EVIDENCE-BASED GI
AN ACG PUBLICATION

AI-Assisted Endoscopy: Can AI Change Gastric Cancer Detection?

Margaret J. Zhou, MD, MSc

Clinical Assistant Professor of Medicine, Gastroenterology & Hepatology, Stanford University School of Medicine, Stanford, CA

This summary reviews Dong Z, Wu L, Du H, et al. Effect of a computer-aided device for detecting gastric neoplasms: a multicenter, randomized controlled trial. Gastroenterology. 2026; 170(7): 1518-1532.

Correspondence to Margaret J. Zhou, MD, MSc. Associate Editor. Email: EBGI@gi.org

Keywords: Artificial intelligence, gastric cancer, gastric neoplasms, endoscopy

ENDOSCOPY

STRUCTURED ABSTRACT

Question: Does artificial intelligence (AI)-assisted computer-aided detection during esophagogastroduodenoscopy (EGD) improve the detection rate of gastric neoplasms compared with standard EGD?

Design: This is a multicenter, single-blind open-label, parallel-group, randomized controlled trial (RCT) with 1:1 allocation.

Setting: It was set across 24 hospitals in China with enrollment from December 21, 2021, to November 11, 2023.

Patients: The patients were adults ≥18 years undergoing sedated EGD for screening, surveillance, or evaluation of upper gastrointestinal (GI) symptoms. Participating endoscopists were required to perform at least 500 EGDs and complete a training module with at least 10 AI-assisted exams.

Intervention: Patients were randomized to AI-assisted EGD (using the ENDOANGEL-GN system) or standard non-assisted EGD. The AI system, built

on deep convolutional neural networks, provided real-time anatomic landmark identification and blind spot monitoring during the procedure as well as lesion detection with risk stratification. Detected lesions are demarcated by a blue box, and if the lesion is predicted to be high-risk for neoplasia, the box turns red, and if low-risk the box stays blue (**Figure 1**).[1] Use of image-enhanced modalities and decisions about taking biopsies for lesions detected with a blue box were decided at the discretion of the endoscopist. Lesions highlighted with a red box required biopsy.

Outcomes: The primary outcome was the detection rate of gastric neoplasms after centralized endoscopic and pathologic review. Secondary outcomes included: 1) detection rate of gastric neoplasms based on original clinical pathology (before centralized pathologic review); 2) relative early gastric cancer (EGC) detection ratio with or without centralized pathology review; 3) detection rate of intestinal metaplasia (IM) or gastric atrophy before or after pathologic review; 4) biopsy rate; 5) procedure time; 6) inspection time; and 7) number of blind spots. EGC was defined as cancer limited to the mucosa or submucosa with or without regional lymph node metastasis (including high-grade intraepithelial neoplasia [HGIN]).

Original pathology is defined as the initial pathologic diagnosis provided by each center before centralized review. Centralized endoscopic and pathologic review occurred for all gastric neoplasms and non-neoplastic lesions that had suspicious features on endoscopy.

Expert panels consisted of endoscopists with >10 years of experience and pathologists with >15 years of experience. Expert pathologists were blinded to the initial pathologic diagnosis. Both panels were blinded to lesion randomization status. For lesions that underwent centralized review, the final diagnosis required agreement from ≥1 expert pathologist for preliminary neoplastic diagnosis, and agreement from ≥2 expert pathologists for preliminary non-neoplastic diagnosis to reclassify to neoplastic.

Data Analysis: Intention-to-treat (ITT), per-protocol (PP), and exploratory subgroup analyses were performed. Risk ratios (RR) with 95% confidence intervals (CI) were reported.

Funding: National Key Research and Development Program of China, Natural Science Foundation of Wuhan, Key Research and Development Program of Hubei Province, College-enterprise Deepening Reform Project of Wuhan University, National Natural Science Foundation of China-Youth Science Fund

Results: In the ITT analysis, 29,514 patients were included and 26,001 were included in the PP analysis. Some patients were excluded from the PP analysis

due to uncertainty about group allocation and deviations from the protocol. Some endoscopists in the control group activated the AI function to evaluate lesions.

The primary outcome of the study was that AI did not significantly improve the detection rate of gastric neoplasms after centralized pathologic review (1.42% vs 1.25%; RR 1.13, 95% CI 0.92–1.38; $P = 0.25$) in the ITT analysis.

The secondary outcomes were that, based on original pathology (before expert review), AI-assisted EGD showed a statistically significantly higher rate of detection of gastric neoplasms (4.06% vs 3.57%; RR 1.14, 95% CI 1.0–1.28; $P = 0.03$). Comparing EGD with versus AI assistance, relative EGC detection ratio was similar (RR, 1.06; 95% CI, 0.62-1.81; 41.46% vs 39.05%; $P = 0.81$) and detection of IM or gastric atrophy was similar (2,703 of 14,767 [18.30%] vs 2650 of 14,747 [17.97%]; $P = 0.46$). Rate of biopsy was also comparable (54% vs 53%; $P = 0.20$). Similar findings were seen in the PP and ITT analyses.

AI reduced the mean number of blind spots from 2.52 to 1.07 ($P <0.001$), and AI prolonged both procedure and inspection time (procedure time with AI 7.69 vs without AI 7.33 min; $P <0.001$; inspection time with AI 7.35 min vs 6.99 min; $P <0.001$).

Among endoscopists with <3 years experience, there was a statistically significant higher detection rate of gastric neoplasms with AI (1.44% vs 0.78%; RR, 1.83; 95% CI, 1.04–3.21). In the PP analysis, RR of neoplasm detection was higher when endoscopists were in the fatigue period than at the nonfatigue period (RR 1.60, 1.07–2.40 vs 1.03, 0.80–1.33, P for interaction =0.07).

In the experimental group, AI correctly identified 100% of pathologically confirmed gastric adenocarcinomas, 91.9% of high-grade intraepithelial neoplasia, and 57.1% of low-grade intraepithelial neoplasia.

COMMENTARY

Why Is This Important?

This is the largest RCT to date on AI-assisted gastric neoplasm detection and provides critical real-world evidence on the impact of AI-assisted EGD. Prior retrospective and smaller prospective studies demonstrated high accuracy of AI for upper GI neoplasia detection, but robust multicenter RCT evidence was lacking. A recent meta-analysis of 11 RCTs (57,512 participants) found that AI-assisted EGD improved any upper GI neoplasm detection overall (RR 1.57, 95% CI 1.23–2.01).[2] While there was significantly increased detection for lesions ≤10 mm (RR, 2.24; 95% CI, 1.72-2.91), detection rate for lesions >10

mm was similar. Another recent meta-analysis of 26 studies (43,088 patients) found that AI-assisted EGD achieved a sensitivity of 90%, specificity of 92%, and AUC of 0.96 for early gastric cancer detection, significantly outperforming clinician-level performance (AUC 0.85–0.90).[3] Existing literature seems to suggest a potential benefit of AI-assisted EGD for upper GI neoplasm detection. However, the significant discrepancy seen in the impact of AI-assisted EGD using central pathology review versus not in this trial does call into question whether pathology review may significantly impact any detected differences seen with AI assistance. Most AI studies on this question have not required centralized pathology review, and this study importantly raises the question of whether this should be included for training AI models moving forward.

This trial also offers additional support to potential use of AI-assisted EGD to improve endoscopic quality. The miss rate for early gastric cancer is around 10%, and higher among inexperienced endoscopists, in part due to subtle features of early lesions.[4] One mechanism for this is likely the reduction in blind spots seen with AI. This trial builds on prior evidence studying the ENDOANGEL system that showed that the AI group has fewer blind spots compared to conventional EGD (mean 5.4 vs 9.8, P <0.001).[5] A prior RCT using the ENDOANGEL system at a single-center tandem trial of 1,812 patients found that AI significantly reduced the miss rate of gastric neoplasms (RR 0.224, 95% CI 0.068–0.744), with particular benefit for lesions ≤10 mm.[6]

Key Study Findings

AI-assisted EGD did not significantly improve gastric neoplasm detection after centralized pathologic review. Of note, the original diagnosis of 84% (804 of 962) low-grade intraepithelial neoplasia and 14% (7 of 50) HGIN were downgraded to benign lesions after the centralized pathology review, while only 0.72% (6/837) of non-neoplastic lesions were upgraded to neoplasms. This accounted for a substantial difference in the rate of gastric neoplasms that was reported for results based on preliminary review versus centralized pathology review. Interestingly, AI did improve detection based on original pathology.

Detection of EGC or IM/gastric atrophy was similar with vs without AI assistance based on both preliminary and centralized review pathology. AI significantly reduced endoscopic blind spots (from 2.52 to 1.07, P <0.001), supporting its value as a quality assurance tool for mucosal inspection completeness, consistent with earlier ENDOANGEL trials. The subgroup analysis showing greater benefit of the AI system among less experienced endoscopists and more fatigued operators suggests AI may be most impactful in settings where baseline endoscopy quality may be suboptimal.

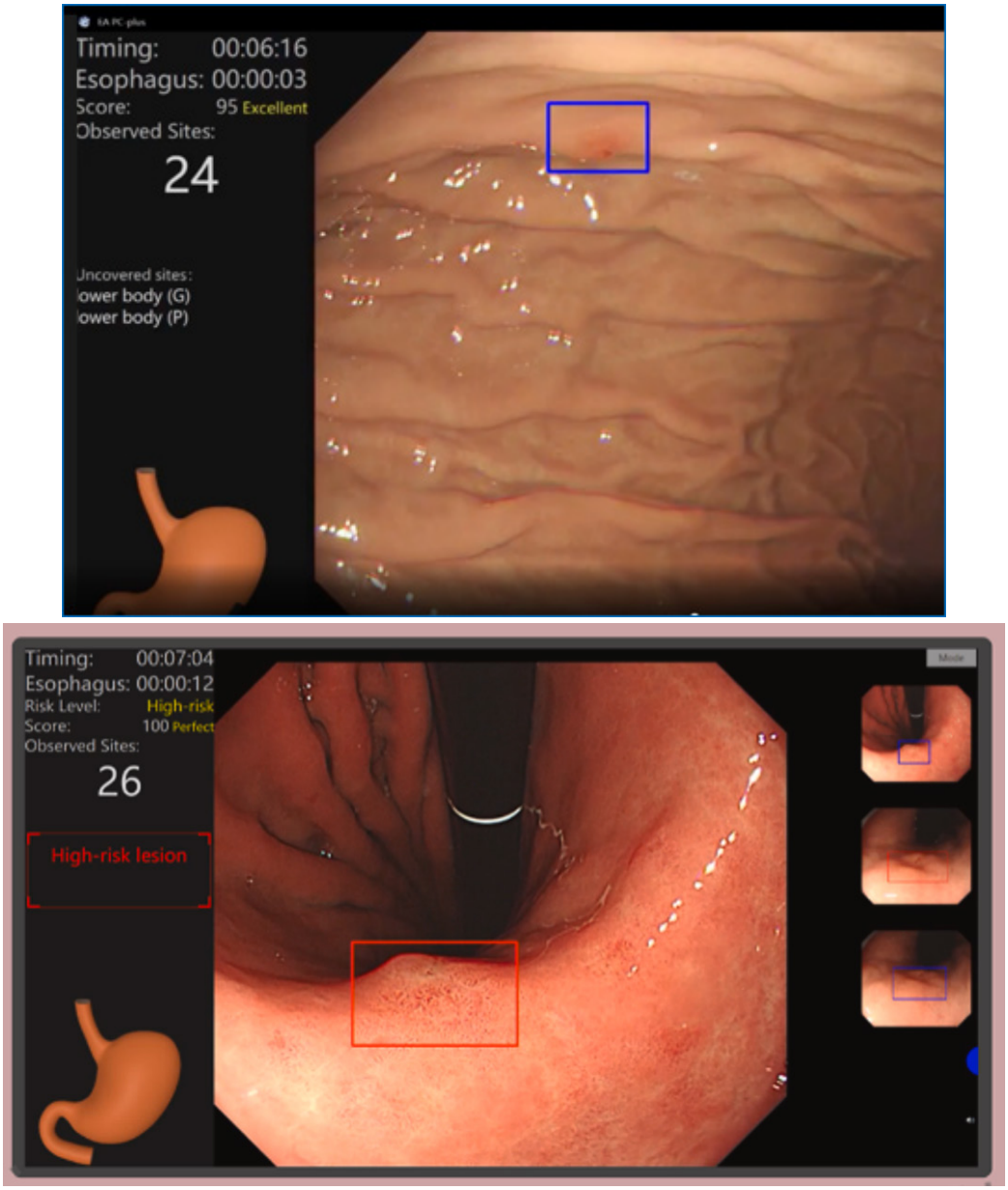

Figure 1. Top: Blue bounding box demarcates a low-risk lesion detected by the ENDOANGEL-GN system. Bottom: Red bounding box demarcates a high-risk lesion detected by the ENDOANGEL-GN system.

Caution

This trial was conducted in China, a region with high gastric cancer prevalence and where endoscopists are generally more experienced in gastric evaluation, and the majority of procedures were performed at tertiary hospitals or by senior endoscopists. This may limit generalizability to lower-prevalence settings or endoscopists with less experience with diagnosis of gastric neoplasms. The study authors also highlight the high quality of EGD performed in the control group in this study, with mean procedure time in the control group of 7.3 minutes and only 2.5 blind spots in the control group. The authors explored neoplasm detection rates among AI-exposed vs AI-naïve endoscopists and found higher detection rates among AI-exposed endoscopists, which may suggest a possible training

bias among those with AI exposure. Further study of downstream effects of AI exposure should be conducted to explore this.

My Practice

This trial tempers enthusiasm for immediate widespread adoption of AI-assisted EGD as a tool to improve gastric neoplasm detection rates in high-volume, experienced endoscopy settings. However, the significant reduction in blind spots supports the use of AI as a quality monitoring tool during EGD. In practice, AI may be most valuable for improving gastric neoplasm detection in lower-volume settings with endoscopists less experienced with neoplasm detection, when operator fatigue may be a concern, and as an endoscopic training tool.

Future Research

Future trials should evaluate AI performance in less experienced endoscopy units to evaluate use in more generalizable real-world settings, as well as stratified b y endoscopist experience level. Future studies should also incorporate advanced imaging modalities, like narrow-band imaging and linked color imaging, with AI.

Conflict of Interest

Dr. Zhou reports no relevant conflicts of interest.

REFERENCES

1. Dong Z, Wu L, Du H, et al. Effect of a computer-aided device for detecting gastric neoplasms: A multicenter, randomized controlled trial. Gastroenterology. 2026; 170(7): 1518-1532.
2. Al Hayek M, Barberio B, Al Hayek O, et al. Efficacy of artificial intelligence-assisted upper gastrointestinal endoscopy for neoplasm detection: A systematic review and meta-analysis of randomized controlled trials. Gastrointest Endosc. 2026; 104(1): 6-18.
3. Lv M, Chen F, Li Q, Xue M, Wang J. Comparative diagnostic accuracy of different artificial intelligence models for early gastric cancer: a systematic review and meta-analysis. Front Oncol. 2025; 15.
4. Pimenta-Melo AR, Monteiro-Soares M, Libânio D,et al. Missing rate for gastric cancer during upper gastrointestinal endoscopy: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol. 2016; 28(9): 1041-9.
5. Wu L, He X, Liu M, et al. Evaluation of the effects of an artificial intelligence system on endoscopy quality and preliminary testing of its performance in detecting early gastric cancer: a randomized controlled trial. Endoscopy. 2021; 53(12): 1199-1207.
6. Wu L, Shang R, Sharma P, et al. Effect of a deep learning-based system on the miss rate of gastric neoplasms during upper gastrointestinal endoscopy: a single-centre, tandem, randomised controlled trial. Lancet Gastroenterol Hepatol. 2021; 6(9): 700-708.