trainable approach, which is different from the sliding window approach of the CAD-WLI system that densely generates a large number of candidate boxes with different sizes and ratios on a given image. Thus, the CAD-NBI system has the advantage of computational speed when compared with the

CAD-WLI system. Finally, NBI and WLI have different characteristics, which can help us obtain high accuracy, sensitivity, and specificity simultaneously. On the basis of the key observation, we can develop a multichannel DNN to extract and fuse the features of NBI and WLI simultaneously in future studies.

The present study has several limitations. First, the sample size, including images in the training and validation datasets, was small. The low detection rate of early ESCC limits our ability to obtain more images. In addition, our work on CAD for the early detection of ESCC was validated only on still images with a limited scale. Second, the performance of endoscopists in the latter phase may be slightly affected by the image impression of the previous phase. Third, the CAD diagnosis was based on high-quality images, and bias might occur with poor-quality images such as out-of-focus images or blurred images caused by mucus during real-time gastroscopy.

CONCLUSION

In conclusion, we have constructed a CAD system under the NBI model for screening early ESCC, and this CAD-NBI system has higher accuracy and specificity than the CAD-WLI system reported previously. Endoscopists could achieve the best diagnostic efficacy by using both CAD-WLI and CAD-NBI. Therefore, a novel system combining the characteristics of these two systems under WLI and NBI is needed.

ARTICLE HIGHLIGHTS

Research background

Non-magnifying endoscopy with narrow-band imaging (NM-NBI) has been frequently used in routine screening of esophagus squamous cell carcinoma (ESCC). The performance of NBI for screening of early ESCC is, however, significantly affected by operator experience. Artificial intelligence may be a unique approach to compensate for the lack of operator experience.

Research motivation

In our previous research, we reported a novel system of computer-aided detection (CAD) to localize and identify early ESCC under conventional endoscopic white-light imaging (WLI) with sensitivity of above 97%. The construction of another CAD system for application in NM-NBI was the next step in the continuation of our research.

Research objectives

To construct a CAD system for application in NM-NBI to identify early ESCC and compare it with our previously reported CAD system with endoscopic WLI.

Research methods

We collected a total of 2167 abnormal NM-NBI images of early ESCC and 2568 normal images from three institutions (Zhongshan Hospital of Fudan University, Xuhui Hospital, and Kiang Wu Hospital) as the training dataset, and 316 pairs of images, each pair including images obtained with WLI and NBI (same part), were collected for validation. Twenty endoscopists participated in this study to review the validation images with or without the assistance of the CAD systems. The diagnostic results of the two CAD systems and the improvement in the diagnostic efficacy of endoscopists were compared in terms of sensitivity, specificity, accuracy, positive predictive value, and negative predictive value.

Research results

The area under receiver operating characteristic curve for CAD-NBI was 0.9761. For the validation dataset, the sensitivity, specificity, accuracy, positive predictive value, and negative predictive value of CAD-NBI were 91.0%, 96.7%, 94.3%, 95.3%, and 93.6%, respectively, while those of CAD-WLI were 98.5%, 83.1%, 89.5%, 80.8%, and 98.7%, respectively. CAD-NBI showed superior accuracy and specificity than CAD-WLI ($P = 0.028$ and $P \leq 0.001$, respectively), while CAD-WLI had higher sensitivity than CAD-NBI ($P = 0.006$). By using both CAD-WLI and CAD-NBI, the endoscopists could improve their diagnostic efficacy to the highest level, with accuracy, sensitivity, and specificity of 94.9%, 92.4%, and 96.7%, respectively.

Research conclusions

The CAD-NBI system for screening early ESCC has higher accuracy and specificity than CAD-WLI. Endoscopists can achieve the best diagnostic efficacy by using both CAD-WLI and CAD-NBI.

Research perspectives

According to the results, the two CAD systems had different advantages in avoiding missed diagnosis and excessive biopsy, which could help endoscopists, especially those with less experience, in screening of early ESCC more efficiently. As the two CAD systems have unique characteristics, we plan to develop a multichannel deep neural network to extract and combine the features of NBI and WLI simultaneously in our future work.

REFERENCES

- 1 **Wei WQ**, Chen ZF, He YT, Feng H, Hou J, Lin DM, Li XQ, Guo CL, Li SS, Wang GQ, Dong ZW, Abnet CC, Qiao YL. Long-Term Follow-Up of a Community Assignment, One-Time Endoscopic Screening Study of Esophageal Cancer in China. *J Clin Oncol* 2015; **33**: 1951-1957 [PMID: 25940715 DOI: 10.1200/JCO.2014.58.0423]
- 2 **He Z**, Liu Z, Liu M, Guo C, Xu R, Li F, Liu A, Yang H, Shen L, Wu Q, Duan L, Li X, Zhang C, Pan Y, Cai H, Ke Y. Efficacy of endoscopic screening for esophageal cancer in China (ESECC): design and preliminary results of a population-based randomised controlled trial. *Gut* 2019; **68**: 198-206 [PMID: 29306867 DOI: 10.1136/gutjnl-2017-315520]
- 3 **Lee YC**, Wang CP, Chen CC, Chiu HM, Ko JY, Lou PJ, Yang TL, Huang HY, Wu MS, Lin JT, Hsiu-Hsi Chen T, Wang HP. Transnasal endoscopy with narrow-band imaging and Lugol staining to screen patients with head and neck cancer whose condition limits oral intubation with standard endoscope (with video). *Gastrointest Endosc* 2009; **69**: 408-417 [PMID: 19019362 DOI: 10.1016/j.gie.2008.05.033]
- 4 **Morita FH**, Bernardo WM, Ide E, Rocha RS, Aquino JC, Minata MK, Yamazaki K, Marques SB, Sakai P, de Moura EG. Narrow band imaging versus lugol chromoendoscopy to diagnose squamous cell carcinoma of the esophagus: a systematic review and meta-analysis. *BMC Cancer* 2017; **17**: 54 [PMID: 28086818 DOI: 10.1186/s12885-016-3011-9]
- 5 **Muto M**, Katada C, Sano Y, Yoshida S. Narrow band imaging: a new diagnostic approach to visualize angiogenesis in superficial neoplasia. *Clin Gastroenterol Hepatol* 2005; **3**: S16-S20 [PMID: 16012987 DOI: 10.1016/s1542-3565(05)00262-4]
- 6 **Nagami Y**, Tominaga K, Machida H, Nakatani M, Kameda N, Sugimori S, Okazaki H, Tanigawa T, Yamagami H, Kubo N, Shiba M, Watanabe K, Watanabe T, Iguchi H, Fujiwara Y, Ohira M, Hirakawa K, Arakawa T. Usefulness of non-magnifying narrow-band imaging in screening of early esophageal squamous cell carcinoma: a prospective comparative study using propensity score matching. *Am J Gastroenterol* 2014; **109**: 845-854 [PMID: 24751580 DOI: 10.1038/ajg.2014.94]
- 7 **Muto M**, Minashi K, Yano T, Saito Y, Oda I, Nonaka S, Omori T, Sugiura H, Goda K, Kaise M, Inoue H, Ishikawa H, Ochiai A, Shimoda T, Watanabe H, Tajiri H, Saito D. Early detection of superficial squamous cell carcinoma in the head and neck region and esophagus by narrow band imaging: a multicenter randomized controlled trial. *J Clin Oncol* 2010; **28**: 1566-1572 [PMID: 20177025 DOI: 10.1200/JCO.2009.25.4680]
- 8 **Ishihara R**, Takeuchi Y, Chatani R, Kidu T, Inoue T, Hanaoka N, Yamamoto S, Higashino K, Uedo N, Iishi H, Tatsuta M, Tomita Y, Ishiguro S. Prospective evaluation of narrow-band imaging endoscopy for screening of esophageal squamous mucosal high-grade neoplasia in experienced and less experienced endoscopists. *Dis Esophagus* 2010; **23**: 480-486 [PMID: 20095991 DOI: 10.1111/j.1442-2050.2009.01039.x]
- 9 **Rodríguez de Santiago E**, Hernanz N, Marcos-Prieto HM, De-Jorge-Turrión MÁ, Barreiro-Alonso E, Rodríguez-Escaja C, Jiménez-Jurado A, Sierra-Morales M, Pérez-Valle I, Machado-Volpato N, García-Prada M, Núñez-Gómez L, Castaño-García A, García García de Paredes A, Peñas B, Vázquez-Sequeiros E, Albillos A. Rate of missed oesophageal cancer at routine endoscopy and survival outcomes: A multicentric cohort study. *United European Gastroenterol J* 2019; **7**: 189-198 [PMID: 31080603 DOI: 10.1177/2050640618811477]
- 10 **Topol EJ**. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019; **25**: 44-56 [PMID: 30617339 DOI: 10.1038/s41591-018-0300-7]
- 11 **Mori Y**, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, Ohtsuka K, Urushibara F, Kataoka S, Ogawa Y, Maeda Y, Takeda K, Nakamura H, Ichimasa K, Kudo T, Hayashi T, Wakamura K, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. *Ann Intern Med* 2018; **169**: 357-366 [PMID: 30105375 DOI: 10.7326/M18-0249]
- 12 **Cai SL**, Li B, Tan WM, Niu XJ, Yu HH, Yao LQ, Zhou PH, Yan B, Zhong YS. Using a deep learning system in endoscopy for screening of early esophageal squamous cell carcinoma (with video). *Gastrointest Endosc* 2019; **90**: 745-753. e2 [PMID: 31302091 DOI: 10.1016/j.gie.2019.05.033]

- 10.1016/j.gie.2019.06.044]
- 13 **McKinney SM**, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, Back T, Chesu M, Corrado GS, Darzi A, Etemadi M, Garcia-Vicente F, Gilbert FJ, Halling-Brown M, Hassabis D, Jansen S, Karthikesalingam A, Kelly CJ, King D, Ledsam JR, Melnick D, Mostofi H, Peng L, Reicher JJ, Romera-Paredes B, Sidebottom R, Suleyman M, Tse D, Young KC, De Fauw J, Shetty S. International evaluation of an AI system for breast cancer screening. *Nature* 2020; **577**: 89-94 [PMID: 31894144 DOI: 10.1038/s41586-019-1799-6]
 - 14 **Chen PJ**, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology* 2018; **154**: 568-575 [PMID: 29042219 DOI: 10.1053/j.gastro.2017.10.010]
 - 15 **Zhu Y**, Wang QC, Xu MD, Zhang Z, Cheng J, Zhong YS, Zhang YQ, Chen WF, Yao LQ, Zhou PH, Li QL. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. *Gastrointest Endosc* 2019; **89**: 806-815. e1 [PMID: 30452913 DOI: 10.1016/j.gie.2018.11.011]
 - 16 **Ding Z**, Shi H, Zhang H, Meng L, Fan M, Han C, Zhang K, Ming F, Xie X, Liu H, Liu J, Lin R, Hou X. Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. *Gastroenterology* 2019; **157**: 1044-1054. e5 [PMID: 31251929 DOI: 10.1053/j.gastro.2019.06.025]
 - 17 **He J**, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019; **25**: 30-36 [PMID: 30617336 DOI: 10.1038/s41591-018-0307-0]
 - 18 **Arnold M**, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut* 2015; **64**: 381-387 [PMID: 25320104 DOI: 10.1136/gutjnl-2014-308124]
 - 19 **Horie Y**, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, Hirasawa T, Tsuchida T, Ozawa T, Ishihara S, Kumagai Y, Fujishiro M, Maetani I, Fujisaki J, Tada T. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest Endosc* 2019; **89**: 25-32 [PMID: 30120958 DOI: 10.1016/j.gie.2018.07.037]
 - 20 **Zhao YY**, Xue DX, Wang YL, Zhang R, Sun B, Cai YP, Feng H, Cai Y, Xu JM. Computer-assisted diagnosis of early esophageal squamous cell carcinoma using narrow-band imaging magnifying endoscopy. *Endoscopy* 2019; **51**: 333-341 [PMID: 30469155 DOI: 10.1055/a-0756-8754]
 - 21 **Guo L**, Xiao X, Wu C, Zeng X, Zhang Y, Du J, Bai S, Xie J, Zhang Z, Li Y, Wang X, Cheung O, Sharma M, Liu J, Hu B. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). *Gastrointest Endosc* 2020; **91**: 41-51 [PMID: 31445040 DOI: 10.1016/j.gie.2019.08.018]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: <https://www.f6publishing.com/helpdesk>
<https://www.wjgnet.com>

