

# Artificial Intelligence as a Decision-Support Tool to Enhance Workflow Efficiency of Routine Esophagogastroduodenoscopy: A Randomized Crossover Trial

**Yusuke Suzuki**

Kyushu University

**Yosuke Minoda**

`minoda.yosuke.633@m.kyushu-u.ac.jp`

Kyushu University

**Tsubasa Takeshima**

Kyushu University

**Kazuki Inamura**

Kyushu University

**Masafumi Wada**

Kyushu University

**Yoshitaka Hata**

International University of Health and Welfare School of Medicine

**Yoshimasa Tanaka**

Kyushu University

**Takatoshi Chinen**

Kyushu University

**Eikichi Ihara**

International University of Health and Welfare School of Medicine

**Haruei Ogino**

Kyushu University

**Yoshihiro Ogawa**

Kyushu University

---

## Article

**Keywords:** Artificial Intelligence, Endoscopy, Gastrointestinal, Stomach Neoplasms, Early Detection of Cancer, Diagnosis, Computer-Assisted

**Posted Date:** July 13th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-9840490/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** Competing interest reported. T. Chinen is affiliated with an endowed course supported by Miyarisan Pharmaceutical Co., Ltd.; Fujifilm Medical Co., Ltd.; Terumo Corporation; FANCL Corporation; Seiwa Medical Corporation; Hakuai Hospital; Japanese Society of Preventive Medicine; TFC Co., Ltd.; and Iharakasuga Clinic. The other authors declare that they have no conflict of interest.

---

# Abstract

Artificial intelligence (AI) has recently emerged as a tool to support lesion detection and characterization, but its effect on the efficiency of routine esophagogastroduodenoscopy (EGD) remains unclear. We conducted a prospective, randomized, two-period crossover trial at Kyushu University Hospital to evaluate whether AI-assisted endoscopy could reduce intragastric observation time compared with conventional magnifying endoscopy. 36 patients with upper gastrointestinal symptoms underwent two consecutive systematic intragastric observations, one with AI-assisted endoscopy and one with conventional magnifying endoscopy, in randomized order during a single sedated EGD. The primary endpoint was intragastric observation time per observation. Secondary endpoints included diagnostic performance and differences in observation time by endoscopist expertise. Mean observation time was significantly shorter with AI-assisted endoscopy than with conventional magnifying endoscopy ( $178.6 \pm 52.6$  vs.  $242.9 \pm 65.8$  seconds), with a mean difference of  $-64.4$  seconds (95% CI,  $-92.4$  to  $-36.3$ ;  $P < 0.001$ ), corresponding to a 27% reduction. The time-saving effect was observed in both expert and non-expert endoscopists. Diagnostic performance was assessed in 42 areas of interest from 32 patients who underwent follow-up endoscopy. AI achieved 100% sensitivity and 43.6% specificity. No adverse events were observed. AI-assisted endoscopy may improve the efficiency of intragastric observation during routine EGD.

## Introduction

Gastric cancer remains a major cause of cancer-related mortality worldwide, and esophagogastroduodenoscopy (EGD) is the gold standard for early detection [1–3]. Early detection enables curative resection and is associated with an excellent prognosis, with reported 5-year overall and disease-specific survival rates of approximately 90% and 99.4%, respectively, for stage IA gastric cancer [4].

Early-stage gastric neoplasms are often overlooked because they exhibit only minimal changes in color or morphology, making high-quality endoscopy essential for their detection [5, 6]. In routine EGD, suspicious lesions identified during screening observation typically undergo further optical evaluation before biopsy. Among currently available diagnostic approaches, magnifying endoscopy with assessment of microvascular and microsurface patterns based on the vessel plus surface (VS) classification system, provides the highest diagnostic accuracy [7–9]. However, this approach is technically demanding, requires substantial training to achieve proficiency, and depends on specialized magnifying endoscopes, which may limit its widespread adoption in routine EGD.

Artificial intelligence (AI) has recently emerged as a tool to support lesion detection and characterization [10–12]. Although AI has shown promising diagnostic performance, its impact on procedural efficiency in real-world routine EGD has not been well established. Specifically, its impact on the time required for evaluating suspicious areas and the overall intragastric observation time has not been well characterized in routine clinical settings.

To address this question, we conducted a prospective, randomized crossover trial comparing AI-assisted endoscopy with conventional magnifying endoscopy during EGD. Using a crossover design to minimize inter-patient anatomical variability, we aimed to quantify the impact of AI assistance on intragastric observation time (primary endpoint) and to assess diagnostic performance and the influence of endoscopist expertise (secondary endpoints).

## Results

### Patient and Procedural Characteristics

A Flow diagram of enrolled patients is shown in Fig. 1. During the study period, 38 patients were assessed for eligibility, of whom 2 declined to provide informed consent. The remaining 36 patients were enrolled and randomized. All 36 patients completed both AI-assisted and conventional endoscopy and were included in the primary analysis of intragastric observation time. Among them, 4 patients were lost to follow-up and were excluded from the diagnostic performance analysis. Therefore, 32 patients were included in the diagnostic performance analysis. Baseline characteristics are summarized in Table 1. The median age was 69.0 years (interquartile range [IQR], 61.0–74.8), and 23 patients were men (63.9%). Patients were equally distributed according to endoscopist expertise, endoscopy equipment, and observation sequence, with 18 patients in each category. The trial ended after the prespecified target sample size was reached and was not stopped early.

Table 1  
Baseline characteristics of the study population. IQR, interquartile range. AI, artificial intelligence.

| Characteristic                 | All Patients (N = 36) |
|--------------------------------|-----------------------|
| Age                            | 69.0 (61.0–74.8)      |
| Median (IQR)                   |                       |
| Sex, n (%)                     |                       |
| Male                           | 23 (63.9)             |
| Female                         | 13 (36.1)             |
| Endoscopist expertise, n (%)   |                       |
| Expert                         | 18 (50.0)             |
| Non-expert                     | 18 (50.0)             |
| Equipment, n (%)               |                       |
| Fujifilm system                | 18 (50.0)             |
| Olympus system                 | 18 (50.0)             |
| Observation sequence, n (%)    |                       |
| AI-assisted first              | 18 (50.0)             |
| Conventional first             | 18 (50.0)             |
| Gastric mucosal atrophy, n (%) |                       |
| None                           | 9 (25.0)              |
| Closed type                    | 17 (47.2)             |
| Open type                      | 10 (27.8)             |
| No. of detected AOI            |                       |
| Median (IQR)                   | 1 (1–2)               |

## AOI Characteristics and Histological Findings

During the procedures, 49 Areas of Interest (AOIs) in the 36 patients were identified (Table 2). Lesions were most frequently located in the lower third (49.0%) or middle third (42.9%) of the stomach, and depressed morphology was the most common (57.1%). The median AOI size was 5 mm (IQR 4–10). The AI system classified 30 of 49 AOIs (61.2%) as “Consider biopsy,” whereas conventional magnifying endoscopy identified suspected neoplastic findings in 6 AOIs (12.2%).

Table 2  
Characteristics of Areas of Interest (AOIs) and histological findings.

| Characteristic                             | Total AOIs (N = 49) |
|--------------------------------------------|---------------------|
| Location, n (%)                            |                     |
| Upper third                                | 4 (8.2)             |
| Middle third                               | 21 (42.9)           |
| Lower third                                | 24 (49.0)           |
| Morphology, n (%)                          |                     |
| Elevated                                   | 11 (22.4)           |
| Flat                                       | 10 (20.4)           |
| Depressed                                  | 28 (57.1)           |
| Lesion size, mm                            | 5 (4–10)            |
| Median (IQR)                               |                     |
| AI classification, n (%)                   |                     |
| Consider biopsy                            | 30 (61.2)           |
| Low confidence                             | 19 (38.8)           |
| Magnifying endoscopy classification, n (%) |                     |
| Neoplasm-suspicious                        | 6 (12.2)            |
| Non-neoplastic                             | 43 (87.8)           |
| Biopsy performed, n (%)                    | 12 (24.5)           |
| Histopathological diagnosis, n             |                     |
| Intestinal metaplasia                      | 6                   |
| Gastric cancer                             | 2                   |
| Neuroendocrine tumor                       | 1                   |
| Hyperplastic lesion                        | 1                   |
| Inflammation                               | 2                   |

Histopathology results included inflammation (n = 2), intestinal metaplasia (n = 6), a hyperplastic lesion (n = 1), and 3 neoplastic lesions, comprising 2 gastric cancers and 1 neuroendocrine tumor (NET). The NET case was not definitively diagnosed at the initial biopsy but was confirmed by repeat biopsy of the same AOI during follow-up EGD.

# Primary and Secondary Outcome: Intra gastric Observation Time

Intra gastric observation time is summarized in Fig. 3, Tables 3 and 4. Mean observation time was significantly shorter with AI-assisted endoscopy than with conventional observation ( $178.6 \pm 52.6$  vs.  $242.9 \pm 65.8$  seconds). The mean difference was  $-64.4$  seconds (95% confidence interval [CI],  $-92.4$  to  $-36.3$ ;  $P < 0.001$ ), corresponding to a 27% reduction.

Table 3  
Comparison of intra gastric observation time between AI-assisted and conventional endoscopy using Student's t-test. CI, confidence interval. SD, standard deviation.

| Total                 | Observation Time (s) | Mean Difference (s)       | P Value |
|-----------------------|----------------------|---------------------------|---------|
|                       | Mean $\pm$ SD        | Difference (95%CI)        |         |
| Conventional (n = 36) | 242.9 $\pm$ 65.8     |                           |         |
| AI-assisted (n = 36)  | 178.6 $\pm$ 52.6     | -64.4 (- 92.4 to - 36.3)  | < 0.001 |
| First observation     |                      |                           |         |
| Conventional (n = 18) | 244.4 $\pm$ 66.0     |                           |         |
| AI-assisted (n = 18)  | 199.6 $\pm$ 62.8     | -44.8 (- 88.4 to - 1.1)   | 0.045   |
| Second observation    |                      |                           |         |
| Conventional (n = 18) | 241.4 $\pm$ 67.5     |                           |         |
| AI-assisted (n = 18)  | 157.5 $\pm$ 28.4     | -83.9 (- 119.7 to - 48.2) | < 0.001 |
| Experts               |                      |                           |         |
| Conventional (n = 18) | 254.9 $\pm$ 75.3     |                           |         |
| AI-assisted (n = 18)  | 194.8 $\pm$ 64.5     | -60.1(- 107.6 to - 12.6)  | 0.015   |
| Non-Experts           |                      |                           |         |
| Conventional (n = 18) | 230.9 $\pm$ 54.3     |                           |         |
| AI-assisted (n = 18)  | 162.3 $\pm$ 31.2     | -68.6 (- 98.6 to - 38.6)  | < 0.001 |

Table 4

Linear mixed-effects model analysis of factors associated with intragastric observation time. The model included patients and endoscopists as random intercepts and fixed effects for treatment, period, sequence, treatment-by-expertise interaction, age, sex, expertise, equipment, mucosal atrophy, and number of AOlS. CI, confidence interval. SE, standard error.

| Variable                             | Estimate (s) | SE   | 95% CI         | P Value |
|--------------------------------------|--------------|------|----------------|---------|
| Treatment Effect                     |              |      |                |         |
| AI-assisted vs. conventional         | -60.1        | 13.5 | -87.7 to -32.6 | < 0.001 |
| Period Effect                        |              |      |                |         |
| First vs. second observation         | 22.5         | 9.6  | 3.0 to 42.0    | 0.025   |
| Sequence Effect                      |              |      |                |         |
| AI-first vs. conventional-first      | 16.6         | 16.9 | -18.1 to 51.2  | 0.34    |
| Treatment × Expertise Interaction    | -8.5         | 19.2 | -47.5 to 30.5  | 0.66    |
| Covariate                            |              |      |                |         |
| Age (per 1-year increase)            | 0.2          | 0.8  | -1.4 to 1.7    | 0.81    |
| Sex (Male vs. Female)                | -18.2        | 18.3 | -55.8 to 19.4  | 0.33    |
| Expertise (non-Expert vs. Expert)    | -26.3        | 27.8 | -95.0 to 42.4  | 0.38    |
| Equipment (Olympus vs. Fujifilm)     | -8.8         | 17.4 | -44.4 to 26.8  | 0.62    |
| Mucosal atrophy (Present vs. Absent) | 8.9          | 18.4 | -29.0 to 46.8  | 0.63    |
| Number of AOlS (per lesion)          | 17.6         | 7.8  | 1.5 to 33.6    | 0.033   |

In the linear mixed-effects model with random intercepts for patients and endoscopists (Table 4), AI-assisted endoscopy was associated with a significantly shorter intragastric observation time than conventional observation (-60.1 s, AI minus conventional; 95% CI, -87.7 to -32.6;  $P < 0.001$ ). A significant period effect was observed, with the first observation taking longer than the second (+22.5 s; 95% CI, 3.0 to 42.0;  $P = 0.025$ ). However, AI-assisted endoscopy was shorter than conventional observation in both periods (Table 3). The sequence effect was not significant (AI-first vs. conventional-first: +16.6 s; 95% CI, -18.1 to 51.2;  $P = 0.34$ ). The number of AOlS was independently associated with longer observation time (+17.6 s, per AOl; 95% CI, 1.5 to 33.6;  $P = 0.033$ ).

Subgroup analyses showed that observation time decreased with AI assistance, regardless of expertise (Table 3). The mean difference (AI minus conventional) was -60.1 seconds (95% CI, -107.6 to -12.6;  $P = 0.015$ ) for experts and -68.6 seconds (95% CI, -98.6 to -38.6;  $P < 0.001$ ) for non-experts. The treatment-by-expertise interaction was not statistically significant ( $P = 0.66$ ), indicating that the time-saving effect did not differ by expertise (Table 4).

# Diagnostic Performance

Diagnostic performance was evaluated for 42 AOIs in the 32 patients who underwent follow-up EGD within 1 year (Table 5). This cohort consisted of 3 neoplastic lesions (2 gastric cancers and 1 NET) and 39 non-neoplastic lesions. No additional neoplastic lesions were detected during follow-up.

Table 5  
Diagnostic performance of AI-assisted endoscopy versus conventional magnifying endoscopy for gastric neoplasia. Analysis was performed on 42 AOIs in 32 patients who underwent follow-up EGD within 1 year. PPV, positive predictive value. NPV, negative predictive value.

| Diagnostic Metric | AI system, % (95% CI) | Magnifying Endoscopy, % (95% CI) |
|-------------------|-----------------------|----------------------------------|
| Sensitivity       | 100 (29.2–100)        | 66.7 (9.4–99.2)                  |
| Specificity       | 43.6 (27.8–60.4)      | 89.7 (75.8–97.1)                 |
| PPV               | 12.0 (2.5–31.2)       | 33.3 (4.3–77.7)                  |
| NPV               | 100 (80.5–100)        | 97.2 (85.5–99.9)                 |
| Accuracy          | 47.6 (32.0–63.6)      | 88.1 (77.4–96.0)                 |

The AI system correctly identified all 3 neoplastic lesions as “Consider biopsy,” yielding a sensitivity of 100% (95% CI, 29.2–100). Magnifying endoscopy identified 2 of the 3 neoplastic lesions, yielding a sensitivity of 66.7% (95% CI, 9.4–99.2). Specificity was 43.6% (95% CI, 27.8–60.4) for the AI system and 89.7% (95% CI, 75.8–97.1) for magnifying endoscopy. The positive predictive value (PPV) was 12.0% (95% CI, 2.5–31.2) for AI and 33.3% (95% CI, 4.3–77.7) for magnifying endoscopy, while the negative predictive value (NPV) was 100% (95% CI, 80.5–100) and 97.2% (95% CI, 85.5–99.9), respectively. Accuracy was 47.6% (95% CI, 32.0–63.6) and 88.1% (95% CI, 77.4–96.0), respectively.

One neoplastic lesion classified as non-neoplastic by magnifying endoscopy was difficult to assess in detail because of White Opaque Substance (WOS) (Fig. 2a-c). In contrast, the AI system classified the lesion as “Consider biopsy”. Discrepancies between AI and magnifying endoscopy were observed in 20 of 42 AOIs (47.6%). All were categorized as “AI: Consider biopsy / magnifying endoscopy: non-neoplastic,” and there were no cases of “AI: Low confidence / magnifying endoscopy: neoplastic”. No clinical or endoscopic factor was significantly associated with discrepancy, although discrepancies were more frequent in open-type atrophy (8/13, 61.5%) and depressed lesions (13/22, 59.1%).

## Adverse Events

No adverse events related to endoscopy or sedation were observed.

## Discussion

This randomized crossover trial demonstrated that AI-assisted endoscopy significantly reduced intragastric observation time compared with conventional observation, with a mean reduction of approximately 64 seconds and a relative reduction of 27%. This benefit was achieved without compromising sensitivity for neoplastic lesions in this cohort. These findings suggest that AI-assisted endoscopy may improve the efficiency of routine EGD while maintaining the ability to detect neoplasms.

In this study, AI was used as an alternative decision-support tool for evaluating AOs within the EGD workflow. The reason we selected magnifying endoscopy as the control group was that it was the most accurate modality for evaluating optical characteristics and represents the upper limit of diagnostic performance. By comparing AI with magnifying endoscopy, this study aimed to determine whether AI could streamline the evaluation process while maintaining diagnostic accuracy.

The reduction in examination time with AI-assisted endoscopy is likely attributable to differences in the evaluation process for AOs. In conventional magnifying endoscopy, assessing a lesion requires time-consuming several steps, such as approaching the lesion, stabilizing the scope, and magnifying the lesion and the surrounding mucosa with detailed observation. In contrast, the AI system provides a result ("Consider biopsy" or "Low confidence") within a few seconds after the endoscopist freezes the image. This difference in evaluation workflow likely contributed to the overall reduction in intragastric observation time. Importantly, the time-saving effect was consistent across expert and non-expert endoscopists, with no significant treatment-by-expertise interaction. Although AI is often considered a support tool for novices, our findings suggest that AI has the potential to improve workflow for expert endoscopists as well. Because this time-saving effect may benefit endoscopists with different levels of expertise, accumulated time savings could improve the overall efficiency of endoscopy units in routine practice.

Regarding diagnostic performance, AI showed high sensitivity and high NPV, whereas its specificity and PPV were lower than those of magnifying endoscopy. The lower specificity of AI was mainly due to false-positive outputs, in which benign lesions such as intestinal metaplasia and inflammation were classified as "Consider biopsy." (Fig. 2d–f). Conversely, one neoplastic lesion that was difficult to assess by magnifying endoscopy because of WOS was classified by AI as "Consider biopsy." This finding suggests that AI may have a complementary role in lesions for which detailed magnifying assessment is limited. Although improving specificity will be required for the development of AI, sensitivity is particularly important in screening endoscopy to minimize missed lesions. From this perspective, this AI system appeared to provide acceptable safety.

Evidence on AI in upper gastrointestinal endoscopy remains limited. Most previous studies have focused on lesion detection, miss rates, diagnostic performance, or quality control, whereas its practical impact on workflow efficiency in routine clinical practice remains unclear [10, 13–15]. This study addresses this gap by demonstrating the impact of AI-assisted endoscopy on procedural efficiency in routine EGD. Magnifying endoscopy with the VS classification remains one of the most reliable methods for evaluating gastric neoplasms [7, 8, 16]. However, it requires substantial training and specialized

equipment, and diagnostic performance may vary among endoscopists [17–19]. AI may provide more standardized lesion evaluation without extending examination time. This may be particularly useful where access to magnifying endoscopy is limited, or where endoscopists have limited training or experience with gastric neoplasia. However, the specificity of AI remains lower than that of magnifying endoscopy. Therefore, AI cannot fully replace magnifying endoscopy at present, but it has the potential to reshape the diagnostic workflow by providing rapid, standardized decision support during routine EGD.

The strength of this study is its randomized crossover design, which minimized anatomical variability between patients and enabled direct comparisons within the same patient. The findings were further supported by a linear mixed-effects model incorporating random intercepts for patients and endoscopists and adjusting for prespecified covariates.

Our study has several limitations. First, this was a single-center study, and the sample size was small, which may limit generalizability. Second, only three neoplastic lesions were identified, resulting in wide confidence intervals for diagnostic performance. Larger multicenter studies are needed to assess diagnostic performance more precisely. Third, blinding was not feasible because of the intervention, and performance bias cannot be excluded, although randomization of the observation order may have reduced this concern.

In conclusion, AI-assisted endoscopy significantly shortened intragastric observation time compared with conventional magnifying endoscopy, without compromising neoplasia detection. AI-assisted endoscopy may improve the efficiency of the EGD workflow in routine practice.

## Methods

### Study design and ethics

We conducted a prospective, randomized, open-label, two-period crossover trial at Kyushu University Hospital comparing AI-assisted endoscopy with conventional observation (Fig. 1). Each participant underwent a single sedated EGD, during which two consecutive systematic intragastric inspections were performed, one with AI-assisted and one with conventional magnifying endoscopy, in a randomized order. This randomized controlled trial was designed and reported in accordance with the CONSORT statement for randomized crossover trials.

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board of Kyushu University Hospital (approval number: 23020). Written informed consent was obtained from all participants before enrollment. The trial was registered in the University Hospital Medical Information Network Clinical Trials Registry (UMIN-CTR; registration number: UMIN000054466; date of registration: May 24, 2024).

### Participants

Between July 2024 and October 2024, 38 consecutive patients were assessed for eligibility in this study. After exclusion of 2 patients (Decline to participate n = 2), 36 patients were enrolled and included in the primary analysis of intragastric observation time. Among these, 32 patients who underwent follow-up EGD within 1 year were included in the analysis of diagnostic performance. The inclusion and exclusion criteria were as follows;

Inclusion criteria: (1) patients aged  $\geq 20$  years; (2) presenting with upper gastrointestinal symptoms suggestive of potential upper gastrointestinal neoplasia and scheduled for sedated EGD; and (3) provided written informed consent.

Exclusion criteria: (1) known poor tolerance to sedation during previous endoscopy; (2) obstructive sleep apnea, heart failure, or unstable hemodynamics at risk for sedation; (3) contraindications for endoscopy (ileus, perforation, severe respiratory or cardiovascular disease); and (4) other severe comorbidities deemed inappropriate by the investigators.

## Endoscopists

All examinations were performed by gastroenterologists with experience of more than 1,000 EGDs. For analysis, participating endoscopists were classified according to expertise. Experts were defined as endoscopists who held the specialist certification from the Japanese Gastroenterological Endoscopy Society (JGES) and had performed more than 10,000 EGD procedures, whereas non-experts were those who did not meet these criteria. A total of 7 endoscopists participated in this study (4 experts and 3 non-experts).

## Randomization

A computer-generated randomization list was created by the data manager (Yoshitaka Hata). Randomization was stratified according to endoscopist expertise (expert vs. non-expert), and separate permuted-block schedules (block size of two) were prepared for each stratum. Patients were randomly assigned in a 1:1 ratio to either the AI-first or conventional-first sequence. The allocation sequence was concealed from the enrolling investigators until patient enrollment. Because of the nature of the endoscopic intervention, endoscopists were not blinded to the observation sequence.

## Endoscopic Procedures and Observation Workflow

All examinations were performed in accordance with the systematic screening protocol for the stomach (SSS), a standardized photo-documentation method widely adopted in Japan to ensure comprehensive visualization of the entire gastric mucosa [20]. The degree of gastric mucosal atrophy was assessed according to the Kimura-Takemoto classification [21] and categorized into three groups: none, closed type, and open type.

Systematic intragastric observation started after adequate visualization was obtained through mucosal cleaning. AOIs were assessed at the time of identification. Within each patient, the same AOIs were assessed by both AI-assisted and magnifying endoscopy. During AI-assisted endoscopy, AOIs were

evaluated by the AI system using frozen images. During conventional observation, AOIs were evaluated by magnifying endoscopy according to the vessel plus surface (VS) classification system. Biopsy sampling was performed after completion of both observations. The decision to perform a biopsy was made by the endoscopist based on an overall assessment of conventional white-light imaging findings and magnifying endoscopy findings. Intra-gastric observation time was calculated from photo timestamps as the interval from the first gastric image to completion of the systematic inspection. Accordingly, observation time did not include mucosal cleaning and biopsy sampling time. Timestamps were recorded automatically by the recording device.

## Artificial Intelligence System

For AI-assisted endoscopy, we used the commercially available gastroAI-model G (AI Medical Service Inc., Tokyo, Japan). This system analyzes endoscopic images to provide real-time feedback categorized as either “Consider biopsy” for suspected neoplastic lesions or “Low confidence” for non-neoplastic changes, such as fundic gland polyps or gastritis. In practice, the endoscopist froze the endoscopic image to trigger the analysis, and the result was displayed within approximately 2–3 seconds. This AI system has been approved based on static image data, with a reported diagnostic performance of 58.4% sensitivity, 87.3% specificity, and a negative predictive value of 96.5% [22].

## Definition of Area of Interest (AOI)

An Area of Interest (AOI) was defined as any mucosal lesion with potential for gastric cancer, including an irregular elevated or depressed area with abnormal color change and/or mucosal surface structures, as well as a flat pale or discolored area.

## Equipment and sedation

EGD was performed using high-definition endoscopy systems and magnifying gastroscopes (GIF-H290Z with EVIS X1, Olympus, Tokyo, Japan; or EG-L600ZW7 with ELUXEO 7000, Fujifilm, Tokyo, Japan). Sedation was provided with midazolam 0.02–0.03 mg/kg, titrated as needed. Oxygen saturation, blood pressure, and heart rate were continuously monitored.

## Outcomes

The primary endpoint was the intra-gastric observation time per observation. Secondary endpoints were: (1) Differences in intra-gastric observation time by endoscopist expertise (expert vs non-expert); (2) Diagnostic performance of the AI system (sensitivity, specificity, positive predictive value, negative predictive value, and accuracy); (3) Adverse events related to endoscopy or sedation.

The evaluation of diagnostic performance was restricted to the 32 patients who underwent follow-up EGD within 1 year. AOIs were defined as negative (non-neoplastic) if gastric neoplasm was not identified during the follow-up endoscopy.

## Histopathology

Biopsy specimens were processed per routine institutional protocols and evaluated by hospital pathologists who were blinded to the study arm. Final diagnoses were extracted from medical records. In cases where a repeat biopsy of the same AOI was performed during follow-up, the histological diagnosis from the follow-up examination was used as the final diagnosis.

## Sample-size calculation

Sample size was estimated based on pilot data demonstrating a mean reduction of approximately 75 seconds in intragastric observation time with AI assistance, with an estimated standard deviation of 106 seconds. To allow a margin of conservatism in detecting a clinically meaningful between-condition difference, a moderate effect size (Cohen's  $d = 0.50$ ) was prespecified for the paired comparison. Using a two-sided paired  $t$ -test with  $\alpha = 0.05$  and a power of 0.80, 34 patients were required. Allowing for potential dropout, the target enrollment was set at 36.

## Statistical Analysis

The primary endpoint was analyzed using Student's  $t$ -test and a linear mixed-effects model with fixed effects for treatment, period, and sequence, and random intercepts for patients and endoscopists. The model was adjusted for prespecified covariates: age, sex, expertise, equipment, mucosal atrophy, AOI count, and treatment-by-expertise interaction. Diagnostic performance was calculated with exact 95% confidence intervals (Clopper–Pearson). Analyses were performed using JMP Pro version 18, with two-sided  $P$  values  $< 0.05$  considered significant.

## Data availability

The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

### Author contributions

Y.S., Y.M., T.T., K.I., M.W.: study conception and design, perform endoscopy, drafting the article, analysis and interpretation of the data. Y.H.: data manager, never perform endoscopy. Y.T., T.C., E.I., H.O., Y.O.: critical revision of the article for important intellectual content. All authors approved the final article.

### Additional Information

### Funding

This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

### Competing interests

T. Chinen is affiliated with an endowed course supported by Miyarisan Pharmaceutical Co., Ltd.; Fujifilm Medical Co., Ltd.; Terumo Corporation; FANCL Corporation; Seiwa Medical Corporation; Hakuai Hospital; Japanese Society of Preventive Medicine; TFC Co., Ltd.; and Iharakasuga Clinic.

The other authors declare that they have no conflict of interest.

## References

1. Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **74**, 229–263, DOI: <https://doi.org/10.3322/caac.21834> (2024).
2. Mabe, K. et al. Endoscopic screening for gastric cancer in Japan: Current status and future perspectives. *Dig. Endosc.* **34**, 412–419, DOI: <https://doi.org/10.1111/den.14063> (2022).
3. Yao, K. et al. Guidelines for endoscopic diagnosis of early gastric cancer. *Dig. Endosc.* **32**, 663–698, DOI: <https://doi.org/10.1111/den.13684> (2020).
4. Kakeji, Y. et al. A retrospective 5-year survival analysis of surgically resected gastric cancer cases from the Japanese Gastric Cancer Association nationwide registry (2001–2013). *Gastric Cancer* **25**, 1082–1093, DOI: <https://doi.org/10.1007/s10120-022-01317-6> (2022).
5. Yashima, K. et al. Current status and future perspective of linked color imaging for gastric cancer screening: a literature review. *J. Gastroenterol.* **58**, 1–13, DOI: <https://doi.org/10.1007/s00535-022-01934-z> (2023).
6. Alexandre, L., Tsilegeridis-Legeris, T. & Lam, S. Clinical and endoscopic characteristics associated with post-endoscopy upper gastrointestinal cancers: a systematic review and meta-analysis. *Gastroenterology* **162**, 1123–1135, DOI: <https://doi.org/10.1053/j.gastro.2021.12.270> (2022).
7. Muto, M. et al. Magnifying endoscopy simple diagnostic algorithm for early gastric cancer (MESDA-G). *Dig. Endosc.* **28**, 379–393, DOI: <https://doi.org/10.1111/den.12638> (2016).
8. Ezoe, Y. et al. Magnifying narrowband imaging is more accurate than conventional white-light imaging in diagnosis of gastric mucosal cancer. *Gastroenterology* **141**, 2017–2025.e3, DOI: <https://doi.org/10.1053/j.gastro.2011.08.007> (2011).
9. Yao, K., Anagnostopoulos, G. K. & Ragunath, K. Magnifying endoscopy for diagnosing and delineating early gastric cancer. *Endoscopy* **41**, 462–467, DOI: <https://doi.org/10.1055/s-0029-1214594> (2009).
10. Wu, L. 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.* **6**, 700–708, DOI: [https://doi.org/10.1016/S2468-1253\(21\)00216-8](https://doi.org/10.1016/S2468-1253(21)00216-8) (2021).
11. Niikura, R. et al. Artificial intelligence versus expert endoscopists for diagnosis of gastric cancer in patients who have undergone upper gastrointestinal endoscopy. *Endoscopy* **54**, 780–784, DOI: <https://doi.org/10.1055/a-1660-6500> (2022).

12. Karsenti, D. et al. Effect of real-time computer-aided detection of colorectal adenoma in routine colonoscopy (COLO-GENIUS): a single-centre randomised controlled trial. *Lancet Gastroenterol. Hepatol.***8**, 726–734, DOI: [https://doi.org/10.1016/S2468-1253\(23\)00104-8](https://doi.org/10.1016/S2468-1253(23)00104-8) (2023).
13. Zhou, R. et al. Efficacy of a real-time intelligent quality-control system for the detection of early upper gastrointestinal neoplasms: a multicentre, single-blinded, randomised controlled trial. *EClinicalMedicine***75**, 102803, DOI: <https://doi.org/10.1016/j.eclinm.2024.102803> (2024).
14. Wu, L. et al. Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy. *Gut***68**, 2161–2169, DOI: <https://doi.org/10.1136/gutjnl-2018-317366> (2019).
15. Ishioka, M. et al. Performance of an artificial intelligence-based diagnostic support tool for early gastric cancers: retrospective study. *Dig. Endosc.***35**, 483–491, DOI: <https://doi.org/10.1111/den.14455> (2023).
16. Nagahama, T. et al. Delineation of the extent of early gastric cancer by magnifying narrow-band imaging and chromoendoscopy: a multicenter randomized controlled trial. *Endoscopy***50**, 566–576, DOI: <https://doi.org/10.1055/s-0044-100790> (2018).
17. Nakanishi, H. et al. Evaluation of an e-learning system for diagnosis of gastric lesions using magnifying narrow-band imaging: a multicenter randomized controlled study. *Endoscopy***49**, 957–967, DOI: <https://doi.org/10.1055/s-0043-111888> (2017).
18. Yoo, C. H. et al. Observer variability in gastric neoplasm assessment using the vessel plus surface classification for magnifying endoscopy with narrow band imaging. *Clin. Endosc.***47**, 74–78, DOI: <https://doi.org/10.5946/ce.2014.47.1.74> (2014).
19. Mabe, K., Yao, K., Nojima, M., Tanuma, T. & Kato, M. An educational intervention to improve the endoscopist's ability to correctly diagnose small gastric lesions using magnifying endoscopy with narrow-band imaging. *Ann. Gastroenterol.***27**, 149–155 (2014).
20. Yao, K. The endoscopic diagnosis of early gastric cancer. *Ann. Gastroenterol.***26**, 11–22 (2013).
21. Kimura, K. & Takemoto, T. An endoscopic recognition of the atrophic border and its significance in chronic gastritis. *Endoscopy***1**, 87–97, DOI: <https://doi.org/10.1055/s-0028-1098086> (1969).
22. Ikenoyama, Y. et al. Detecting early gastric cancer: Comparison between the diagnostic ability of convolutional neural networks and endoscopists. *Dig. Endosc.***33**, 141–150, DOI: <https://doi.org/10.1111/den.13688> (2021).

## Figures

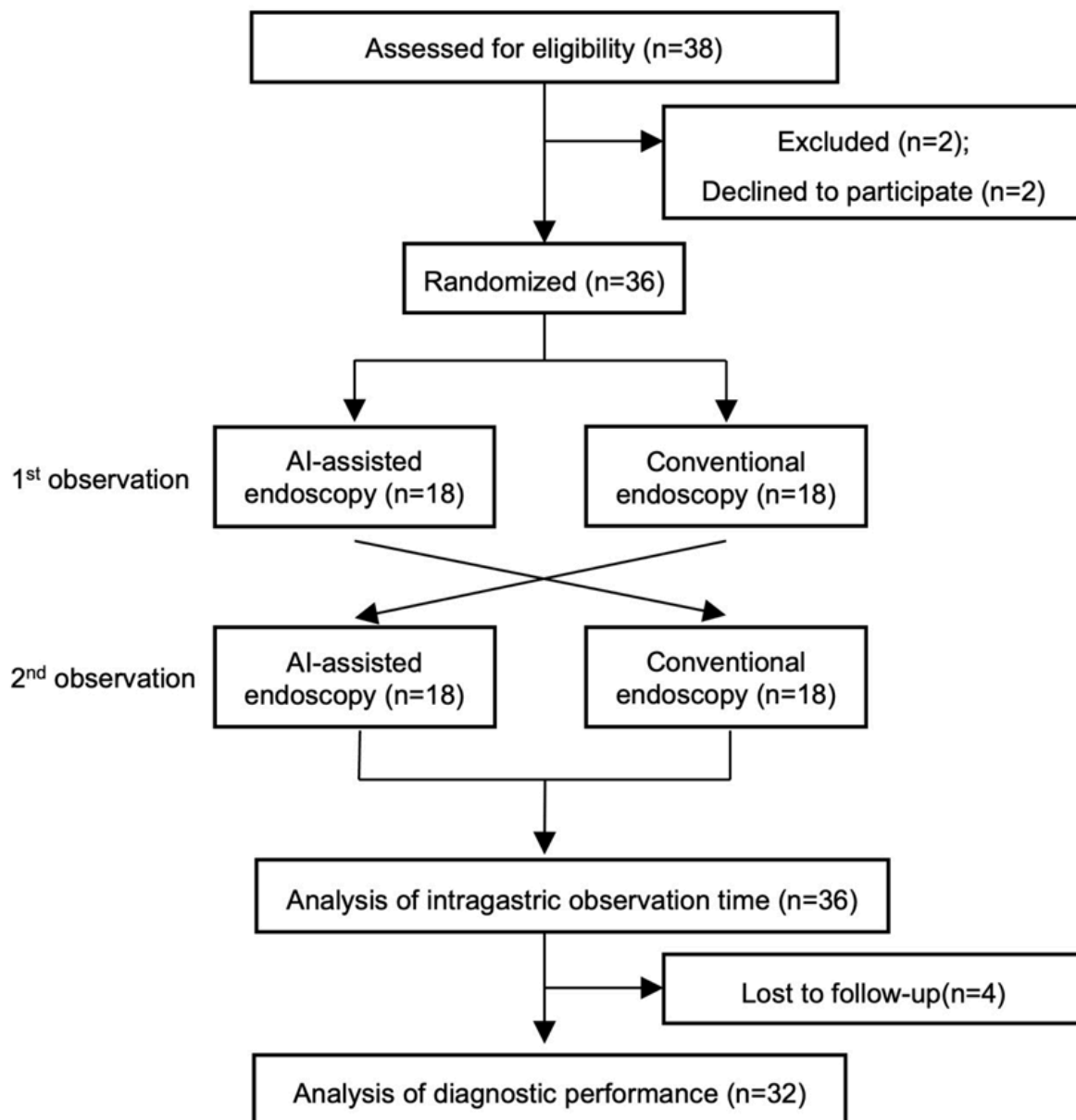


**Figure 1**

Flow diagram of the study

36 patients were randomized to either the AI-first sequence (AI-assisted followed by conventional endoscopy) or the Conventional-first sequence. All 36 patients were included in the primary analysis of intragastric observation time. Diagnostic performance was evaluated in the 32 patients who underwent follow-up EGD within 12 months, excluding 4 patients lost to follow-up.

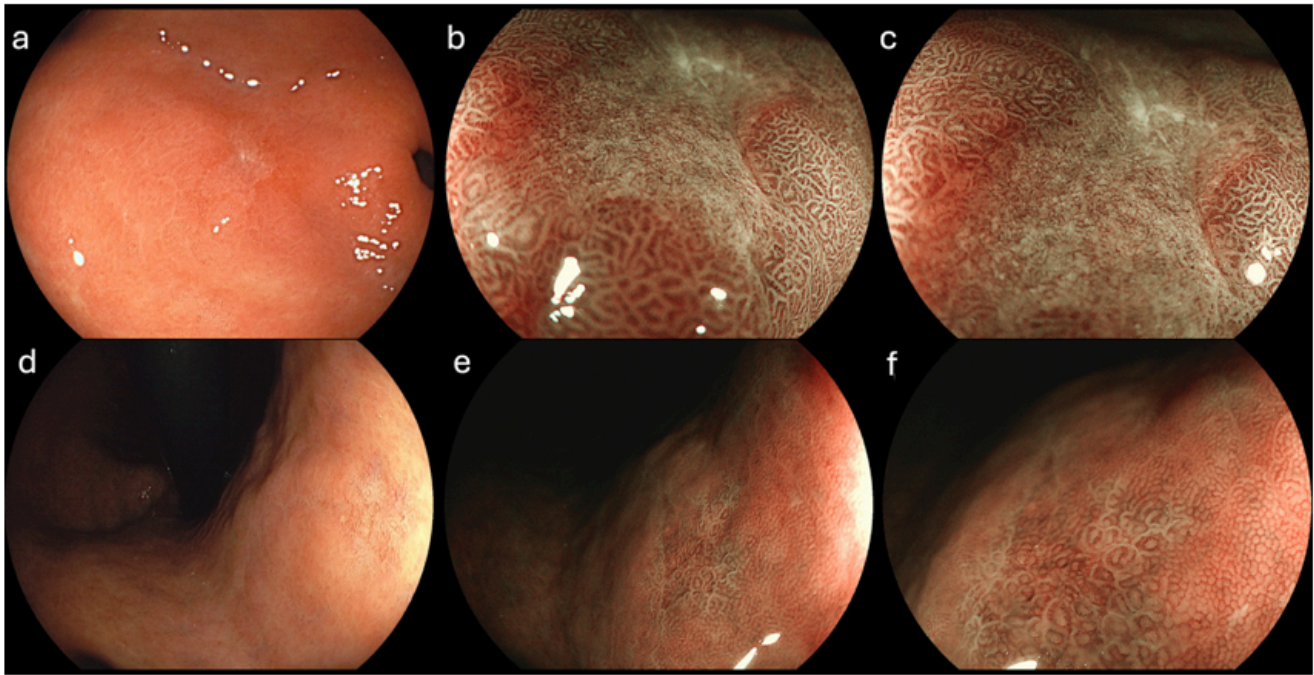


**Figure 2**

Representative endoscopic images of an adenocarcinoma lesion (a–c) and an intestinal metaplasia lesion (d–f) in which AI and magnifying assessment were discordant.

a) White-light imaging (WLI) shows a slightly reddish depressed lesion, approximately 6 mm in diameter, in the gastric antrum. The AI classified the lesion as “Consider biopsy.”

b) Magnifying blue laser imaging (BLI) demonstrates a demarcation line.

c) Detailed evaluation of the lesion was difficult because of the presence of white opaque substance (WOS), and no clear microvascular or microsurface irregularity was identified.

d) WLI shows a slightly reddish depressed lesion, approximately 8 mm in diameter, on the lesser curvature. The AI classified the lesion as “Consider biopsy.”

e, f) Magnifying BLI shows an indistinct demarcation line. The surface pattern and vascular pattern were regular, and the lesion was assessed to be non-neoplastic.

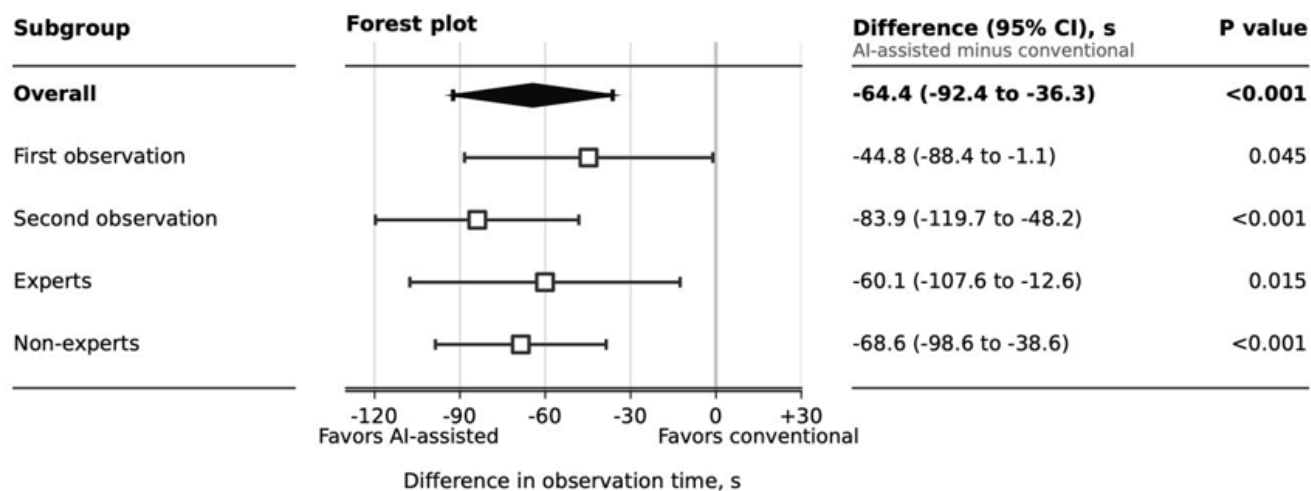


**Figure 3**

Forest plot of Mean differences in intragastric observation time between AI-assisted and conventional magnifying endoscopy using Student’s t-test. Squares indicate mean differences, and horizontal bars indicate 95% confidence intervals.