Narrow band imaging and serology in the assessment of premalignant gastric pathology

Jonathan R. White, Sarmed S. Sami, Dona Reddiar, Jayan Mannath, Jacobo Ortiz-Fernández-Sordo, Sabina Beg, Robert Scott, Prarthana Thiagarajan, Saqib Ahmad, Adolfo Parra-Blanco, Madhavi Kasi, Emmanouil Telakis, Alyshah A. Sultan, Jillian Davis, Adam Figgins, Philip Kaye, Karen Robinson, John C. Atherton & Krish Ragunath


To link to this article: https://doi.org/10.1080/00365521.2018.1542455
Narrow band imaging and serology in the assessment of premalignant gastric pathology

Jonathan R. White a,b, Sarmed S. Sami c, Dona Reddian a,b, Jayan Mannath d, Jacobo Ortiz-Fernández-Sordo a,b, Sabina Beg a,b, Robert Scott a,b, Prarthana Thiagarajan a,b, Saqib Ahmad e, Adolfo Parra-Blanco a,b, Madhavi Kasi a,b, Emmanouil Telakis f, Alyshah A. Sultan g, Jillian Davis h, Adam Figgins h, Philip Kaye h, Karen Robinson a,b, John C. Atherton a,b and Krish Ragunatha b,

1NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and The University of Nottingham, Nottingham, UK; 2Nottingham Digestive Diseases Centre, The University of Nottingham, Nottingham, UK; 3Mayo Clinic Division of Gastroenterology and Hepatology, Rochester, MN, USA; 4Department of Gastroenterology, University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK; 5Sherwood Forest Hospitals NHS Foundation Trust, Kings Mill Hospital, Nottinghamshire, UK; 6Department of Gastroenterology, Hellenic Red Cross Hospital, Athens, Greece; 7Research Institute for Primary Care and Health Sciences, Primary Care Sciences, Keele University, Staffordshire, UK; 8Department of Pathology, Nottingham University Hospitals NHS Trust, Queen’s Medical Centre Campus, Nottingham, UK.

ABSTRACT

Background: Patient outcomes in gastric adenocarcinoma are poor due to late diagnosis. Detecting and treating at the premalignant stage has the potential to improve this. Helicobacter pylori is also a strong risk factor for this disease.

Aims: Primary aims were to assess the diagnostic accuracy of magnified narrow band imaging (NBI-Z) endoscopy and serology in detecting normal mucosa, H. pylori gastritis and gastric atrophy. Secondary aims were to compare the diagnostic accuracies of two classification systems using both NBI-Z and white light endoscopy with magnification (WLE-Z) and evaluate the inter-observer agreement.

Methods: Patients were prospectively recruited. Images of gastric mucosa were stored with histology and serum for IgG H. pylori and Pepsinogen (PG) I/II ELISAs. Blinded expert endoscopists agreed on mucosal pattern. Mucosal images and serological markers were compared with histology. Kappa statistics determined inter-observer variability for randomly allocated images among four experts and four non-experts.

Results: 116 patients were prospectively recruited. Diagnostic accuracy of NBI-Z for determining normal gastric mucosa was 0.87 (95% CI 0.82–0.92), H. pylori gastritis 0.65 (95% CI 0.55–0.75) and gastric atrophy 0.88 (95% CI 0.81–0.94). NBI-Z was superior to serology at detecting gastric atrophy: NBI-Z gastric atrophy 0.88 (95% CI 0.81–0.94) vs PG/I/II ratio < 3 0.74 (95% CI 0.62–0.85) p < .0001. Overall NBI-Z was superior to WLE-Z in detecting disease using two validated classifications. Inter-observer agreement was 0.63 (95% CI 0.51–0.73).

Conclusions: NBI-Z accurately detects changes in the GI mucosa which currently depend on histology. NBI-Z is useful in the detection of precancersous conditions, potentially improving patient outcomes with early intervention to prevent gastric cancer.

Introduction

Helicobacter pylori (H. pylori) colonizes the gastric mucosa of approximately 50% of the world’s population, although the prevalence varies between countries the infection rates are higher in developing countries [1,2]. H. pylori infects individuals during childhood and typically persists lifelong in the absence of effective eradication therapy [3]. In 15% of individuals infection leads to serious complications such as peptic ulcer disease, distal gastric adenocarcinoma or primary gastric mucosa associated lymphoid tissue (MALT) lymphoma [4]. Other conditions associated with H. pylori include iron deficiency anemia, gastric atrophy and idiopathic thrombocytopenia purpura [5,6]. H. pylori infection has also been suggested to be a significant contributor to the hygiene hypothesis and the development of a healthy immune system [4,7].

Gastric atrophy occurs as a result of a chronic gastric mucosal inflammation, usually due to H. pylori infection, leading to loss of specialized glandular cells and replacement with an intestinal type surface and fibrous tissue [8]. This may then progress to intestinal metaplasia then dysplasia and finally intestinal type gastric cancer. The development from normal mucosa to gastric cancer is gradual and in a stepwise manner [9]. The annual incidence of progression from intestinal metaplasia to gastric cancer varies from 0.1% to 0.9% [10]. The annual incidence increases to 6% in dysplasia [11,12]. Gastric atrophy is, therefore, considered a pre-cancerous condition and current European guidelines...
recommend endoscopic surveillance of these high risk patients with severe atrophy [13]. Unfortunately, at present these premalignant mucosal changes are often disregarded in routine clinical practice or result in variable surveillance frequency [14]. Gastric cancer is ranked the fifth most common malignancy worldwide [15] and is the third most common cause of cancer related death [16]. In order to reduce this associated mortality, one possibility is to detect premalignant disease using advanced endoscopy techniques. However, due to variety in endoscopic techniques and mucosal classifications the diagnosis is often dependent on histology.

High definition magnified white light endoscopy (WLE-Z) allows for detailed assessment of mucosal pit and vascular pattern, magnifying images by factors greater than 100 with resolutions smaller than 7.9 μm [17–19]. Narrow band imaging (NBI) relies on specific wavelengths of light to produce a sharper contrast between mucosal and vascular structures enhancing detail [20,21]. NBI endoscopes enable the endoscopist to easily switch from WLE to NBI to optimize visualization and sampling.

Standard WLE mucosal appearances often correlate poorly with histology and there is no widely accepted consensus on the macroscopic appearance, therefore, clinicians are reliant on histology which is dependent on multiple factors such as biopsy area, biopsy numbers and experience of the histopathologist [22,23]. Anagnostopoulos et al proposed a classification using WLE-Z which was highly accurate for detecting normal, _H. pylori_ gastritis and gastric atrophy with sensitivity and specificity of 90–100%. Using this classification a normal gastric corpus microvasculature consists of a honeycomb type subepithelial capillary network (SECN) with a regular arrangement of collecting venules and round pits. In the inflamed corpus, the normal SECN pattern and collecting venules are lost with enlarged white pits surrounded by erythema. Gastric atrophy is characterized by loss of the normal SECN and round pits, with an irregular distribution of collecting venules [18]. Pimentel et al described a more detailed NBI classification that also included intestinal metaplasia and dysplasia with accuracy of 84% and 95% respectively [24]. The light blue crest (blue-whitish patchy reflections sited on the epithelial margins) appearance in magnified NBI (NBI-Z) of gastric mucosa was first described by Uedo et al and correlated well to intestinal metaplasia with an accuracy of 91% [25]. White Opaque Substance (WOS) is also sometimes associated with epithelial tumors and intestinal metaplasia [26]. The use of NBI in the detection of these conditions was also accurate in other studies but due to study limitations reported results vary [27–30].

IgG ELISA serology detects _H. pylori_ antibodies and is relatively cheap and noninvasive but a positive test cannot distinguish between current and previously treated infection. Accuracy is variable but some commercial kits report accuracy of greater than 90% [31] with the benefit that results are not affected by proton pump inhibitor (PPI) treatment.

Serum pepsinogen can be used to predict the extent of gastric atrophy. Pepsinogen I (PGI) and Pepsinogen II (PGII) are usually released from secretory cells found in the gastric mucosa. With the progression to atrophy this causes the loss of these secreting cells and reduces pepsinogen levels. PGI is more affected by this and thus the ratio is decreased further. Low pepsinogen I or pepsinogen I/I ratio less than 3 can signify moderate to severe corpus atrophy [13]. For estimating gastric atrophy extent studies have described a huge variation in results with sensitivity and specificity from 9.4% to 92.3% and 9.9% to 100% respectively [13,32–34]. Serological markers of gastric atrophy are not routinely used in Western countries.

No study has yet evaluated the utility of both magnified narrow band imaging and serology in the detection of premalignant gastric pathology. Patients referred for endoscopic investigation for iron deficiency anemia were selected for this study due to the association with both _H. pylori_ and gastric atrophy [6,35,36]. The Anagnostopoulos et al classification was adapted to use NBI-Z rather than WLE-Z in this study and termed the Nottingham classification.

The primary aim of this pilot study was to assess the diagnostic accuracy of magnified NBI endoscopy and serology in diagnosing normal mucosa, _H. pylori_ gastritis and gastric atrophy using the Nottingham classification. Secondary aims were to compare the diagnostic accuracies of two validated classification systems using both NBI-Z and WLE-Z and evaluate the inter-observer agreement.

**Material and methods**

**Participants and clinical samples**

150 adult patients (18–85 yrs) attending Nottingham University Hospitals (NUH) NHS Trust for a diagnostic gastroscopy as part of their investigation into iron deficiency anemia were prospectively recruited between August 2010 and December 2014. Derbyshire Research Ethical Committee (REC Ref: 10/H0401/33) approved the protocol and written informed consent was gained. Patients with low hemoglobin (<130 g/l men & <120 g/l women) and either a low MCV (<84 fl) or low ferritin (<25 mcg/l men & <13 mcg/l women) were recruited. Patients taking anticoagulants and proton pump inhibitors or individuals for whom biopsy sampling was contraindicated were not recruited. Patients who had an overt cause for iron deficiency anemia after both upper and lower GI investigations (e.g. malignancy (n = 22), celiac disease (n = 9), poor quality digital images (n = 2) or patients without biopsies (n = 1)) were excluded. 116 patients were included in the final analysis.

**Laboratory investigations**

Fasting blood samples were collected and analyzed for full blood count, B12, folate, serum ferritin, transferrin saturation, transferrin, serum iron and iron binding capacity. Plasma was also stored at –80°C until processed for _H. pylori_ IgG (Biohit), PG I (Biohit) and PG II (Biohit) with ELISA kits according to manufacturer’s instructions.

**Endoscopy procedures**

Procedures were performed by expert endoscopists using pharyngeal local anesthetic spray Xylocaine (AstraZeneca,
Luton, UK) or conscious sedation (midazolam/pethidine) according to patient preference. A black soft rubber hood (MB46, MAJ-1990, Olympus) was attached to the endoscope tip to allow a fixed 2 mm distance between gastric mucosa and gastroscope. All procedures were done with high definition and magnification gastrosopes (GIF-FQ260Z; Olympus Optical, Tokyo, Japan) and Lucera Elite CV290 video processor. The video images were viewed on a high definition video monitor (OEV-191H, Olympus). During the procedure the mucosa was washed with a mixture containing 100 ml of water mixed with 2 ml of acetylcysteine (200 mg/ml, Parvolex, Celltech, UK) and 0.5 ml (40 mg/ml) dimethicone (Infacol, Forrest Laboratories, UK). Detailed examination of the gastric mucosa was then carried out, in WLE and then NBI using both low magnification and magnified views. Still digital images were recorded in both WLE and NBI, with biopsies taken from the areas where the digital images were recorded. A minimum of 8 images were recorded for each patient. All still images were transferred to an external hard drive.

Post endoscopy image production

Images were stored as JPEG files (200–300 kilobytes, 1093 x 948 pixels, 32-bit color), edited, anonymized and given random numbers generated in Excel Office 2010 (Microsoft Corporation, Redmond, Washington, USA) before transfer into an evaluation set of folders according to area and classification system. Two principal endoscopists (Ragunath/Sami), experts in advanced endoscopy, agreed on the magnified appearance of each selected image of the gastric mucosa in WLE-Z and NBI-Z according to specific criteria. Both were blinded to clinical, histological and serological findings. Images were reviewed and graded according to mucosal morphology using the classifications described below. These scores were used to assess magnification endoscopy performance in terms of sensitivity, specificity, accuracy, positive and negative likelihood ratios.

Image classification

Two validated classification systems were used for the gastric corpus: the Nottingham classification (Figure 1) [18] and a modified Pimentel-Nunes et al classification [24] to include gastric atrophy (Db) termed the modified Porto classification (Table 1, Figure 2).

Inter-observer variability study

Eight blinded endoscopists assessed the digital still images: four experienced, fully accredited endoscopists (experts) and

![Figure 1](image)

Figure 1. The Nottingham classification in gastric corpus. Type I (a) normal gastric body microvasculature = a honeycomb type subepithelial capillary network (SECN) and collecting veins in a regular arrangement. Type II (b) = honeycomb – type SECN with regular round pits, either with or without sulci, but with loss of collecting venules. Type III (c) = loss of the normal SECN and collecting venules, with enlarged white pits surrounded by erythema. Type IV (d) = loss of the normal SECN and round pits, with irregular arrangement of the collecting venules [18].
four trainee endoscopists (non-experts). Prior to grading the images all endoscopists underwent a comprehensive, self-directed training package which covered all the gastric mucosa classifications. Each endoscopist viewed folders containing WLE-Z and NBI-Z images of the corpus for both classifications. Approximately 464 anonymized randomized images were reviewed over an unrestricted length of time. Data on image classification from a drop down menu of pre-defined options and image quality according to a 10 point Visual Analogue Scale (VAS) [37] were recorded onto an Excel spreadsheet.

**Histopathology analysis**

Two gastric antrum, two gastric corpus and four duodenal samples were fixed in formalin, then embedded in paraffin and cut into approximately 4 μm sections. Sections were stained with haematoxalin and eosin (H&E) to allow histological scores to be carried out according to the updated Sydney scoring system [38]. All specimens were classified as none (0), mild (1), moderate (2) or severe (3) for the following histological features activity (polymorphonuclear cell infiltration), inflammation (mononuclear cell infiltration), atrophy (loss of specialized glands) and intestinal metaplasia (replacement gastric mucosa by metaplastic columnar absorptive cells and goblet cells with intestinal morphologic features). Toluidine blue staining was also carried out for *H. pylori* density grading. A single blinded expert GI histopathologist carried out the histological grading. One biopsy from the antrum was placed in urease medium for rapid urease detection, and another sample in iso-sensitest broth (Oxoid, Cambridge, UK)/10% glycerol for *H. pylori* isolation and culture.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>H. pylori+</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mucosal pattern</strong></td>
<td>Regular circular/oval</td>
<td>Regular Ridge/tubulo-villous with or without light blue crest</td>
<td>Regular</td>
<td>Irregular/ absent with or without white opaque substance</td>
<td>Absent</td>
</tr>
<tr>
<td><strong>Vascular pattern (SECN)</strong></td>
<td>Regular mucosa and SECN</td>
<td>Vascular pattern central in antrum/peripheral in corpus of gland</td>
<td>Regular/ variable vascular density</td>
<td>Irregular</td>
<td>Absent</td>
</tr>
<tr>
<td><strong>Expected histology</strong></td>
<td>Normal</td>
<td>Intestinal Metaplasia</td>
<td>H. pylori infection</td>
<td>Dysplasia</td>
<td>Atrophic gastritis</td>
</tr>
</tbody>
</table>

Table 1. Modified Porto classification.

The letter b for corpus is added to the class type to signify the site. For *H. pylori* positive pattern, the symbol (+) is added to the end of the pattern class site [24].

**Figure 2.** Modified Porto gastric corpus classification. Normal corpus (Ab) in image (a), *H. pylori* gastritis corpus (Ab+) (b), intestinal metaplasia in corpus with some light blue crest (Bb) in image (c) and atrophy in corpus (Db) in image (d). Magnification 115X.
Table 2. Patient description and laboratory data.

<table>
<thead>
<tr>
<th>Numbers</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male: Female ratio</td>
<td>1:1.1</td>
</tr>
<tr>
<td>Age, median (range) years</td>
<td>67 (19–85)</td>
</tr>
<tr>
<td>% Pharyngeal local anesthetic: conscious sedation</td>
<td>38%:62%</td>
</tr>
<tr>
<td>Midazolam median dose (range) mg</td>
<td>3 mg (1–5)</td>
</tr>
<tr>
<td>Pethidine median dose (range) mg</td>
<td>25 mg (25–100)</td>
</tr>
<tr>
<td>Hemoglobin males (130–180 g/L) median (range) g/L</td>
<td>110 g/L (85–129)</td>
</tr>
<tr>
<td>Hemoglobin females (120–165 g/L) median (range) g/L</td>
<td>101 g/L (42–118)</td>
</tr>
<tr>
<td>MCV (84–102 fl) median (range)</td>
<td>83 fl (56–111)</td>
</tr>
<tr>
<td>Ferritin (25–350 µg/L) median (range) µg/L</td>
<td>12 µg/L (1–474)</td>
</tr>
<tr>
<td>Iron (9–31 µmol/L) median (range) µmol/L</td>
<td>11 µmol/L (1–83)</td>
</tr>
<tr>
<td>% H. pylori infection</td>
<td>24%</td>
</tr>
<tr>
<td>% Atrophy corpus</td>
<td>35%</td>
</tr>
<tr>
<td>% Intestinal metaplasia</td>
<td>14.7%</td>
</tr>
</tbody>
</table>

Statistical methods

Sensitivity, specificity, positive and negative likelihood ratios along with 95% confidence intervals (CI) for magnification endoscopy appearances and serological markers were compared with histology. A receiver operating characteristic (ROC) curve was used to assess the diagnostic accuracy. Chi-squared tests were used to compare the diagnostic accuracies. Kappa (k) statistics were calculated to determine inter-observer agreement among experts and non-experts. Interpretation of k values was as follows: <0 = no agreement; 0–0.20 = slight agreement; 0.21–0.40 = fair agreement; 0.41–0.60 = moderate agreement; 0.61–0.80 = substantial agreement; and 0.81–1 = almost perfect agreement [39]. Differences in image quality were assessed using the Mann-Whitney U test. p values < .05 were considered statistically significant. Stata version 14 (Stata Corporation, College Station, Texas) was used for the statistical analysis.

Results

Patient characteristics are described in Table 2. 24% of patients had histological evidence of H. pylori infection, 35% had gastric atrophy and 14.7% had intestinal metaplasia.

Diagnostic accuracy of NBI-Z and serology

NBI-Z diagnostic accuracy for determining normal corpus (Type I) was 0.87 (95% CI 0.82–0.92), H. pylori gastritis (Type II/III) 0.65 (95% CI 0.55–0.75) and gastric atrophy (Type IV) 0.88 (95% CI 0.81–0.94) respectively.

When NBI-Z was compared with serology, NBI-Z was superior for detecting gastric atrophy: NBI-Z Type IV 0.88 (95% CI 0.81–0.94) vs PG I/II ratio < 3 0.74 (95% CI 0.62–0.85), p < .0001 and NBI-Z 0.88 (95% CI 0.81–0.94) vs PG < 30 µg/L/L 0.75 (95% CI 0.64–0.87), p < .0001. Although H. pylori IgG had a numerically higher accuracy this did not achieve statistical significance, NBI-Z Type II/III 0.65 (95% CI 0.55–0.75) vs H. pylori IgG 0.82 (0.73–0.9), p = .078.

Diagnostic accuracy of the modified Porto classification

NBI was more accurate than WLE for detecting normal mucosa, H. pylori gastritis and atrophy in the corpus. Normal corpus (Ab): NBI-Z 0.81 (95% CI 0.74–0.87) vs WLE-Z 0.79 (95% CI 0.72–0.87), p < .0001. Atrophy (Type Ab+): NBI-Z 0.77 (95% CI 0.67–0.87) vs WLE-Z 0.62 (95% CI 0.51–0.74), p = .01. Gastric atrophy (Db): NBI-Z 0.71 (95% CI 0.62–0.79) vs WLE-Z 0.65 (95% CI 0.56–0.73), p < .0001. For detecting intestinal metaplasia, there was no statistical difference in accuracy. Intestinal metaplasia (Bb): NBI-Z 0.66 (95% CI 0.53–0.79) vs WLE-Z 0.59 (95% CI 0.45–0.73), p = .28 (Table 4).

Inter-observer agreement

The mean kappa values for inter-observer agreement for NBI endoscopy images were higher among expert than non-expert endoscopists. For describing the corpus using the Nottingham classification the agreement was 0.63(95% CI 0.51–0.73) vs 0.5 (95% CI 0.39–0.62). When describing atrophy the agreement was 0.65 (95% CI 0.53–0.75) vs 0.47 (95% CI 0.2–0.68). Using the modified Porto classification when describing the corpus was 0.33 (95% CI 0.21–0.43) vs 0.2 (95% CI 0.12–0.28). When describing intestinal metaplasia the agreement was 0.36 (95% CI 0.24–0.47) vs 0.21 (95% CI 0.1–0.34). Both expert and non-expert endoscopists rated the overall image quality VAS higher for NBI than WLE (7 vs 6 p < .0001).

Discussion

Standard WLE is limited in its ability to accurately diagnose gastric lesions, which has led to the development of a variety of techniques such as chromoendoscopy, NBI, flexible spectral imaging color enhancement (FICE), autofluorescence (AFI) and confocal laser endomicroscopy [19] in addition to gold standard histology. NBI has been extensively studied for its use in detecting dysplasia in the esophagus, bronchus and in colonic polyps [19,40–42].

We conducted this pilot study to assess the role of NBI and serology in predicting the histological diagnosis of gastric mucosa. We demonstrated good correlation with normal and gastric atrophy using magnification endoscopy. However, for detecting H. pylori gastritis our results were limited with an accuracy of 0.65 using the Nottingham classification. Endoscopic diagnosis of H. pylori infection requires a subjective assessment of vascular density of subepithelial capillary
shows the majority have good correlation with histology. The use of magnified endoscopy in detecting other disease are a true reflection of the structural differences seen in vitro historically. This in turn allows NBI to predict histological diagnosis [43]. The endoscopic appearance of gastric atrophy has been established since the 1970’s [44]. The normal gastric body consists of a regular pattern of honeycomb type SECN and collecting venules but in H. pylori gastritis this is irregular and accompanied by surrounding edema [18,45].

Very few studies have investigated the endoscopic appearance of H. pylori gastritis, with nodular mucosal and gastric fold hypertrophy the most consistent features despite low sensitivities when compared to histology [46,47]. Thus the exact features that describe H. pylori gastritis are not entirely known which makes description and interpretation difficult. The findings using the Nottingham classification were lower than when initially described using WLE [18].

Potential reasons for this include the fact that the principal endoscopists scoring the images did not perform the procedures and, therefore, were blinded to potential clinical information that could influence decision making. Although the modified Porto classification required more time to examine the images it also offered more details in terms of the presence of intestinal metaplasia and dysplasia. The specificity was high for normal mucosa, gastric atrophy and intestinal metaplasia. Potentially this could enable endoscopists to confidently avoid taking biopsies in the corpus. The evidence for the use of magnified endoscopy in detecting other disease shows the majority have good correlation with histology [18,28,48–50]. Currently, the gold standard for the diagnosis of intestinal metaplasia and dysplasia remains histology despite promising results with NBI.

When compared with serological markers, NBI-Z overall performed better. Previous studies have also shown that NBI is accurate in detecting premalignant lesions [29,30,51]. The serological data presented in this study was similar to previous studies in terms of sensitivity and specificity [13,31,52,53]. H. pylori serology is commonly used in clinical practice in the UK but markers of atrophy are not. These results suggest the PG I/II ratio or PG I alone cannot replace endoscopy surveillance or detection of gastric atrophy. However, using this in clinical practice could reduce the need to obtain histology if serology is negative and NBI-Z does not suggest disease, therefore, reducing associated costs and time.

Inter-observer agreement was unsurprisingly higher among expert endoscopists when compared with non-experts. Endoscopy assessment is dependent on experience and training, so these techniques are likely to perform better in the hands of experts [54,55]. The training of Western endoscopists is also likely to be different compared to Asian countries where the incidence of gastric cancer is higher. The observers VAS median scores were higher for NBI-Z than WLE-Z suggesting NBI provided more clarity to enable a diagnosis.

Our results are more applicable to specialist centers that routinely use NBI endoscopy as it requires a certain skill level and so may not always be practical to use in routine clinical practice. In the same way, the captured images in this study were captured by endoscopists with advanced diagnostic imaging experience and this is unlikely to happen routinely. As shown by the lower agreement amongst non-experts, NBI requires training in mucosal pattern recognition. The main comparison was done with experts so if this was used in a real world setting the results may be poorer. Also the time duration each endoscopist spent reviewing the still images in a real world setting the results may be poorer. Also the time constraints seen in clinical practice, possibly influencing the decision making process. In terms of training, this study

### Table 3. Sensitivity, specificity and likelihood ratios with 95% CI of NBI-Z and WLE-Z using the modified Porto classification and serology.

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Positive likelihood ratio (95% CI)</th>
<th>Negative likelihood ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBI-Z Type I</td>
<td>75.6% (64.6–84.7)</td>
<td>98.6% (92.2–100)</td>
<td>54.0 (7.4–366.8)</td>
<td>0.25 (0.17–0.37)</td>
</tr>
<tr>
<td>WLE-Z Type I</td>
<td>66.2% (54.6–76.6)</td>
<td>96.8% (89–99.6)</td>
<td>20.7 (5.3–82.4)</td>
<td>0.35 (0.25–0.48)</td>
</tr>
<tr>
<td>NBI-Z Type II/III</td>
<td>62.1% (42.3–79.3)</td>
<td>68.1% (58.8–76.4)</td>
<td>1.95 (1.32–2.87)</td>
<td>0.56 (0.34–0.9)</td>
</tr>
<tr>
<td>WLE-Z Type II/III</td>
<td>74.1% (53.7–88.9)</td>
<td>64.9% (55–73.7)</td>
<td>2.11 (1.5–2.95)</td>
<td>0.40 (0.21–0.77)</td>
</tr>
<tr>
<td>NBI-Z Type IV</td>
<td>75.6% (59.7–87.6)</td>
<td>100% (96.6–100)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>WLE-Z Type IV</td>
<td>64.1% (47.2–78.8)</td>
<td>98% (93–99.8)</td>
<td>32.1 (8–130)</td>
<td>0.37 (0.24–0.56)</td>
</tr>
<tr>
<td>H. pylori IgG</td>
<td>90% (68.3–98.8)</td>
<td>73.9% (61.9–83.7)</td>
<td>3.45 (2.26–5.27)</td>
<td>0.14 (0.04–0.51)</td>
</tr>
<tr>
<td>PG I/II ratio &lt;3</td>
<td>73.7% (48.8–90.9)</td>
<td>73.8% (60.9–84.2)</td>
<td>2.81 (1.71–4.63)</td>
<td>0.36 (0.17–0.77)</td>
</tr>
<tr>
<td>PG I &lt;30 µg/l</td>
<td>52.6% (28.9–75.6)</td>
<td>98.4% (91.2–100)</td>
<td>32.11 (4.4–235)</td>
<td>0.48 (0.3–0.77)</td>
</tr>
</tbody>
</table>

### Table 4. Sensitivity, specificity and likelihood ratios with 95% confidence intervals of NBI-Z and WLE-Z using the modified Porto classification in the corpus.

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Positive likelihood ratio (95% CI)</th>
<th>Negative likelihood ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBI-Z Ab</td>
<td>70.5% (59.8–79.7)</td>
<td>90.9% (78.3–97.5)</td>
<td>7.75 (3–19.9)</td>
<td>0.32 (0.23–0.45)</td>
</tr>
<tr>
<td>WLE-Z Ab</td>
<td>77.8% (66.4–86.7)</td>
<td>81% (65.9–91.4)</td>
<td>4.09 (2.2–7.7)</td>
<td>0.27 (0.17–0.43)</td>
</tr>
<tr>
<td>NBI-Z Ab+</td>
<td>75% (53–90.2)</td>
<td>78.7% (69.8–86)</td>
<td>3.52 (2.3–5.4)</td>
<td>0.32 (0.16–0.64)</td>
</tr>
<tr>
<td>WLE-Z Ab+</td>
<td>43.5% (23.2–65.5)</td>
<td>81.3% (71.8–88.7)</td>
<td>2.33 (1.24–4.4)</td>
<td>0.69 (0.48–1.01)</td>
</tr>
<tr>
<td>NBI-Z Bb</td>
<td>35.7% (12.8–64.9)</td>
<td>95.8% (90.4–98.6)</td>
<td>8.43 (2.78–25.5)</td>
<td>0.67 (0.45–0.99)</td>
</tr>
<tr>
<td>WLE-Z Bb</td>
<td>27.3% (6–61)</td>
<td>90.1% (82.5–95.1)</td>
<td>2.75 (0.89–8.53)</td>
<td>0.81 (0.56–1.17)</td>
</tr>
<tr>
<td>NBI-Z Db</td>
<td>70% (52–81)</td>
<td>99% (94.5–100)</td>
<td>70.0 (10.74–307)</td>
<td>0.30 (0.23–0.78)</td>
</tr>
<tr>
<td>WLE-Z Db</td>
<td>61% (45–76)</td>
<td>100% (95.7–100)</td>
<td>–</td>
<td>0.39 (0.27–0.50)</td>
</tr>
</tbody>
</table>
did not measure inter-observer agreement before and after the training session which would have given some insight into the learning required for NBI use in a routine clinical setting.

Study strengths include the use of two endoscopy classification criteria to describe the magnified gastric mucosa in addition to serological markers in a large cohort of prospectively recruited patients. Post endoscopy image assessment controlled for clinician influence on pretest probability. Bias was also reduced by blinding both the endoscopists and histopathologist to clinical data. Also the consistent use of the same gastroscope reduced image quality variability. Patients on proton pump inhibitors were also excluded to lower the number of false negatives in terms of H. pylori infection and to avoid missing early gastric neoplasia [56]. This work has provided further evidence to support routinely investigating the presence of gastric atrophy and H. pylori gastritis in iron deficiency anemia patients [31,36].

With regards to study limitations, by including only iron deficiency anemia patients this makes the study prone to selection bias and thus may not be representative at a population scale. However, as the annual incidence of gastric atrophy is low (0–10.9%) [57] we needed a larger cohort to enable estimates of sensitivity and specificity to be made to guide sample size calculations in future larger studies. Extent of disease also influenced the results. For example, atrophy serology tests only detect moderate to severe atrophy so endoscopy is more likely to perform better as these classifications only detect the presence of atrophy and not the degree. Inter-observer agreement for the modified Porto classification may also be lower due to the examiners being more familiar with the Nottingham classification at this specialist center. Finally, although the addition of atrophy to the modified Porto classification provided a more detailed description which more accurately resembled histology it also made the classification more complex and time consuming. The diagnosis of intestinal metaplasia is more reliable in terms of both histological grading and disease progression.

In conclusion, NBI-Z can detect changes in the GI mucosa which are usually dependent on histology. Although serology performs well, NBI endoscopy performs better in terms of disease detection with a high specificity and moderate to substantial observer agreement. A detailed examination with NBI-Z could potentially help identify early precancerous gastric mucosa, which could enable patients to be promptly enrolled in appropriate endoscopic surveillance with improved disease outcomes. Also NBI-Z use may allow stratification of the need for histology and thus minimize associated costs, time and sampling error. These study findings will help to design future trials to evaluate NBI techniques in the gastric mucosa.

Acknowledgements

We would like to thank the faculty staff especially Samantha Warburton, Susan Henry and Melanie Lingaya and the equipment support provided by the NIHR Nottingham Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

Disclosure statement

The authors report no conflict of interest.

Funding

This work was supported by the National Institute for Health Research Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and The University of Nottingham, United Kingdom.

ORCID

Jonathan R. White http://orcid.org/0000-0002-6264-4560

References


