Variation in the rate of detection of minute and small early gastric cancers at diagnostic endoscopy may reflect the performance of individual endoscopists

BMJ Open Gastroenterology

Variation in the rate of detection of minute and small early gastric cancers at diagnostic endoscopy may reflect the performance of individual endoscopists

Daisuke Murakami ,[1,2,3] Masayuki Yamato,[2] Yuji Amano,[3] Takayoshi Nishino,[1] Makoto Arai[1]

To cite: Murakami D, Yamato M, Amano Y, *et al*. Variation in the rate of detection of minute and small early gastric cancers at diagnostic endoscopy may reflect the performance of individual endoscopists. *BMJ Open Gastroenterol* 2023;**10**:e001143. doi:10.1136/bmjgast-2023-001143

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjgast-2023-001143).

Received 6 March 2023
Accepted 18 June 2023

[1]Department of Gastroenterology, Tokyo Women's Medical University Yachiyo Medical Centre, Yachiyo, Chiba, Japan
[2]Institute of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Shinjuku, Tokyo, Japan
[3]Department of Gastroenterology and Endoscopy, New Tokyo Hospital, Mtsudo, Chiba, Japan

Correspondence to
Dr Daisuke Murakami; daisuke.murakami@gmail.com

ABSTRACT

Objective The documented variation in gastric cancer (GC) detection among endoscopists has often been dismissed as a coincidental artefact of the low incidence of gastric neoplasms; it is not considered associated with differences in physicians' performance of the esophagogastroduodenoscopy procedure. This study is to confirm whether significant variations among endoscopists in early GC detection suggest the individual performance of the upper endoscopy.

Design A retrospective observational study at a single centre in Japan assessed the results of 218 early GCs detected during 25 688 routine esophagogastroduodenoscopies by 12 endoscopists. The main outcome was the rate of early GC detection for each endoscopist under the same circumstances. Other measures included the major diameters and locations of the lesions, *Helicobacter pylori* infection status, and baseline patient characteristics that could affect the prevalence of GC.

Results The early GC detection rates exhibited wide variation among endoscopists (0.09%–2.87%) despite performing routine esophagogastroduodenoscopies in a population with a similar background. Endoscopists were assigned to a low-detection group (n=6; detection rate: 0.47% (range: 0.09%–0.55%)) and a high-detection group (n=5; detection rate: 0.83% (range: 0.63%–1.12%)), with the single highest detector analysed separately due to his distinct detection rate (2.87%). Endoscopists in the high-detection group had better detection rates for minute (major diameter ≤5 mm) and small (major diameter 6–10 mm) GCs than the low-detection group (0.19%/0.23% vs 0.085%/0.098%). These differences were significant (p<0.01), although there were no significant differences in detection of larger tumours (major diameter ≥11 mm; 0.40% vs 0.28%; p=0.13). The tumour location and *H. pylori* status were similar in the low-detection group, high-detection group and for the highest detector.

Conclusion Significant variation in the detection of hard-to-find, smaller GCs may reflect individual performance of the examination.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Detection of early gastric cancers (GCs) requires refined diagnostic techniques from the endoscopists, resulting in interobserver variation in early GC detection.

⇒ Although the differences in early GC detection among individual physicians have been empirically recognised, this has not been considered associated with differences in physicians' performance during the upper endoscopy procedure, but rather attributed to a coincidental artefact of the low incidence of gastric neoplasms.

WHAT THIS STUDY ADDS

⇒ In a retrospective analysis of early GCs detected by 12 endoscopists at a high-volume endoscopy centre, the high-detection group of physicians had significantly better detection rates for smaller early GCs (≤10 mm) than the low-detection group in a similar patient population.

⇒ There were no significant differences in detection rates of easy-to-find, larger early GCs (≥11 mm).

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Significant variation in the detection of hard-to-find, smaller early GCs may reflect differences in the physicians' performance during the examination; this is an important first step in discussing performance measures for upper endoscopy procedures.

INTRODUCTION

For early gastric cancers (GCs) without lymph node metastasis, less invasive endoscopic resection (ER) rather than conventional surgical treatment has become more common globally. Early detection of lesions at the stage where ER is possible in esophagogastroduodenoscopy (EGD) screening will greatly contribute to improving the quality of life of patients. However, detection of early GC requires more refined diagnostic proficiency from the endoscopist than does detection of other tumours of the gastrointestinal tract.[1 2] Physicians must discern the subtle elevation or depression of early GCs against the coarse

background of gastritis with limited time for examination. Therefore, it is sometimes difficult for even certified endoscopists to detect early GCs that demonstrate insubstantial morphologic changes and varying colour tones. Currently, detection of such hard-to-find early GCs depends greatly on the individual techniques of the endoscopist, resulting in interobserver variation in early GC detection rates.[3] This means that many early GCs are potentially overlooked during routine EGD.[4 5] Indeed, the reported false-negative rate for detecting GC with EGD varies from 4.6% to 25.8%.[2 4–7] This interobserver variation in detecting GCs has not typically been regarded as an issue of endoscopy quality for global clinical practice as detection rates of gastric neoplasms during screening EGD are much lower (generally <0.5%, at most 1.0%, in Japanese GC screening)[8–10] than the colonic adenoma detection rate (ADR), which is a well-known performance indicator for colonoscopy.[11 12] As ADR is considered the key benchmark of colonoscopic performance, it is surprising that it also exhibits drastic variations between endoscopists.[13] This variation is particularly worrisome because the ADR is inversely associated with the incidence and mortality of postcolonoscopy colorectal cancer.[14] Although variation in early GC detection rates (ie, missed GCs) for individual endoscopists has been recognised empirically, this has not traditionally been considered a difference in the physician's performance during the examination. This is because gastric neoplasm detection rate itself has not been established as a feasible performance indicator when performing EGD[15 16]; the variations noted in early GC detection among endoscopists are considered clinically coincidental due to the low prevalence of gastric neoplasm overall, because the detection rate could be greatly affected by the incidental and continuous encounter of the easy-to-find lesions (eg, larger early GCs and protruded lesions).

To confirm whether significant variations among endoscopists in early GC detection rates reflect the individual performance during the EGD examination, we retrospectively analysed the clinical data and background factors for patients with early GC, along with the overall routine EGD data at a single high-volume endoscopy centre in Japan.

METHODS

Study population

In this retrospective observational study, we reviewed the records for all consecutive patients who underwent EGD at New Tokyo Hospital, a district general hospital, between April 2017 and March 2020. Of the 28 582 EGDs performed, we excluded 1592 for including endoscopic treatment (eg, haemostasis, foreign-body removal, ER of the upper gastrointestinal tract). The remaining 26 990 EGDs were termed 'routine' (online supplemental file 1) and included screening studies performed in symptomatic or asymptomatic patients, and surveillance EGDs (eg, gastritis, post-*Helicobacter pylori* (HP) eradication, post-ER). All procedures were performed using a standard video-endoscope (GIF-H290Z, GIF-HQ290 or GIF-H260Z; Olympus Medical Systems, Tokyo, Japan) and endoscopic video system (EVIS LUCERA ELITE; Olympus).

This study focused on the identification of early-stage GCs amenable to endoscopic treatment, as the primary aim is to explicate the variations of GC detection rates among endoscopists, and advanced GCs were likely to be detectable by any endoscopist. Therefore, we purposefully excluded 96 cases of GCs that were identified during routine EGDs and considered to require surgical treatment; these cases mainly comprised advanced stage GCs, but also included particular lesions that were noted as requiring surgery during the initial EGD, even when subsequent surgical pathology revealed them to be early GCs. Consequently, we analysed the data recorded for 222 early GC lesions in 199 patients that were detected by target biopsies on routine EGDs and for which ER was indicated (online supplemental file 1). All GC lesions had pathologically confirmed carcinoma in the biopsy (ie, Group 5) and/or resected specimen using the pathologic criteria of the Japanese Gastric Cancer Association (adenomas were excluded).[17 18] It is worth noting that lesions classified as high-grade dysplasia in Western countries are included in the classification of GC in Japan because of the Japanese nuclear and structural criteria for diagnosis, even when invasion is absent.[19] All GCs were classified by their major diameter into minute GC (≤5 mm), small GC (6–10 mm) and larger GC (≥ 11 mm). The minute GCs included a small number of cases in which the cancer was resected completely by the biopsy procedure so there were no cancer cells in the endoscopically resected specimen after a Group 5 biopsy diagnosis.

GC detection rate

We assessed 12 faculty gastroenterologists competent to perform both upper and lower endoscopy procedure: five trainees (completed 2 years of junior residency but with <4 years of specialised endoscopy training) and seven board-certified fellows of the Japan Gastroenterological Endoscopy Society (with endoscopic subspecialties); all of them performed routine EGDs at the facility during the study period. We excluded faculty clinicians who performed fewer than 500 routine EGDs during the study period, which accounted for 1302 EGDs and 4 early GCs detected. We conducted a retrospective review of the remaining 25 688 EGDs and 218 GC lesions detected by the 12 endoscopists in 195 patients (online supplemental file 1).

We calculated the detection rate for each endoscopist by dividing the number of detected GCs by the number of routine EGDs performed. We defined the 'detector' as the endoscopist who first detected the GC lesion on routine EGD and who also performed the biopsy resulting in Group 2–5 classification. In the Japanese Group Classification system, Group 1 means normal tissue or a non-neoplastic lesion, whereas Group 2 lesions contain

atypical cells but are not definitively neoplasia.[17] Therefore, we did not define endoscopists performing biopsies of the lesions with Group 1 results as detectors. We calculated the biopsy rate (BR) by dividing the number of EGDs with at least one gastric biopsy by the total number of EGDs performed by each endoscopist. We calculated the positive predictive value (PPV) of endoscopic biopsy for each endoscopist by dividing the number of detected GCs by the number of EGDs with at least one gastric biopsy.

Baseline characteristics of patients undergoing routine EGD, such as age, sex, background mucosal atrophy and a history of prior gastric ER (mainly follow-up examinations after resection of gastric neoplasms), also were examined. The severity of endoscopic gastric atrophy was divided into closed- or open-type, according to the Kimura-Takemoto classification.[20]

Previously missed GCs and associated factors

Previously missed GC was defined as a cancer previously undiagnosed by EGD in the same facility within 3 years of the confirmatory diagnosis (ie, identified as 'hard-to-find' GC).[2 4 6 7] Previously missed GCs were also classified by their major diameter, and the detection rate was calculated as the number of previously missed GCs divided by the total number of routine EGDs.

As associated factors for failed and difficult GC detection, we noted the HP infection status and tumour location for each patient with early GC. Non-infection status was defined as having no gastric atrophy and confirmed negative results on either anti-HP immunoglobulin G (IgG), faecal antigen or urea breath testing. Currently infected patients were defined as those endoscopically exhibiting diffuse redness and atrophy with at least 1 of the following: positive anti-HP IgG, HP organisms detected in the biopsy specimen, positive faecal antigen testing or positive urea breath testing. Patients with past infection were defined as those endoscopically having obvious atrophy but no diffuse redness, with at least 1 of the following: negative anti-HP IgG, negative faecal antigen testing or negative urea breath testing. The patients with past infection were classified into a posteradication group and those with no history of eradication (eg, spontaneous resolution of HP). The location of detected GCs was classified by dividing the stomach into three equal sections along the long axis, as the upper (cardia, fundus and upper body), middle (midbody, lower body and angle) and lower (antrum and prepylorus) sections.

Statistical analysis for differences between observed frequencies was performed using the χ^2 test (1-sided) or t-test (2-sided) using Microsoft Excel to compare factors between the two groups and to calculate p values. In statistical comparative analyses, we excluded the endoscopists' data with the highest and lowest detection rate as outliers. Statistical significance was defined as a p value <0.05.

RESULTS

Of the 218 early GC lesions detected during 25 688 EGDs by the 12 endoscopists between April 2017 and March 2020, we excluded 3 GC lesions that were not identified by exact major diameters (because ER could not be performed en bloc with cancer-free margins). We also excluded one lesion that represented a local recurrence after GC treatment by ER, and two lesions that were biopsy-confirmed (Group 5) but left untreated per the patients' wishes. In our final analysis, we included 212 early GC lesions for which ER was performed, with or without subsequent surgical treatment (online supplemental file 1).

The mean early GC detection rate for our clinical endoscopy faculty was 0.83% (range: 0.09%–2.87%) (table 1). The GC detection rate exhibited enormous disparity among our endoscopists, especially for detection of minute and small GCs, regardless of whether the endoscopists were board-certified or trainees (figure 1). The BR of the trainees was significantly higher than that of the certified endoscopists (22.4% vs 11.0%; p<0.05), but the PPV for biopsy was not significantly different between trainees and certified endoscopists (4.0% vs 5.7%; p=0.39) (calculated results in table 1). We divided the 12 endoscopists into 2 groups of 6 each: the low-detection group (figure 1A–F; 71 GCs detected) and the high-detection group (figure 1G–K; 64 GCs detected) plus the single highest detector (figure 1L; 77 GCs detected). The high-detection endoscopists plus the highest had significantly greater GC detection rates for every GC size than the low-detection endoscopists, although there were no significant differences in baseline patient characteristics that might affect the prevalence of GC (eg, severe mucosal atrophy, history of post-ER of gastric neoplasms) (online supplemental file 2).

Since the single highest detector demonstrated an exceptional GC detection rate (77 GCs/2684 EGDs, 2.87%) compared with the other endoscopists, we treated this endoscopist's data separately for further comparative analysis. Endoscopists in the high-detection group had better detection rates for minute and small GCs than the low-detection group (0.19%/0.23% vs 0.085%/0.098%; p<0.05) in patients of similar background; these differences were significant (p<0.05) even though there were no significant differences in the detection of larger tumours (more than 11 mm in major diameter; 0.40% vs 0.28%; p=0.13) (online supplemental file 3). Incidentally, as the endoscopists with the lowest detection rates (0.09%) were obviously different from the detection rates of the other endoscopists, we excluded the data as an outlier. Thus, endoscopists in the high-detection group had significantly greater detection rates for minute and small GCs than those in the low-detection group after the exclusion of the lowest detector (0.19%/0.23% vs 0.092%/0.106%; p<0.05). No significant differences were observed in the detection rates of larger tumours (0.40% vs 0.30%; p=0.20) (table 2). It is noted that endoscopists in the low-detection group performed routine EGDs on

Table 1 Early GCs detected by 12 endoscopists and baseline characteristics of patients undergoing routine EGD

Endoscopist number		A	B	C	D	E	F	G	H	I	J	K	L
Detection Group		Low-detection group						High-detection group					Highest
Certified		(+)	(+)	(+)	(+)	(–)	(–)	(+)	(+)	(+)	(–)	(–)	(–)
Subspeciality		Pathology	ERCP/ EUS	BAE	ESD			Surgery	ESD	ERCP/EUS			
Detected GCs (n)		1	11	22	23	2	12	10	9	12	24	9	77
Routine EGDs (n)		1091	2163	4253	4149	1058	2547	1596	1129	1069	3140	809	2684
GC detection rate (%)		0.09	0.51	0.52	0.55	0.19	0.47	0.63	0.80	1.12	0.76	1.11	2.87
Atrophy	Closed type (n, (%))	270 (24.7)	338 (15.6)	907 (21.3)	1099 (26.5)	191 (18.1)	656 (25.8)	371 (23.2)	166 (14.7)	163 (15.2)	700 (22.3)	155 (19.2)	552 (20.6)
	Open type (n, (%))	351 (32.2)	746 (34.5)	1593 (37.5)	1316 (31.7)	471 (44.5)	808 (31.7)	549 (34.4)	337 (29.8)	261 (24.4)	1075 (34.2)	332 (41.0)	1118 (41.7)
Postgastric ER (n, (%))		134 (12.3)	117 (5.4)	269 (6.3)	165 (4.0)	54 (5.1)	124 (4.9)	38 (2.4)	48 (4.3)	63 (5.9)	160 (5.1)	33 (4.1)	224 (8.3)
EGDs with biopsy (n)		137	279	516	554	155	559	229	156	133	743	221	1130
EGDs with gastric biopsy (n)		116	209	452	485	117	464	199	140	102	604	201	1030
Biopsy rate (%)		10.6	9.7	10.6	11.7	11.1	18.2	12.5	12.4	9.5	19.2	24.8	38.4
Positive predictive value (%)		0.9	5.3	4.9	4.7	1.7	2.6	5.0	6.4	11.8	4.0	4.5	7.5

The early GC detection rates differed by over 10-fold (mean: 0.83%; range: 0.09%–2.87%) in a population with similar background. Endoscopists were divided into a low-detection group (A–F; detection rate: 0.09%–0.55%) and a high-detection group (G–K; detection rate: 0.63%–1.12%), with the single highest detector (L; detection rate: 2.87%). We discriminated trainees and board-certified fellows with endoscopic subspecialties.
Certified refers to board certified fellow of the Japan Gastroenterological Endoscopy Society.
BAE, balloon assisted endoscopy; EGD, esophagogastroduodenoscopy; ER, endoscopic resection; ERCP, endoscopic retrograde cholangiopancreatography; ESD, endoscopic submucosal dissection; EUS, endoscopic ultrasonography; GC, gastric cancer.

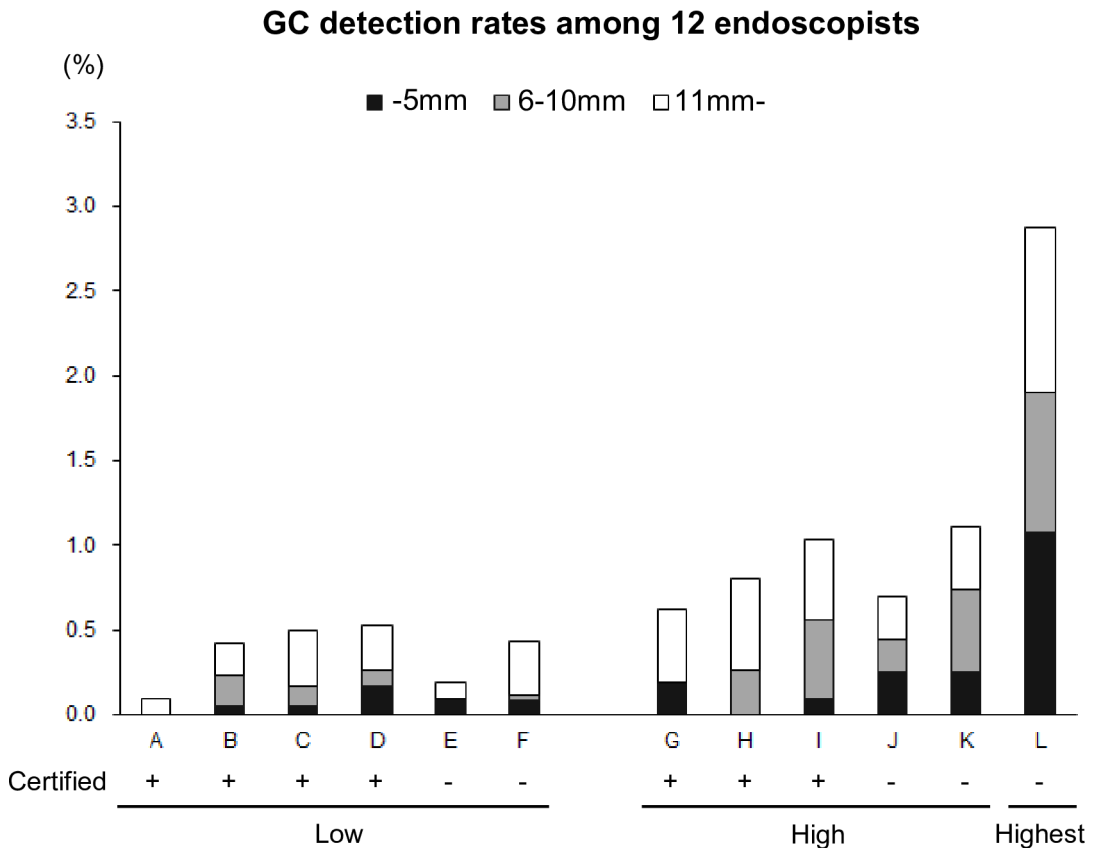


Figure 1 Early GC detection rates for 12 endoscopists. We observed a wide variation among endoscopists in early GC detection rates, especially for minute (major diameter ≤5 mm) and small (major diameter 6–10 mm) GCs, under the same circumstances at a district general hospital. Endoscopists were divided into a low-detection group (A-F; detection rate, 0.09%–0.55%) and a high-detection group (G–K; detection rate, 0.63%–1.12%), while the single highest detector (L; detection rate, 2.87%) was treated independently, due to their markedly disparate detection rate. +: Board certified fellow of the Japan Gastroenterological Endoscopy Society. GC, gastric cancer.

a patient population with a slightly increased risk for GC, necessitating more careful observation by endoscopists to identify GCs, in contrast to the high-detection group. This encompassed a marginally higher ratio of patients exhibiting mucosal atrophy and those receiving follow-up care after ER of gastric neoplasms (refer to online supplemental file 3 and table 2). Moreover, endoscopists in the high-detection versus low-detection group, excluding the lowest detector, exhibited no significant differences in the BR (15.7% vs 12.2%; p=0.31) and PPV for biopsy (6.3% vs 3.8%; p=0.15) (calculated results in table 1).

The detection rate of previously missed GCs in the low-detection group was also significantly different than that in the high-detection group (0.18% vs 0.32%; p<0.05) (table 3). The single highest detector had a significantly better detection rate for previously missed GCs than either the low-detection or high-detection groups (1.5%; p<0.01); this difference was also apparently caused by a higher ratio of the smaller-sized GCs detected in patients with previously missed early GC results (table 3). The HP infection status (non-infection, current infection, past-infection (posteradication, no history of eradication)) was similar in the low-detection group, high-detection group and for the highest detector (table 4). No trend in tumour location was observed among the three groups (table 4).

DISCUSSION

Using endoscopy results from controlled clinical circumstances in a large population of similar background, we proved that there are obvious, significant differences in early GC detection rates among endoscopists, especially for detection of smaller lesions (ie, minute and small GCs). Detection rates of previously missed GCs—lesions that might have been missed during prior EGDs—were also significantly different among observers, in accordance with the differences noted for both total and minute-and-small GC detection. Because detecting smaller GCs and previously missed GCs typically requires more refined diagnostic techniques from the endoscopist than detecting easy-to-find larger lesions, this observed variation in early GC detection was not merely an incidental artefact of the low incidence of gastric neoplasm overall but may reflect a difference in endoscopists' performance of the examination.

Surprisingly, the wide individual variation in early GC detection rates seen at our district general hospital

Table 2 Comparison of baseline EGD characteristics and early GC detection rates between the low-detection group except the lowest detector and high-detection group

			Low (except the lowest)		High		P value	
Characteristics of routine EGDs (n, (%))								
Total			14 170		7743			
Sex, m (n, (%))			7843 (55.3)		4265 (55.1)		0.70	
Mean age (range)			64.3 (16–99)		64.3 (17–98)		0.85	
Atrophy	Closed (n, (%))		3188 (22.5)		1552 (20.0)		<0.01	
	Open (n, (%))		4935 (34.8)		2551 (32.9)		<0.01	
Postgastric ER (n, (%))			700 (4.94)		342 (4.42)		0.08	
Detected GCs (n, (%))								
Total			70 (0.49)		64 (0.83)		<0.01	
Tumour size (mm)	≤5	≤10	13 (0.092)	28 (0.20)	15 (0.19)	33 (0.43)	<0.05	<0.01
	6–10		15 (0.106)		18 (0.23)		<0.05	
	≥11		42 (0.30)		31 (0.40)		0.20	

Since the endoscopists with the highest and lowest detection rates (2.87% and 0.09%) were so far removed from the detection rates of the other endoscopists, we excluded those data as outliers from this comparative analysis. Therefore, endoscopists were assigned to the low-detection group after exclusion of the lowest detector (n=5; detection rate: 0.19%–0.55%) and the high-detection group (n=5; detection rate: 0.63%–1.12%) for statistical comparative analysis.
EGD, esophagogastroduodenoscopy; ER, endoscopic resection; GC, gastric cancer.

appears to reflect the actual clinical situation as our mean early GC detection rate (0.83%) matches the national standard level in Japan (generally 0.5%–1.0%).[8–10] American Society for Gastrointestinal Endoscopy guidelines for endoscopic training recommend a minimum of 100 supervised EGDs to acquire proficiency in the procedure.[21] All endoscopists in this study, including trainees, possess considerable experience performing numerous endoscopic procedures at our high-volume endoscopy facility. Consequently, they were sufficiently competent to perform the examination, ensuring the consistency and accuracy of image acquisition within the limited examination time. Recently, interesting article was published from a district general hospital suggesting that assigning a different endoscopist for each annual follow-up contributed to improved GC detection rates.[22] This report may suggest that there are the differences in physicians' techniques and peculiar habit of the EGD procedure. Considering that the main purpose of the screening/surveillance EGD is early detection of cancerous lesions that are candidates for ER and surgery, our results should be enlightening and impactful for general clinicians as well as for patients getting EGDs. Although it is unclear whether the missed smaller-sized GCs could affect patients' prognosis, it might have at least some impacts on the patient's quality of life. Because the relationship between GC size and submucosal invasion at the same facility indicated that approximately 4%–5% of smaller GCs (≤10 mm) had already developed submucosal invasion (data not shown), that could be an indication for surgical treatment. Interestingly, the highest-scoring physician detector in our study demonstrated exceptionally high performance on routine EGD and effectively detected a variety of hard-to-find larger GCs as well as smaller-sized GCs, which most certified endoscopists at our facility were unable to find. Hence, the differences in endoscopic performance among physicians sometimes might lead them to critically miss of larger, invasive GCs, that would require surgical resection. We will report this physician's diagnostic technique with image-enhanced endoscopy (IEE), and his strategies for routine EGD, in more detail in a different venue.

Table 3 Previously missed GCs detected by the low-detection group, the high-detection group and the single highest detector

	Low	High	Highest
Previously missed GCs (n (detection rate, %))	28 (0.18)*	25 (0.32)* **	40 (1.5)**
Tumour size (mm)			
≤5	6	6	18
6–10	10	7	13
≥11	12	12	9

*p<0.05, **p<0.01.
GC, gastric cancer.

Reported detection and missed rates of gastric neoplasm for trainees are not significantly different than those for experienced endoscopists,[22–25] even though endoscopic training and education were essential ways to increase the GC detection.[26 27] Our study had similar findings: the observed GC detection rates in routine

Table 4 HP status and tumour location of early GCs detected by the low-detection group, the high-detection group and the single highest detector

	Low		High		Highest	
HP status (n (%))						
Non-infection	3 (4.2)		1 (1.6)		3 (3.9)	
Current infection	17 (24)		20 (31)		26 (34)	
Pastinfection_history of eradication (+)	27 (38)	49 (69)	19 (30)	42 (66)	25 (33)	46 (60)
(–)	22 (31)		23 (36)		21 (27)	
Uncertain	2 (2.8)		1 (1.6)		2 (2.6)	
Location (n (%))						
U	12 (17)		10 (16)		13 (17)	
M	34 (48)		29 (45)		36 (47)	
L	25 (35)		25 (39)		28 (36)	

HP, *Helicobacter pylori*; L, lower third of the stomach; M, middle third of the stomach; U, upper third of the stomach.

EGDs bore no relation to whether the endoscopists were trainees or experienced, board-certified endoscopists. Moreover, the BR of trainees was approximately twice that of certified endoscopists. In contrast, the PPV of target biopsy had no significant differences between trainees and certified endoscopists. Therefore, we hypothesise that the performance of early GC detection (ie, identifying early GC findings at a longer-distance view) might not be correlated with years of endoscopy experience, even though it has been reported that diagnostic proficiency in 'optical biopsy' (ie, effectively classifying GCs and benign endoscopic findings at short-distance views without obtaining biopsies) can be acquired with experience.[28 29]

To our knowledge, only a few publications describe the parameters influencing gastric neoplasm detection (eg, rate of detection of gastric subepithelial lesions and diverticula,[30] photodocumentation of the ampulla,[31] endoscopist BR,[32] average examination time[10 33] and use of IEE[34 35]); however, unlike the ADR in colonoscopy, these parameters are not feasible candidates of easy-to-understand performance indicator for EGD because gastric neoplasm detection has not been an identified performance measure of EGD. The minimum requirements for maintaining endoscopy quality might be defined by photodocumentation of endoscopic images, the endoscopist BR or examination time. However, a larger number of endoscopic images and biopsies, or longer examination time are not always better for routine EGD performed under time constraints. Indeed, our data showed there was no association between the BR and GC detection rate. We did not investigate the number of endoscopic images and the examination time in this research, but it is noted that the minimum requirements for photodocumentation vary worldwide,[36] and no association between examination time and upper gastrointestinal neoplasm detection rate has been reported in cancer hub hospitals in Japan.[37] Therefore, we believe that these parameters merely demonstrated the minimum necessary condition for endoscopists to carefully perform the EGD procedure and did not directly lead to greater GC detection. On the other hand, the adoption of IEE might potentially affect the detection of smaller GCs, though we could not evaluate all routine EGDs to determine whether each endoscopist used IEE or not. In our facility, the routine application of IEE in high-risk patient screening (eg, severe atrophy and post-ER follow-up) remained at the discretion of individual endoscopists, although they consistently used IEE when gastric neoplasm-suggestive lesions were detected during EGDs. In the current situation, Japanese guidelines for endoscopic diagnosis of early GC state that white-light observation is the basic method for gastric endoscopy and that not all endoscopists routinely use IEE as a mandatory method of detecting early GC.[38] Hence, the use of IEE, the length of examination time and the number of images taken are distinctive traits of each endoscopist that greatly depend on the GC risk of patients. We propose that these characteristics contribute to the observed variations (ie, individual performance) in GC detection rates among endoscopists.

In this study, we particularly emphasise the variation (ie, the 'missed GC rate') among endoscopists in detection of smaller GCs as reflecting the overall rate of early GC detection. By contrast, individual variability in detection of early GCs was not affected by detection of posteradication GCs or tumour location. Recently, early GCs detected after HP eradication have been identified as 'hard-to-find' GCs, because posteradication GCs exhibit a gastritis-like appearance, similar to the surrounding background mucosa.[39–41] Though missed GCs might also be associated with tumour location,[4 42] some reports claim, controversially, that HP infection status and tumour location do not significantly alter GC detection rates.[43] Further research is necessary to clarify these

discrepant conclusions. However, by this definition the small size of GC lesions may be more of an indication of 'harder-to-find' status than the HP infection status or the tumour location. Some paper reported that missed GCs were usually small-sized (<20 mm) intramucosal cancer[44] and the median size of GCs newly detected in an annual surveillance EGD were 10 mm.[22 45] According to published data for artificial intelligence (AI)-based computer-aided detection systems in surveillance EGD, the sensitivity of detecting minute GCs by AI was 16.7% though 98.6% of GCs with a diameter of 6 mm or more were correctly detected.[46] Another paper showed that smaller GC size (1–13 mm) results in statistically lower accuracy for GC detection using AI on multivariable analysis.[47] These results make sense, because larger lesions could be hardly missed by both endoscopists and AI, but the detection of smaller GCs tends to differ among observers. The overall GC detection rate has not been considered a viable performance measure, as it is influenced by the incidental and continuous encounter of larger GCs, which are easily detected by any observer. Detecting minute and/or small GCs might more accurately reflect the performance of the endoscopist than the overall gastric neoplasm detection rate. Therefore, the rate of detection of smaller GCs could potentially serve as a performance measure for routine EGDs conducted under similar circumstances or, at least, within the same clinical facility. However, the prevalence of GC varies greatly depending on the patient's background, including differences in HP prevalence and the number of post-ER GCs they have had. Further study is needed to lend the difference of smaller GC detection among endoscopists to broader applications as a comparable measure for EGD in different regions.

Our study has certain limitations. First, this was a single-centre, retrospective study, making it difficult to accurately align patient backgrounds. However, prospective study designs that assess endoscopy performance tend to produce better detection rates; these may be influenced by potential bias resulting from the possibility that the endoscopists, who are informed about and enrolled in such studies, make their observations more carefully. This phenomenon, specifically in a research setting, can be attributed to the Hawthorne effect,[48] in which performance improvement occurs in situations where evaluations and monitoring are being conducted. It has been reported that there is a correlation between a Hawthorne effect and an increased ADR.[49] Second, we did not investigate the aforementioned parameters of routine EGDs potentially influencing GC detection. Therefore, the possibility cannot be ruled out that endoscopists in low-detection group did not spend the minimum required observation time and rarely use IEE, as each endoscopist at our facility had to perform many routine EGDs within similar time constraints. Third, no exclusion criteria in routine EGDs were applied beyond the requirement that they be performed by the 12 participating physicians; however, the individual objectives of each EGD, such as screening or surveillance, and first-time or follow-up examinations, were not assessed. Fourth, though the presence of gastric atrophy and post-ER follow-up were evaluated in routine EGDs, the gastric atrophy was not evaluated histologically but based on each endoscopist's assessment; thus, this could potentially lead to bias in the overdiagnosing or underdiagnosing gastric atrophy.

In conclusion, we found that the wide variability in early GC detection rates among physicians, in patients with similar background, contributes to variable rates of detection of minute and small GC lesions. The detection rate for these 'harder-to-find' GCs might have a practical use in assessing individual performance on upper endoscopy in clinical practice.

Acknowledgements The authors gratefully acknowledge Prof David W Grainger (University of Utah, USA) for invaluable technical comments and the clinical work of past colleagues of New Tokyo Hospital.

Contributors Study conception and design, data collection and analysis were performed by DM. The first draft of the manuscript was written by DM. Supervision on drafting the article was performed by MY and MA. Critical revision of the article for important intellectual content was performed by MY, YA, TN and MA. All authors commented on the manuscript. All authors approved the final manuscript. The manuscript's guarantor is MA.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval We certify that the procedures were performed in accordance with the ethical standards of the 1975 Declaration of Helsinki as revised in 2000, as well as national law. Before undergoing endoscopy, all patients provided comprehensive written informed consent. The study was approved by the institutional review boards of New Tokyo Hospital (IRB no. 0181).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Raw data were generated at New Tokyo Hospital. Derived data supporting the findings of this study are available from the corresponding author (DM) on request.

ORCID iD

Daisuke Murakami http://orcid.org/0000-0002-6012-844X

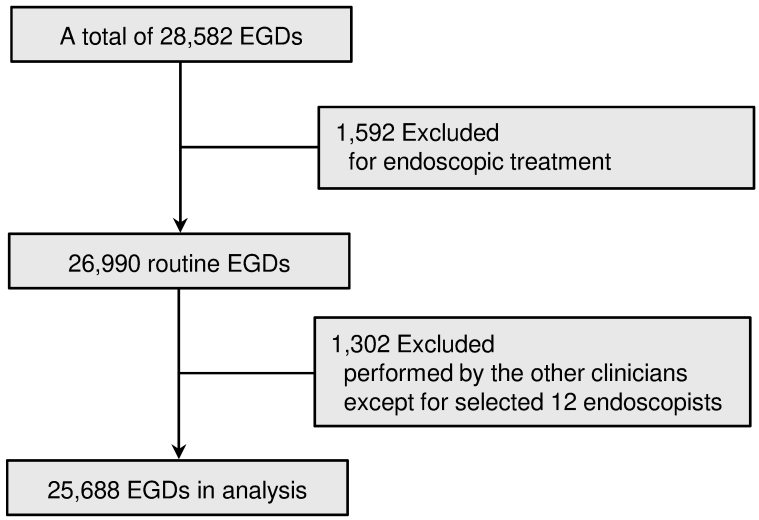
REFERENCES

1 Gotoda T, Uedo N, Yoshinaga S, *et al*. Basic principles and practice of gastric cancer screening using high-definition white-light Gastroscopy: eyes can only see what the brain knows. *Dig Endosc* 2016;28 Suppl 1:2–15.
2 Hosokawa O, Hattori M, Douden K, *et al*. Difference in accuracy between gastroscopy and colonoscopy for detection of cancer. *Hepatogastroenterology* 2007;54:442–4.
3 Murakami D, Yamato M, Amano Y, *et al*. Challenging detection of hard-to-find gastric cancers with artificial intelligence-assisted endoscopy. *Gut* 2021;70:1196–8.

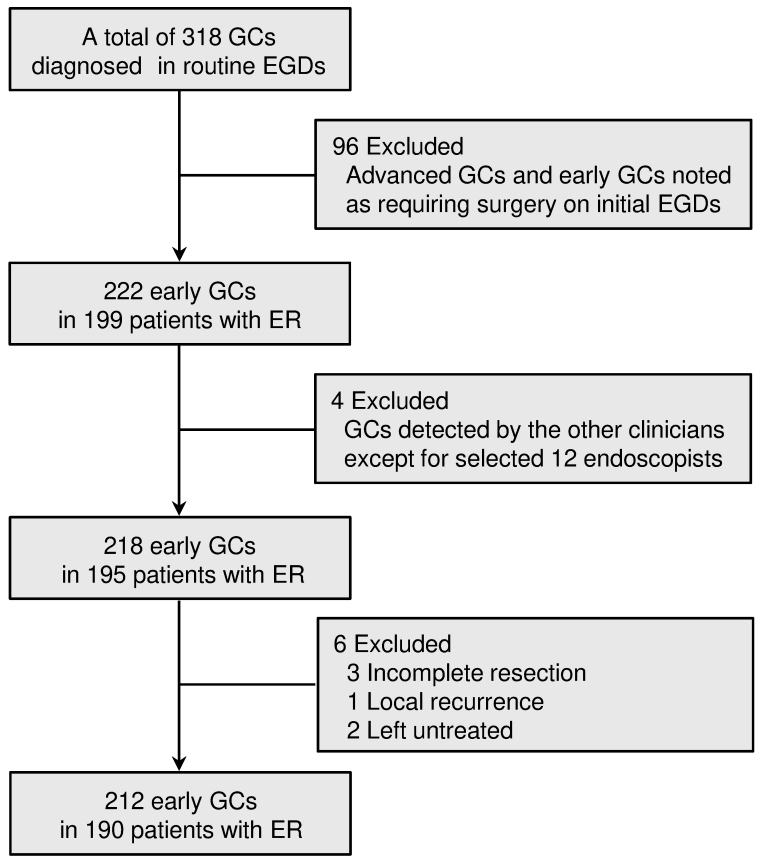
4 Hosokawa O, Tsuda S, Kidani E, *et al*. Diagnosis of gastric cancer up to three years after negative upper gastrointestinal Endoscopy. *Endoscopy* 1998;30:669–74.
5 Lee HL, Eun CS, Lee OY, *et al*. When do we Miss synchronous gastric neoplasms with endoscopy *Gastrointest Endosc* 2010;71:1159–65.
6 Menon S, Trudgill N. How commonly is upper gastrointestinal cancer missed at endoscopy? A meta-analysis. *Endosc Int Open* 2014;2:E46–50.
7 Chadwick G, Groene O, Riley S, *et al*. Gastric cancers missed during endoscopy in England. *Clin Gastroenterol Hepatol* 2015;13:1264–70.
8 Hamashima C, Okamoto M, Shabana M, *et al*. Sensitivity of endoscopic screening for gastric cancer by the incidence method. *Int J Cancer* 2013;133:653–9.
9 JSGCS. *A manual of endoscopic screening for gastric cancer in population-based screening 2015*. Tokyo: The Japanese Society of Gastrointestinal Cancer Screening, c2016. Available: https://www.jsgcs.or.jp
10 Kawamura T, Wada H, Sakiyama N, *et al*. Examination time as a quality indicator of screening upper gastrointestinal endoscopy for asymptomatic examinees. *Dig Endosc* 2017;29:569–75.
11 Kaminski MF, Regula J, Kraszewska E, *et al*. Quality indicators for colonoscopy and the risk of interval cancer. *N Engl J Med* 2010;362:1795–803.
12 Lee TJW, Rutter MD, Blanks RG, *et al*. Colonoscopy quality measures: experience from the NHS bowel cancer screening programme. *Gut* 2012;61:1050–7.
13 Lee RH, Tang RS, Muthusamy VR, *et al*. Quality of colonoscopy withdrawal technique and variability in adenoma detection rates (with videos). *Gastrointest Endosc* 2011;74:128–34.
14 Corley DA, Jensen CD, Marks AR, *et al*. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014;370:1298–306.
15 Park WG, Shaheen NJ, Cohen J, *et al*. Quality indicators for EGD. *Gastrointestinal Endoscopy* 2015;81:17–30.
16 Bisschops R, Areia M, Coron E, *et al*. Performance measures for upper gastrointestinal endoscopy: a European society of gastrointestinal Endoscopy (ESGE) quality improvement initiative. *Endoscopy* 2016;48:843–64.
17 Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma, 3rd English edition. *Gastric Cancer* 2011;14:101–12.
18 Schlemper RJ, Riddell RH, Kato Y, *et al*. The Vienna classification of gastrointestinal epithelial neoplasia. *Gut* 2000;47:251–5.
19 Schlemper RJ, Itabashi M, Kato Y, *et al*. Differences in diagnostic criteria for gastric carcinoma between Japanese and Western pathologists. *Lancet* 1997;349:1725–9.
20 Kimura K, Takemoto T. An endoscopic recognition of the atrophic border and its significance in chronic Gastritis. *Endoscopy* 1969;1:87–97.
21 ASGE. *Alternative Pathways to Training in Gastrointestinal Endoscopy, Training Guideline*. Illinois: American Society for Gastrointestinal Endoscopy, c2020. Available: https://www.asge.org/home/education-meetings/training-trainees/training-core-curriculum#training-guidelines
22 Unno S, Igarashi K, Saito H, *et al*. Assigning a different endoscopist for each annual follow-up may contribute to improved gastric cancer detection rates. *Endosc Int Open* 2022;10:E1333–42.
23 Oka K, Iwai N, Okuda T, *et al*. Clinical features of false-negative early gastric cancers: a retrospective study of endoscopic submucosal dissection cases. *Gastroenterol Res Pract* 2021;2021:6635704.
24 Raftopoulos SC, Segarajasingam DS, Burke V, *et al*. A cohort study of missed and new cancers after esophagogastroduodenoscopy. *Am J Gastroenterol* 2010;105:1292–7.
25 Hernanz N, Rodríguez de Santiago E, Marcos Prieto HM, *et al*. Characteristics and consequences of missed gastric cancer: a multicentric cohort study. *Dig Liver Dis* 2019;51:894–900.
26 Zhang Q, Chen Z, Chen C, *et al*. Training in early gastric cancer diagnosis improves the detection rate of early gastric cancer: an observational study in China. *Medicine* 2015;94:e384.
27 Kumar NL, Smith BN, Lee LS, *et al*. Best practices in teaching endoscopy based on a Delphi survey of gastroenterology program directors and experts in endoscopy education. *Clin Gastroenterol Hepatol* 2020;18:574–9.
28 Dohi O, Yagi N, Majima A, *et al*. Diagnostic ability of magnifying endoscopy with blue laser imaging for early gastric cancer: a prospective study. *Gastric Cancer* 2017;20:297–303.
29 Yao K, Uedo N, Muto M, *et al*. Development of an E-learning system for the endoscopic diagnosis of early gastric cancer: an international multicenter randomized controlled trial. *EBioMedicine* 2016;9:140–7.
30 Park CH, Kim B, Chung H, *et al*. Endoscopic quality indicators for esophagogastroduodenoscopy in gastric cancer screening. *Dig Dis Sci* 2015;60:38–46.
31 Park JM, Lim C-H, Cho YK, *et al*. The effect of photo-documentation of the ampulla on neoplasm detection rate during esophagogastroduodenoscopy. *Endoscopy* 2019;51:115–24.
32 Januszewicz W, Wieszczy P, Bialek A, *et al*. Endoscopist biopsy rate as a quality indicator for outpatient gastroscopy: a multicenter cohort study with validation. *Gastrointest Endosc* 2019;89:1141–9.
33 Park JM, Huo SM, Lee HH, *et al*. Longer observation time increases proportion of neoplasms detected by esophagogastroduodenoscopy. *Gastroenterology* 2017;153:460–9.
34 Dohi O, Ono S, Kawada K, *et al*. Linked color imaging provides enhanced visibility with a high color difference in upper gastrointestinal neoplasms. *J Gastroenterol Hepatol* 2023;38:79–86.
35 Ono S, Kawada K, Dohi O, *et al*. Linked color imaging focused on neoplasm detection in the upper gastrointestinal tract: a randomized trial. *Ann Intern Med* 2021;174:18–24.
36 Kim SY, Park JM. Quality indicators in esophagogastroduodenoscopy. *Clin Endosc* 2022;55:319–31.
37 Yoshimizu S, Hirasawa T, Horiuchi Y, *et al*. Differences in upper gastrointestinal neoplasm detection rates based on inspection time and esophagogastroduodenoscopy training. *Endosc Int Open* 2018;6:E1190–7.
38 Yao K, Uedo N, Kamada T, *et al*. Guidelines for endoscopic diagnosis of early gastric cancer. *Dig Endosc* 2020;32:663–98.
39 Saka A, Yagi K, Nimura S. Endoscopic and histological features of gastric cancers after successful Helicobacter Pylori eradication therapy. *Gastric Cancer* 2016;19:524–30.
40 Kitamura Y, Ito M, Matsuo T, *et al*. Characteristic epithelium with low-grade Atypia appears on the surface of gastric cancer after successful Helicobacter Pylori eradication therapy. *Helicobacter* 2014;19:289–95.
41 Masuda K, Urabe Y, Ito M, *et al*. Genomic landscape of epithelium with low-grade Atypia on gastric cancer after Helicobacter Pylori eradiation therapy. *J Gastroenterol* 2019;54:907–15.
42 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;28:1041–9.
43 Cho YS, Chung I-K, Kim JH, *et al*. Risk factors of developing interval early gastric cancer after negative endoscopy. *Dig Dis Sci* 2015;60:936–43.
44 Abe S, Oda I, Minagawa T, *et al*. Metachronous gastric cancer following curative endoscopic resection of early gastric cancer. *Clin Endosc* 2018;51:253–9.
45 Abe S, Oda I, Suzuki H, *et al*. Long-term surveillance and treatment outcomes of metachronous gastric cancer occurring after curative endoscopic submucosal dissection. *Endoscopy* 2015;47:1113–8.
46 Hirasawa T, Aoyama K, Tanimoto T, *et al*. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer* 2018;21:653–60.
47 Yoon HJ, Kim S, Kim J-H, *et al*. A lesion-based convolutional neural network improves endoscopic detection and depth prediction of early gastric cancer. *J Clin Med* 2019;8:1310.
48 McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *J Clin Epidemiol* 2014;67:267–77.
49 Then EO, Brana C, Dadana S, *et al*. Implementing visual cues to improve the efficacy of screening colonoscopy: exploiting the Hawthorne effect. *Ann Gastroenterol* 2020;33:374–8.

Supplemental 1

Flowchart of selected routine EGDs



Flowchart of selected early GCs detected on EGDs and endoscopically resected



Supplemental 2

Comparison of baseline characteristics of patients undergoing routine esophagogastroduodenoscopy and gastric cancer detection rates between in the low-detection and high-detection group plus the single highest detector

EGD, esophagogastroduodenoscopy; ER, endoscopic resection; GC, gastric cancer

		Low	High+Highest	*P* Value
Characteristics of routine EGDs, No. (%)				
Total		15261	10427	
Sex, m (%)		8404 (55.1)	5777 (55.4)	0.60
Mean age (range)		64.2 (16-99)	64.8 (17-98)	
Atrophy	Closed (%)	3458 (23.2)	2103 (20.2)	< 0.01
	Open (%)	5286 (34.6)	3667 (35.2)	0.38
Post-gastric ER (%)		863 (5.65)	571 (5.48)	0.54
Detected GCs, No. (%)				
Total		71 (0.47)	141 (1.35)	< 0.01
Tumor size (mm)	5 ≤	13 (0.085)	44 (0.42)	< 0.01
	6-10	15 (0.098)	40 (0.38)	< 0.01
	≥ 11	43 (0.28)	57 (0.55)	< 0.01

Supplemental 3

Comparison of baseline characteristics of patients undergoing routine esophagogastroduodenoscopy and gastric cancer detection rates between in the low-detection and high-detection group

EGD, esophagogastroduodenoscopy; ER, endoscopic resection; GC, gastric cancer

		Low		High		*P* Value
Characteristics of routine EGDs, No. (%)						
Total		15261		7743		
Sex, m (%)		8404 (55.1)		4265 (55.1)		0.98
Mean age (range)		64.2 (16-99)		64.3 (17-98)		
Atrophy	Closed (%)	3458 (23.2)		1552 (20.0)		< 0.01
	Open (%)	5286 (34.6)		2551 (32.9)		< 0.05
Post-gastric ER (%)		863 (5.65)		342 (4.42)		< 0.01
Detected GCs, No. (%)						
Total		71 (0.47)		64 (0.83)		< 0.01
Tumor size (mm)	5 ≤	13 (0.085)	28 (0.18)	15 (0.19)	33 (0.43)	< 0.05
	6-10	15 (0.098)		18 (0.23)		< 0.05
	≥ 11	43 (0.28)		31 (0.40)		0.13
	10 ≤					< 0.01