

# Endoscopic grading of gastric intestinal metaplasia using blue light imaging in a low-risk population: Multicenter cross-sectional validation study



## Authors

Marta Rodriguez-Carrasco<sup>1</sup>  Gianluca Esposito<sup>2</sup>, Emanuele Dilaghi<sup>2</sup>, Guido Manfredi<sup>3</sup>, Saverio Alicante<sup>3</sup>, Elisabetta Buscarini<sup>3</sup>, Bruno Annibale<sup>2</sup>, Mário Dinis-Ribeiro<sup>1,4</sup>

## Institutions

- 1 Gastroenterology Department, Instituto Portugues de Oncologia do Porto Francisco Gentil, Porto, Portugal
- 2 Department of Medical-Surgical Sciences and Translational Medicine, Sant'Andrea Hospital, Sapienza University of Rome, Rome, Italy
- 3 Gastroenterology and Endoscopy Department, Maggiore Hospital Crema, Crema, Italy
- 4 Porto Comprehensive Cancer Center & RISE@CI-IPO, University of Porto, Porto, Portugal

## Key words

Endoscopy Upper GI Tract, Precancerous conditions & cancerous lesions (displasia and cancer) stomach, Diagnosis and imaging (inc chromoendoscopy, NBI, iSCAN, FICE, CLE), Endoscopic resection (ESD, EMRC, ...)

received 18.6.2024

accepted after revision 3.12.2024

## Bibliography

Endosc Int Open 2025; 13: a25003748

DOI 10.1055/a-2500-3748

ISSN 2364-3722

© 2025. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Oswald-Hesse-Straße 50, 70469 Stuttgart, Germany

## Corresponding author

Dr. Marta Rodriguez-Carrasco, Instituto Portugues de Oncologia do Porto Francisco Gentil, Gastroenterology Department, Rua Dr. António Bernardino de Almeida, 4200-072 Porto, Portugal  
martarc7@gmail.com

## ABSTRACT

**Background study aims** Detecting gastric intestinal metaplasia (GIM) with white light endoscopy (WLE) remains a challenge and virtual chromoendoscopy methods have been shown to increase accuracy. We aimed to externally validate the Endoscopic Grading of Gastric Intestinal Metaplasia (EGGIM) using blue light imaging (BLI).

**Methods** First, the reliability of BLI and the EGGIM score was evaluated through assessment of 90 images divided into three sets of 30. A multicenter cross-sectional study was conducted at two Italian centers involving 102 patients (510 biopsies). Both per-biopsy and per-patient analyses were performed to ascertain accuracy of BLI in detecting and staging GIM (vs. histology).

**Results** BLI significantly enhanced interobserver agreement of endoscopic diagnosis of GIM, with a Fleiss Kappa of 0.4 (95% confidence interval [CI] 0.3–0.5), compared to 0.2 (95% CI 0.2–0.3) with WLE. Concordance was particularly strong in applying the EGGIM score (weighted Kappa 0.7; 95% CI 0.5–0.9). BLI showed significant improvements in sensitivity over WLE, with an increase observed in both per-biopsy analysis (82%; 95%CI 73.7–89.0 vs. 50%;95% CI 40.6–60.3) and per-patient analysis (96%; 95% CI 84.5–99.4 vs. 68%;95% CI 52.4–81.4). The area under the curve of EGGIM in diagnosing OLGIM III/IV was 0.9 (95% CI 0.8–1.0), confirming EGGIM > 4 being the optimal threshold (sensitivity of 80%, specificity of 88%).

**Conclusions** Our study validates BLI integrated with the EGGIM system as an effective strategy, highlighting its precision in identifying advanced GIM stages. BLI's notable sensitivity enhances its use as a complementary tool to WLE, significantly improving gastric cancer risk assessment.

## Introduction

Gastric intestinal metaplasia (GIM) is a well-established precursor in the carcinogenic process leading to gastric cancer (GC) [1] and represents the most reliable indicator of precancerous activity from a histological perspective [2]. Accurate endoscopic recognition of GIM is essential to identify individuals at increased risk for GC in order to offer proper follow-up [3]. Nevertheless, diagnostic challenges inherent in white light endoscopy (WLE) are well-documented, with its sensitivity for detecting GIM notably varying between 53% and 75% [4, 5, 6]. Such limitations led to development of virtual chromoendoscopy (VCE) technologies, with narrow band imaging (NBI) being the most rigorously evaluated. A recent meta-analysis positions NBI as the most effective technology for diagnosing GIM, demonstrated by its high diagnostic accuracy [7]. However, the value of VCE technologies transcends mere diagnostic capabilities, also offering significant contributions to stratification of individual risk for GC. Although the Operative Link on Gastritis Assessment for Intestinal Metaplasia (OLGIM) system remains the benchmark for histological grading [2], introduction of the Endoscopic Grading of Gastric Intestinal Metaplasia (EGGIM) score represents a noteworthy advancement, enabling comprehensive assessment of the entire mucosa via endoscopic imaging instead of relying on selective biopsy sites [8]. This grading system has been externally validated with NBI, demonstrating excellent diagnostic performance in comparison to OLGIM staging, especially for EGGIM scores above 4 (sensitivity 89.4%, specificity 94.6%) [8].

Despite the proven efficacy of various VCE modalities, the literature remains sparse on the performance of newer technologies. Grounded in the same physical principles as NBI, blue light imaging (BLI) distinguishes itself by foregoing traditional optical filters. Instead, BLI enhances imaging through modulation of LED light intensity, producing high-quality images that closely resemble those obtained with NBI. Although initial evidence suggests a comparable diagnostic accuracy between NBI and BLI, the bulk of research on BLI has been conducted in populations at high risk for GC, and has predominantly concentrated on its predecessor laser technologies and diagnosis of early gastric cancer (EGC) [9, 10, 11, 12, 13, 14, 15]. In a recent study conducted by our group involving a cohort of 37 patients, BLI utilized alongside the EGGIM scoring system exhibited high diagnostic accuracy in identification of advanced GIM stages, with a sensitivity and specificity of 100.0% and 79.2%, respectively for scores greater than 4 [16]. However, the limited sample size and the epidemiological context of being an intermediate-risk country for GC restrict its generalizability to other settings. Although the EGGIM score has emerged as a promising tool for refining risk stratification, its comprehensive validation and practical application with BLI warrant further investigation, especially in low-risk GC populations, defined by an age-standardized incidence rate (ASR) of less than 10 per 100,000 people per year (<https://gco.iarc.fr/>). In these populations, diagnostic performance may vary significantly compared with high-risk groups, potentially influencing clinical endoscopic practice and patient management strategies. Understanding these dif-

ferences is critical for optimizing use of BLI and EGGIM in diverse clinical settings.

This study aimed to assess diagnostic accuracy of BLI in detecting GIM and to externally validate applicability of the EGGIM score in conjunction with BLI for staging GIM in a low-risk GC population.

## Patients and methods

This multicenter study was structured to execute a dual-analysis approach. The first component, Study 1, focused on a reliability analysis to assess and standardize diagnostic processes across participating centers, ensuring that the employed diagnostic techniques remained consistent and reliable irrespective of the performing center. The second component, Study 2, was dedicated to an accuracy analysis between BLI and WLE in detection of GIM.

Our study was designed in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology statement [17], ensuring a robust and transparent reporting framework. The study protocol received prior approval from the Ethics Committee of each participating center, and was further registered to ClinicalTrials.gov on February 17, 2021 (ID: NCT04768218).

### Study 1: Reliability analysis of BLI and EGGIM classification

Interobserver agreement was evaluated among participating endoscopists, which included two endoscopists from each participating center and one endoscopist from the coordinating center. All endoscopists were proficient in NBI or BLI. This evaluation was conducted through assessment of 90 endoscopic images from antrum, incisura, and corpus, without magnification. The initial 60 images included 30 WLE and 30 BLI images selected from a Portuguese cohort, which showcased normal mucosa, AG, GIM, and dysplastic lesions. This step aimed to calculate overall agreement for both technologies among all endoscopists. The remaining 30 images corresponded to BLI images with different EGGIM scores selected from patients included in Study 2. This step aimed to calculate interobserver agreement for EGGIM classification between two endoscopists: one from the participating centers in Study 2 and one from the coordinating center.

### Study 2: Accuracy analysis of BLI and EGGIM classification

#### Study design and patient selection

A multicenter cross-sectional study was proposed by the Portuguese Oncology Institute of Porto, Portugal, which had original proposed use of BLI technology for endoscopic diagnosis and grading of GIM. Selected European centers, recognized for their expertise in BLI technology and located in countries classified as low-risk for GC according to their ASR, were invited to participate. The patient recruitment period spanned from September 2021 to November 2023. During this period, individuals aged 18 years or older who had any clinical indications for un-

dergoing upper gastrointestinal endoscopy were proposed to be enrolled, regardless of whether it was their first endoscopy or a follow-up procedure. Exclusion criteria included any personal history of gastric surgery or gastric neoplasia, presence of contraindications to performing biopsies during the endoscopic procedure, and inability of individuals to provide informed consent. Comprehensive informed consent was obtained from all participants prior to their inclusion in the study.

### Endoscopic procedure

Prior to the procedure, all patients underwent pharyngeal anesthesia and/or were administered conscious sedation to ensure comfort and compliance. Endoscopic examinations were carried out at each center using Fujifilm EG-760ZHD endoscope in conjunction with the ELUXEO VP-7000 processor, by two fully-trained endoscopists proficient in BLI or NBI, with over 100 procedures annually. The endoscopists were allocated for each phase of the procedure through a random assignment process.

Initially, WLE examination was conducted by an endoscopist unaware of the patient’s prior GIM staging, during which an assistant documented all observed mucosal abnormalities. A comprehensive evaluation of the entire gastric mucosa was performed to assess presence and extent of GIM. Subsequently, a different, blinded endoscopist conducted the BLI examination, following the same approach. According to the EGGIM score, five distinct gastric areas were evaluated: the lesser curvature and the greater curvature of the antrum, the angular incisure, and the lesser and the greater curvature of the corpus [8]. These areas were each assigned a score: 0 in the absence of GIM, 1 for GIM presence in less than 30% of the area, and 2 for GIM encompassing more than 30% of the area. Both the per-area scores and the cumulative EGGIM score were accurately recorded by the assistant for each patient. ▶ **Fig. 1** shows endoscopic images of normal and GIM areas, under WLE and BLI vision.

A total of five biopsies were systematically performed during each procedure, focusing on the five specific gastric areas (lesser and greater curvature of the antrum, incisure, lesser and greater curvature of the corpus), as guided by the EGGIM results. In areas with an EGGIM score of 0, targeted biopsies were taken from normal-appearing gastric mucosa. In areas with an EGGIM score of 1 or 2, targeted biopsies were performed from GIM-detected sites. This approach ensured that all five gastric areas were sampled.

### Histopathological evaluation

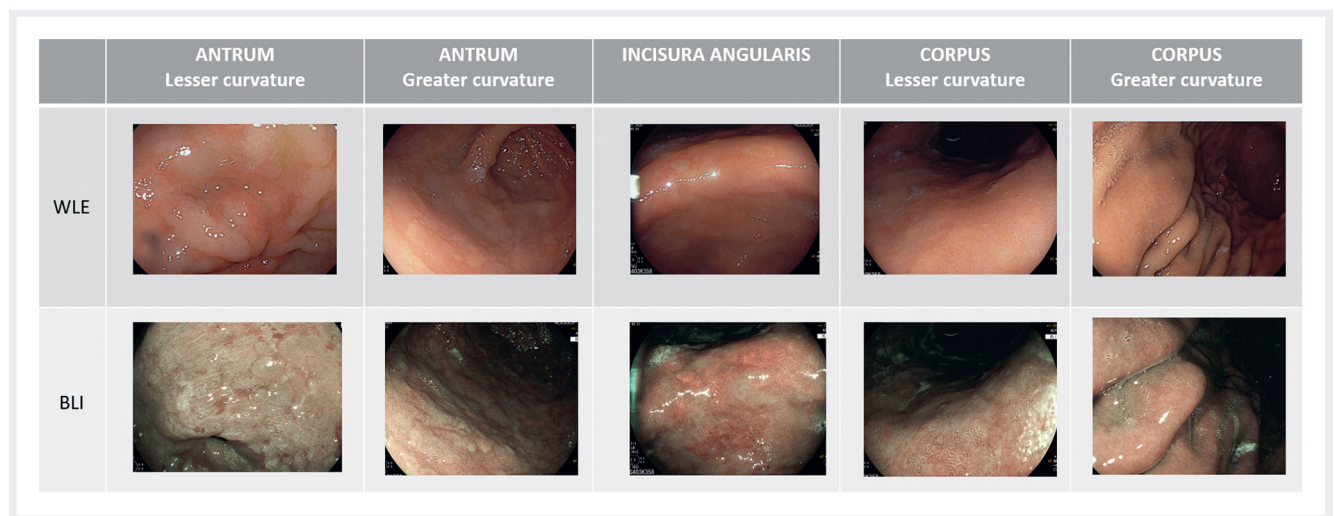
At each participating center, histopathological examinations were conducted by an expert gastrointestinal pathologist who was unaware of the endoscopic results. The final diagnosis was established based on the OLGIM system.

### Accuracy assessment

To evaluate diagnostic accuracy of WLE and BLI, two distinct approaches were employed: per-biopsy analysis and per-patient analysis. Per-biopsy analysis focused on precision of both technologies in identifying GIM at the biopsy level, focusing on their ability to detect GIM accurately. Per-patient analysis expanded the scope of assessment to the patient level, measuring overall accuracy of WLE and BLI in identifying individuals with GIM. In addition, through this approach, we assessed accuracy of the EGGIM score in detecting OLGIM stages III and IV.

### Definitions

Endoscopic findings were classified under WLE and BLI according to Pimenel-Nunes et al. [18] and Uedo et al. [19] as follows. Normal gastric mucosa was defined if it exhibited a regular vascular and glandular pattern, with regular microvascular and glandular pattern (oval morphology in the antral mucosa and circular morphology in the corpus mucosa). GIM was diagnosed if regular ridge or tubulovillous glandular pattern along with



▶ **Fig. 1** Comparative visualization of gastric intestinal metaplasia (GIM) and normal gastric mucosa under white light endoscopy (WLE) and blue light imaging (BLI). Examples of gastric intestinal metaplasia (GIM) areas visualized using WLE and BLI (columns 1–4) alongside normal gastric mucosa (column 5). The images illustrate different gastric regions: the antrum (lesser curvature and greater curvature), the angular incisure, and the corpus (lesser curvature and greater curvature). GIM areas appear as whitish-elevated regions with a tubulovillous pattern.

regular microvascular pattern was identified. Atrophic gastritis was recognized as diminishment of gastric folds in the corpus, presence of pale mucosa, and increased visible deeper vascular pattern. Dysplasia was suspected when a demarcated area with irregular or absent glandular and/or microvascular pattern was identified.

### Data collection

Each procedure performed was documented using an anonymized, standardized electronic case report form designed to capture all relevant data points. This electronic report was structured to include a wide range of information, systematically categorized into six main sections: 1) center data and inclusion date; 2) patient demographic characteristics; 3) family history of GC; 4) proton pump inhibitors, antiplatelet or non-steroidal anti-inflammatory medication; 5) procedure-related information; and 6) histological-related information. Following the conclusion of data collection, case report forms from all procedures and participating centers were integrated into a centralized database, enabling subsequent analytical phase.

### Statistical analysis

In Study 1, interobserver agreement was evaluated using Fleiss' kappa statistic, weighted kappa statistic, and the proportion of agreement, with 95% confidence intervals (CIs) calculated to gauge agreement levels. According to Landis and Koch [20], agreement levels were categorized as follows: almost perfect (0.81 to 1.00), substantial (0.61 to 0.80), moderate (0.41 to 0.60), fair (0.21 to 0.40), slight (0.01 to 0.20), and no agreement (<0.01).

In Study 2, sample size was determined to confidently detect the previously reported 15% difference in sensitivities between BLI and WLE for GIM detection, using histology as the gold standard. Based on global GIM rates, and considering its geographical variations as well as the influence of *Helicobacter pylori* (*H. pylori*) infection [21, 22], we estimated a GIM prevalence of 20% and an OLGIM III-IV prevalence of 5%. To achieve 80% statistical power with a 5% margin of error, we calculated that a sample size of 270 patients (corresponding to 1350 biopsies) would be required. This calculation, performed using EPI INFO, considered the study's multicenter nature and the expected variability in GIM prevalence among centers. Descriptive statistical analysis was conducted to delineate the clinical profile of the patient cohort, presenting data as absolute numbers, percentages, and medians with interquartile ranges (IQRs).

Diagnostic accuracy of WLE and BLI, as analyzed using contingency tables along with the chi-square or Fisher's exact test as appropriate. Sensitivity and specificity comparisons across endoscopic technologies were made using the McNemar test. The receiver operating characteristic (ROC) curve identified the optimal EGGIM threshold for diagnosing OLGIM stages III/IV. Statistical significance was set at a value of  $P < 0.05$ . Analyses were performed using SPSS software (version 24; IBM) and MedCalc Software.

## Results

### Reliability of BLI and EGGIM score

BLI significantly improved overall interobserver agreement of endoscopic findings, evidenced by a Fleiss kappa of 0.4 (95% CI: 0.3–0.5), compared with the 0.2 (95% CI: 0.2–0.3) obtained with WLE. Diagnostic confidence levels for each endoscopist varied, with a range of 60.0% to 83.0% for WLE and 73.0% to 93.0% for BLI. When endoscopists felt confident in their assessments, the proportion of agreement on presence or absence of GIM was observed to be 55.0% for complete consensus (all observers agreed on the diagnosis) and 90.0% for partial consensus (only one observer's diagnosis differed from the others). Notably, regarding the level of concordance for EGGIM score under BLI, the results showed a weighted kappa of 0.7 (95% CI 0.5–0.9), indicating substantial agreement.

### Accuracy of BLI for diagnosis and staging of GIM

A total of 102 patients and 510 biopsies from two Italian centers were included in this analysis. The patient cohort comprised 50 individuals from the first center and 52 from the second. ► **Table 1** shows baseline characteristics for the included patients. The majority of patients were female (65.7%) and had a median age of 60 years (IQR 44.8–73.5). Prevalence of GIM in the cohort was 43.1%, and in most cases (95.5%), it was endoscopically detected. EGGIM score distribution revealed a predominance of lower scores, with EGGIM 0 being the most frequent, observed in 38 patients, followed by EGGIM 4 and 3, with 18 and 14 patients respectively. Similarly, the OLGIM system showed a predominance of absence of GIM (0) or low stages (I-II) of OLGIM (58 individuals and 39 individuals, respectively).

### Per-biopsy analysis

A significant enhancement in sensitivity was observed with BLI, registering a 63% increase (from 50% with WLE, to 82% with BLI;  $P < 0.01$ ) (► **Table 2**). Conversely, the observed improvement in sensitivity was offset by a 12% decrease in specificity, dropping from 92% with WLE to 81% with BLI ( $P < 0.01$ ). This decline in specificity was primarily attributed to a notable false-positive rate of 18.9%. Notably, a significant proportion of these false positives (56/76) were observed in biopsies from the antrum and incisura.

### Per-patient analysis

Among the 102 individuals assessed, BLI correctly identified GIM in 95.5% of the cases (42 of 44 patients; 95% CI 84.5–99.4), demonstrating a significantly higher sensitivity than WLE, which identified GIM in 68.2% of cases (30 of 44 patients; 95% CI 52.4–81.4;  $P < 0.01$ ) (► **Table 2**). Despite BLI's superior sensitivity, it exhibited a reduced specificity of 62.1% (95% CI: 48.4–74.5), compared with 79.3% for WLE (95% CI: 66.7–88.8;  $P = 0.01$ ), reflecting a higher incidence of false positives, which were mainly identified as focal/moderate GIM (20/22).

A pronounced tendency for overestimating both presence and severity of GIM was particularly evident in patients with lower EGGIM scores (1–4), where 20 of 48 patients were histo-

**► Table 1** Baseline characteristics and endoscopic results from the included patients.

| N = 102  |                |
|--|----------------|
| Gender, n (%)  |                |
| ▪ Male   | 34 (33.3)      |
| ▪ Female   | 67 (65.7)      |
| Age, median (IQR)  | 60 (44.8–73.5) |
| First-degree familial history of gastric cancer, n (%)         | 11 (10.8)      |
| Chronic medication, n (%)                                      |                |
| ▪ Proton pump inhibitors                                       | 29 (28.4)      |
| ▪ Antiplatelet/nonsteroidal anti-inflammatory                  | 11 (10.8)      |
| Indication for upper gastrointestinal endoscopy, n (%)         |                |
| ▪ Screening  | 3 (2.9)        |
| ▪ Dyspepsia  | 28 (27.5)      |
| ▪ Gastroesophageal reflux disease                              | 15 (14.7)      |
| ▪ Dysphagia  | 3 (2.9)        |
| ▪ Persistent nausea or vomiting                                | 1 (1.0)        |
| ▪ Anemia   | 5 (4.9)        |
| ▪ Significant weight loss                                      | 2 (2.0)        |
| ▪ Other  | 45 (44.1)      |
| Prevalence of gastric intestinal metaplasia, n (%)             | 44 (43.1)      |
| ▪ Endoscopic diagnosis of gastric intestinal metaplasia, n (%) | 42 (41.2)      |
| Helicobacter pylori active infection                           |                |
| ▪ Positive   | 14 (13.7)      |
| ▪ Negative   | 88 (86.3)      |

logically found to have no GIM (► **Table 3**). In this group, a single patient was misclassified as having focal/moderate GIM, while histological evidence placed it in a more advanced OLGIM stage. Another 10 patients corresponding to EGGIM 5–10 were reclassified to lower OLGIM stages. Notably, among the three instances of underdiagnosis, two were classified within OLGIM stages I-II. Among the 14 patients with active *H. pylori* infection, seven had histological GIM. Misclassification of EGGIM score occurred in half of these patients. Specifically, three patients with OLGIM stage 0 were categorized as EGGIM 1–4 and four patients with OLGIM stages I-II were classified as EGGIM 5–10.

In evaluating the effectiveness of endoscopic diagnosis for advanced stages of GIM (OLGIM III/IV) using the EGGIM scale, the area under the ROC curve (AUC) was found to be 0.9 (95% CI 0.8–1.0). The optimal threshold value for the EGGIM to diagnose OLGIM stages III/IV was determined to be greater than 4 (► **Table 4**). At this cut-off, sensitivity and specificity of EGGIM were 80.0% (95% CI 28.36–99.5) and 87.6% (95% CI 79.4–94.4), respectively, and the positive likelihood ratio and negative like-

**► Table 1** (Continuation)

| N = 102  |           |
|--|-----------|
| EGGIM score system, n (%)  |           |
| ▪ 0  | 38 (37.3) |
| ▪ 1  | 9 (8.8)   |
| ▪ 2  | 7 (6.9)   |
| ▪ 3  | 14 (13.7) |
| ▪ 4  | 18 (17.6) |
| ▪ 5  | 8 (7.8)   |
| ▪ 6  | 5 (4.9)   |
| ▪ 7  | 1 (1.0)   |
| ▪ 8  | 1 (1.0)   |
| ▪ 9  | 1 (1.0)   |
| ▪ 10   | 0 (0.0)   |
| OLGIM score system   |           |
| ▪ 0  | 58 (56.9) |
| ▪ I  | 19 (18.6) |
| ▪ II   | 20 (19.6) |
| ▪ III  | 4 (3.9)   |
| ▪ IV   | 1 (1.0)   |
| EGGIM, endoscopic grading of gastric intestinal metaplasia; IQR, interquartile range; OLGIM, operative link of gastric, intestinal metaplasia. |           |

lihood ratio were 6.5 (95% CI 3.3–12.9) and 0.2 (95% CI 0.0–1.3), respectively.

## Discussion

Our study presents a comprehensive evaluation of the reliability and accuracy of BLI and EGGIM score for diagnosis and staging of GIM in a country with low risk for GC. The primary novelty of this study lies in external validation of the EGGIM system using BLI, particularly in a population that has been understudied in terms of diagnostic accuracy with enhanced imaging technologies. Our results highlight the remarkable diagnostic capability of the EGGIM system for detecting advanced GIM stages (OLGIM III/IV), demonstrated by an AUC of 0.9 (95% CI 0.8–1.0). Specifically, the best accuracy measures for detecting these GIM stages were associated with a EGGIM score greater than 4, achieving a sensitivity of 80.0% and a specificity of 87.6%. These findings closely match the diagnostic accuracy observed with NBI [8], where the highest sensitivity and specificity similarly validated an EGGIM cut-off greater than 4 (89.4% and 94.6%, respectively) and the AUC for identifying OLGIM III/IV stages was reported as 0.96 (95% CI 0.93–0.98). Although global results are slightly higher with NBI, our study was conducted in a low-risk setting for GC where GIM incidence is expected to be lower. These epidemiological differences may in-

**Table 2** Diagnosis accuracy of white light endoscopy and blue light imaging for diagnosis of gastric intestinal metaplasia.

|                      |     | Sensitivity, % (95% IC) | Specificity, % (95% IC) | PLR, % (95% IC) | NLR, % (95% IC) |
|----------------------|-----|-------------------------|-------------------------|-----------------|-----------------|
| Per-biopsy analysis  |     |                         |                         |                 |                 |
| Center 1             | WLE | 54.0 (41.0–66.6)        | 94.1 (89.7–97.0)        | 9.2 (5.0–17.0)  | 0.5 (0.4–0.6)   |
|                      | BLI | 81.0 (69.1–89.8)        | 93.6 (89.1–96.6)        | 12.6 (7.2–22.1) | 0.2 (0.1–0.3)   |
| Center 2             | WLE | 45.5 (30.4–61.2)        | 89.8 (85.0–93.5)        | 4.5 (2.6–7.4)   | 0.6 (0.5–0.9)   |
|                      | BLI | 84.1 (69.9–93.4)        | 70.4 (63.8–76.4)        | 2.8 (2.2–3.6)   | 0.2 (0.1–0.5)   |
| TOTAL                | WLE | 50.5 (40.6–60.3)        | 91.8 (88.7–94.3)        | 6.16 (4.2–9.0)  | 0.5 (0.4–0.7)   |
|                      | BLI | 82.2 (73.7–89.0)        | 81.1 (77.0–84.8)        | 4.4 (3.5–5.4)   | 0.2 (0.2–0.3)   |
| Per-patient analysis |     |                         |                         |                 |                 |
| Center 1             | WLE | 66.7 (44.7–84.3)        | 84.6 (65.1–95.6)        | 4.3 (1.7–11.2)  | 0.4 (0.2–0.7)   |
|                      | BLI | 100.0 (85.8–100.0)      | 80.8 (60.7–93.5)        | 5.2 (2.4–11.4)  | 0.0             |
| Center 2             | WLE | 70.0 (45.7–88.1)        | 75.0 (56.6–88.5)        | 2.8 (1.4–5.4)   | 0.4 (0.2–0.8)   |
|                      | BLI | 90.0 (68.3–98.8)        | 46.9 (29.1–65.3)        | 1.7 (1.2–2.4)   | 0.2 (0.1–0.8)   |
| TOTAL                | WLE | 68.2 (52.4–81.4)        | 79.3 (66.7–88.8)        | 3.3 (1.9–5.7)   | 0.4 (0.3–0.6)   |
|                      | BLI | 95.5 (84.5–99.4)        | 62.1 (48.4–74.5)        | 2.5 (1.8–3.5)   | 0.1 (0.0–0.3)   |

BLI, blue light imaging; NLR, negative likelihood ratio; PLR, positive likelihood ratio; WLE, white light endoscopy.

**Table 3** Correlation between endoscopic grading of gastric intestinal metaplasia and operative link of gastric intestinal metaplasia scores.

|       |        | EGGIM     |           |           |
|-------|--------|-----------|-----------|-----------|
|       |        | 0         | 1–4       | 5–10      |
| OLGIM | 0      | 36 (62.1) | 20 (34.5) | 2 (3.5)   |
|       | I–II   | 2 (5.1)   | 27 (69.2) | 10 (25.6) |
|       | III–IV | 0 (0.0)   | 1 (20.0)  | 4 (80.0)  |

EGGIM, endoscopic grading of gastric intestinal metaplasia; OLGIM, operative link of gastric intestinal metaplasia.

fluence the accuracy of the technologies used, which further underscores the importance of validating BLI in such contexts. In addition, the goal of this study is not to replace NBI but to validate BLI as a robust diagnostic tool for GIM detection and staging, ensuring that where BLI is available – an increasingly common technology – there are validated data to support definitive diagnoses. Further research comparing the overall accuracy of NBI and BLI within the same epidemiological context would provide valuable insights and enhance our understanding of potential differences in their diagnostic efficacy. In addition, our analysis demonstrated a notable precision in categorizing patients with an EGGIM score of 0, with only a 5% misclas-

**Table 4** Accuracy of endoscopic grading of gastric intestinal metaplasia score according to the different thresholds for diagnosis operative link of gastric intestinal metaplasia score III/IV.

| EGGIM | Sensitivity (95% CI) | Specificity (95% CI) | LR+              | LR-           |
|-------|----------------------|----------------------|------------------|---------------|
| >1    | 100.0 (47.8–100.0)   | 48.5 (38.2–58.8)     | 1.9 (1.6–2.4)    | 0.0           |
| >2    | 100.0 (47.8–100.0)   | 55.7 (45.2–65.8)     | 2.3 (1.8–2.8)    | 0.0           |
| >3    | 80.00 (28.4–99.5)    | 69.1 (58.9–78.1)     | 2.6 (1.5–4.4)    | 0.3 (0.1–1.7) |
| >4    | 80.0 (28.4–99.5)     | 87.6 (79.4–94.4)     | 6.5 (3.3–12.9)   | 0.2 (0.0–1.3) |
| >5    | 60.0 (14.7–94.7)     | 94.9 (88.4–98.3)     | 11.6 (3.8–35.5)  | 0.4 (0.1–1.2) |
| >6    | 40.0 (5.3–85.3)      | 99.0 (94.4–100.0)    | 38.8 (4.2–359.4) | 0.6 (0.3–1.2) |
| >7    | 0.2 (0.5–71.6)       | 99.0 (94.4–100.0)    | 19.4 (1.4–267.0) | 0.8 (0.5–1.3) |
| >8    | 0.0 (0.0–52.2)       | 99.0 (94.4–100.0)    | 0.0              | 1.0 (0.9–1.0) |
| >10.0 | 0.0 (0.0–52.2)       | 100.0 (96.3–100.0)   | 0.0              | 1.0 (1.0–1.0) |

CI, confidence interval; EGGIM, endoscopic grading of gastric intestinal metaplasia; LR+, positive likelihood ratio; LR-, negative likelihood ratio.

sification rate. Importantly, these misclassified cases still corresponded to lower OLGIM stages (I–II), indicating a minor discrepancy in accurately ruling out presence of advanced GIM. Although current guidelines strongly recommend performing biopsies during a first-time upper gastrointestinal endoscopy [3], our finding supports the notion that endoscopic biopsies may be unnecessary or at least sent in a single vial in cases where GIM is absent [23], particularly in populations with a low prevalence of extensive GIM. However, potential for GIM overdiagnosis in patients presenting with lower EGGIM scores (1–4), coupled with instances in which the EGGIM score might suggest advanced disease (5–10), emphasizes the continued importance of histological confirmation in these scenarios. Interestingly, overestimation of EGGIM scores ranging from 1 to 4 has also been documented in conditions such as autoimmune atrophic gastritis with pseudopyloric metaplasia, a factor that could confound accurate diagnosis of GIM [24]. Although our study did not delve into this specific scenario in depth, it highlights an area for future research and consideration in clinical practice. In our cohort, even though the EGGIM score was overestimated in half of patients with active *H. pylori* infection, establishing a potential correlation between inflammatory alterations induced by *H. pylori* and misclassification of EGGIM was constrained by the small sample size. Overrating of EGGIM scores has also been documented with NBI, albeit to a lesser extent, and no association with *H. pylori* infection as a confounding factor was found [8]. The greater overestimation observed in our study may be due not to the technology itself, but rather, to other cofounding factors, such as inflammation from other etiologies, or the varying levels of training in EGGIM estimation. Future research addressing these possible discrepancies is needed to arrive at definite conclusions.

Besides the high accuracy of EGGIM score in combination with BLI for detecting OLGIM III/IV, the reliability analysis for this approach demonstrated a substantial level of concordance between observers, as indicated by a weighted kappa value of 0.7. These findings imply that EGGIM score, when utilized alongside BLI, is not only straightforward to understand and apply but also underscores its capacity for widespread use, and suggests its utility in offering a uniform method for evaluating individual risk profiles effectively. Conversely, when considering the level of agreement in the overall endoscopic diagnosis, an improvement was noted with BLI over WLE technology, although to a lesser extent (Fleiss kappa 0.4 for BLI vs. 0.2 for WLE). The enhancement provided by BLI in endoscopic diagnosis was also documented in a recent study performed by our group, evidenced by a weighted kappa of 0.8 for BLI in bright mode, compared with a weighted kappa of 0.4 for WLE [16].

The additional value of BLI was also evidenced in absence of the application of EGGIM score. The significant improvement in sensitivity observed with use of BLI over WLE marks a critical advancement in endoscopic diagnosis, evidenced by a 62.7% increase in per-biopsy analysis and a 40.0% increase in per-patient analysis. Despite the increase in sensitivity, BLI demonstrated a lower specificity compared with WLE (81.1% vs. 91.8% in per-biopsy analyses; 62.1% vs. 79.3% in per-patient analysis) due to a higher false-positive rate, primarily observed in

the antrum and incisura. Marked differences in this rate were observed between the two centers (6% vs. 29%), which may correlate with the varying levels of expertise in GIM estimation among endoscopists, as well as presence of possible confounding factors within the cohort. However, we did not investigate the specific reasons for these differences. We theorize that the overestimation mainly observed in antrum and incisura is likely attributed to similarities between the specific glandular architecture in these gastric areas and the subtle changes in the glandular pattern observed at the initial stages of GIM. From a practical clinical perspective, it is crucial for endoscopists to integrate insights from both BLI and WLE in patient evaluation. The combination of these modalities should not be viewed as a choice between alternatives, but rather, as a synergistic strategy to capitalize on the distinct advantages that each technique offers.

Several limitations warrant careful consideration. First, our sample encompasses the full spectrum of GIM, albeit with a reduced prevalence of extensive GIM. Although this could be considered a potential bias, the OLGIM distribution accurately reflects the natural prevalence of GIM in the studied population, inherently resulting in more cases with lower EGGIM scores. This pattern is consistent with the scenario prevalent in the majority of Western countries, which are categorized as low-risk regions for GC [25]. This delineation highlights the relevance of the study to these populations while acknowledging the potential limitations in generalizability to areas with higher incidences of GIM and GC. Second, we did not consistently perform targeted biopsies from both GIM-positive and GIM-negative sites within the same gastric area due to logistical constraints. However, our protocol ensured systematic sampling of all five gastric areas, with targeted biopsies taken from either GIM-positive or normal-appearing mucosa in each region, and these were compared with histological results. This approach still provided valuable data through comprehensive coverage of the gastric mucosa, making the findings robust and informative for understanding GIM patterns across the stomach. Third, despite the observed improvement in concordance rates with BLI, the achieved concordance level remained at a moderate tier. To assess interobserver agreement, our analysis relied on still images, which do not fully capture the dynamic nature of actual procedures. In real-time examinations, the context is enriched by the ability to have a close and more distant vision of the areas of interest, which can contextualize the mucosal alterations, providing a comprehensive view of the stomach. This observation implies that the most effective analysis of these data may not be fully realized through exclusive use of still images, but by utilization of real-time videos. Fourth, although BLI significantly enhanced sensitivity, it also led to a high rate of false positives in one of the two centers, particularly in the antrum and incisura areas. Nevertheless, given the characteristics of these false positives, their impact on patient management protocols is minimal, ensuring that surveillance remains focused on those with a higher-risk profile. Furthermore, BLI is understood to complement WLE, with its primary advantage being marked increase in sensitivity. Fifth, despite substantial agreement on the EGGIM score, this calculation was based on assess-

ments by only two endoscopists. Although an educational session on EGGIM calculation was conducted for participant endoscopists before starting the study, differences in EGGIM estimation were not assessed before and after the training. Future research should investigate the agreement among endoscopists with varying levels of experience in EGGIM estimation and evaluate the impact of training by comparing agreement before and after EGGIM application training.

## Conclusions

In conclusion, our research pioneers the validation of BLI for the EGGIM classification, demonstrating its consistency and remarkable precision in identifying individuals at advanced GIM stages. Moreover, BLI's high sensitivity underscores its viability as an effective screening tool that can complement WLE, ensuring more precise evaluation of GC risk. Although BLI offers a refined approach for assessing individual risk profiles, the potential for overdiagnosis within EGGIM scores of 1 to 4 calls for further exploration to optimize its application.

## Conflict of Interest

Mário Dinis-Ribeiro has financial relationships with a commercial entity producing health care-related products and/or services relevant to this article: Companies: Medtronic, Roche Type of relationship: Consultancy The remaining authors have no conflicts of interest to declare

## Clinical trial

ClinicalTrials.gov (<http://www.clinicaltrials.gov/>)  
Registration number (trial ID): NCT04768218  
Type of Study: Multi-center, cross-sectional study

## References

- [1] Correa P, Haenszel W, Cuello C et al. A model for gastric cancer epidemiology. *Lancet* 1975; 2: 58–60 doi:10.1016/s0140-6736(75)90498-5
- [2] Capelle LG, de Vries AC, Haringsma J et al. The staging of gastritis with the OLGA system by using intestinal metaplasia as an accurate alternative for atrophic gastritis. *Gastrointest Endosc* 2010; 71: 1150–1158
- [3] Pimentel-Nunes P, Libanio D, Marcos-Pinto R et al. Management of epithelial precancerous conditions and lesions in the stomach (MAPS II): European Society of Gastrointestinal Endoscopy (ESGE), European Helicobacter and Microbiota Study Group (EHMSG), European Society of Pathology (ESP), and Sociedade Portuguesa de Endoscopia Digestiva (SPED) guideline update 2019. *Endoscopy* 2019; 51: 365–388 doi:10.1055/a-0859-1883
- [4] Ang TL, Pittayanon R, Lau JY et al. A multicenter randomized comparison between high-definition white light endoscopy and narrow band imaging for detection of gastric lesions. *Eur J Gastroenterol Hepatol* 2015; 27: 1473–1478
- [5] Panteris V, Nikolopoulou S, Lountou A et al. Diagnostic capabilities of high-definition white light endoscopy for the diagnosis of gastric intestinal metaplasia and correlation with histologic and clinical data. *Eur J Gastroenterol Hepatol* 2014; 26: 594–601
- [6] Pimentel-Nunes P, Libanio D, Lage J et al. A multicenter prospective study of the real-time use of narrow-band imaging in the diagnosis of premalignant gastric conditions and lesions. *Endoscopy* 2016; 48: 723–730 doi:10.1055/s-0042-108435
- [7] Rodriguez-Carrasco M, Esposito G, Libanio D et al. Image-enhanced endoscopy for gastric preneoplastic conditions and neoplastic lesions: a systematic review and meta-analysis. *Endoscopy* 2020; 52: 1048–1065 doi:10.1055/a-1205-0570
- [8] Esposito G, Pimentel-Nunes P, Angeletti S et al. Endoscopic grading of gastric intestinal metaplasia (EGGIM): a multicenter validation study. *Endoscopy* 2019; 51: 515–521 doi:10.1055/a-0808-3186
- [9] Chen H, Liu Y, Lu Y et al. Ability of blue laser imaging with magnifying endoscopy for the diagnosis of gastric intestinal metaplasia. *Lasers Med Sci* 2018; 33: 1757–1762
- [10] Chen H, Wu X, Liu Y et al. Blue laser imaging with acetic acid enhancement improved the detection rate of gastric intestinal metaplasia. *Lasers Med Sci* 2019; 34: 555–559
- [11] Kimura-Tsuchiya R, Dohi O, Fujita Y et al. Magnifying endoscopy with blue laser imaging improves the microstructure visualization in early gastric cancer: Comparison of magnifying endoscopy with narrow-band imaging. *Gastroenterol Res Pract* 2017; 2017: 8303046
- [12] Dohi O, Yagi N, Yoshida S et al. Magnifying blue laser imaging versus magnifying narrow-band imaging for the diagnosis of early gastric cancer: A prospective, multicenter, comparative study. *Digestion* 2017; 96: 127–134 doi:10.1159/000479553
- [13] Dohi O, Yagi N, Naito Y et al. Blue laser imaging-bright improves the real-time detection rate of early gastric cancer: a randomized controlled study. *Gastrointest Endosc* 2019; 89: 47–57 doi:10.1016/j.gie.2018.08.049
- [14] 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
- [15] Takayama S, Dohi O, Naito Y et al. Diagnostic ability of magnifying blue light imaging with a light emitting diode light source for early gastric cancer: A prospective comparative study. *Digestion* 2021; 102: 580–589
- [16] Castro R, Rodriguez M, Libanio D et al. Reliability and accuracy of blue light imaging for staging of intestinal metaplasia in the stomach. *Scand J Gastroenterol* 2019; 54: 1301–1305 doi:10.1080/00365521.2019.1684555
- [17] von Elm E, Altman DG, Egger M et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg* 2014; 12: 1495–1499 doi:10.1016/j.jclinepi.2007.11.008
- [18] Pimentel-Nunes P, Dinis-Ribeiro M, Soares JB et al. A multicenter validation of an endoscopic classification with narrow band imaging for gastric precancerous and cancerous lesions. *Endoscopy* 2012; 44: 236–246
- [19] Uedo N, Yao K. Endoluminal diagnosis of early gastric cancer and its precursors: bridging the gap between endoscopy and pathology. *Adv Exp Med Biol* 2016; 908: 293–316
- [20] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174
- [21] Altayar O, Davitkov P, Shah SC et al. AGA Technical Review on Gastric Intestinal Metaplasia-Epidemiology and Risk Factors. *Gastroenterology* 2020; 158: 732–744 e716 doi:10.1053/j.gastro.2019.12.002
- [22] Hooi JKY, Lai WY, Ng WK et al. Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-Analysis. *Gastroenterology* 2017; 153: 420–429 doi:10.1053/j.gastro.2017.04.022



- [23] Castro R, Esposito G, Libanio D et al. A single vial is enough in the absence of endoscopic suspected intestinal metaplasia – less is more! *Scand J Gastroenterol* 2019; 54: 673–677
- [24] Dilaghi E, Esposito G, Pivetta G et al. Endoscopic diagnosis of gastric intestinal metaplasia in patients with autoimmune gastritis using narrow-band imaging: does pseudopyloric metaplasia muddy the waters? *Endosc Int Open* 2022; 10: E434–E440
- [25] Bray F, Laversanne M, Sung H et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2024; 74: 229–263 doi:10.3322/caac.21834