Diagnostic accuracy of narrow-band imaging endoscopy with targeted biopsies compared with standard endoscopy with random biopsies in patients with Barrett’s esophagus: A systematic review and meta-analysis

Omer Hajelsedig, Leonardo Zorron Cheng Tao Pu, Jacqueline Y Thompson, Anton Lord, Iman El Sayed, Chase Meyer, Faisal Shakaut Ali, Hebatullah M Abdulazeem, Ammar O Kheir, Timo Siepmann, Rajvinder Singh

Key words
Barrett’s esophagus, esophageal adenocarcinoma, narrow-band imaging endoscopy, random biopsy, targeted biopsy.

Abstract
Background and Aim: Endoscopic surveillance for dysplasia in Barrett’s esophagus (BE) with random biopsies is the primary diagnostic tool for monitoring clinical progression into esophageal adenocarcinoma. As an alternative, narrow-band imaging (NBI) endoscopy offers targeted biopsies that can improve dysplasia detection. This study aimed to evaluate NBI-guided targeted biopsies’ diagnostic accuracy for detecting dysplasia in patients undergoing endoscopic BE surveillance compared with the widely used Seattle protocol. Methods: Cochrane DTA Register, MEDLINE/PubMed, EMBASE, OpenGrey, and bibliographies of identified papers were searched until 2018. Two independent investigators resolved discrepancies by consensus, study selection, data extraction, and quality assessment. Data on sensitivity, specificity, and predictive values were pooled and analyzed using a random-effects model. Results: Of 9528 identified articles, six studies comprising 493 participants were eligible for quantitative synthesis. NBI-targeted biopsy showed high diagnostic accuracy in detection of dysplasia in BE with a sensitivity of 76% (95% confidence interval [CI]: 0.61–0.91), specificity of 99% (95% CI: 0.99–1.00), positive predictive value of 97% (95% CI: 0.96–0.99), and negative predictive value of 84% (95% CI: 0.69–0.99) for detection of all grades of dysplasia. The receiver-operating characteristic curve for NBI model performance was 0.8550 for detecting all dysplasia. Conclusion: Narrow-band imaging-guided biopsy demonstrated high diagnostic accuracy and might constitute a valid substitute for random biopsies during endoscopic surveillance for dysplasia in BE.

Background
Esophageal adenocarcinoma (EAC) is the eighth most common cancer and represents the sixth most common cause of cancer-related deaths worldwide.1 It is one of the fastest rising cancers in the world.2 Barrett’s esophagus (BE) is the only identifiable pre-malignant condition linked to worsening clinical outcome.3 BE is an esophagus in which any portion of the normal squamous epithelial mucosa is replaced by metaplastic columnar epithelium, visible under endoscopy (at least 1 cm) above the gastro-esophageal junction and confirmed by histopathological esophageal biopsies.4 BE-associated cancers arise via a sequence of metaplasia, dysplasia to become a carcinoma. Neoplasia in BE
can be prevented, and patient outcomes can be improved by early detection and eradication of dysplasia. The overall 5-year survival of patients with EAC is close to 5% in most countries. This high mortality rate is partially a result of the late detection with a high proportion of patients presenting with advanced disease.3

Barrett’s esophagus prevalence has been estimated at 1–2% in the general population and 10–15% in those with gastroesophageal reflux disease.5 Although there is an increased risk of developing esophageal cancer in patients with BE, the absolute risk of EAC in BE remains low. The annual risk of BE transforming into EAC is 0.12% to 0.5%.6 However, the risk of developing cancer is higher among patients with dysplasia at baseline diagnosis than patients with non-dysplastic BE; 5.6% of patients have a concurrent diagnosis of low-grade dysplasia (LGD) at the time of BE diagnosis.6 In a large cohort of Danish patients with BE observed between 1992 and 2009, the relative risk of EAC among patients with LGD at baseline compared with those who did not have LGD was 4.8 (95% confidence interval [CI]: 2.6 to 8.8). This increased to 21.1 (95% CI: 17.8 to 24.7) among patients with high-grade dysplasia (HGD) at baseline. Symptomatic gastroesophageal reflux disease for 10 years or more has been identified as a significant factor for developing both BE and, consequently, EAC.7 Other risk factors of EAC include being older than 60 years, obese, male, and presenting with long-segment Barrett (> 3 cm).8

The diagnosis of BE is established via endoscopy (see Appendix A for details). BE’s size and extent can be objectively graded using the uniform Prague C & M criteria.9 Of all types of metaplastic columnar epithelium in BE, intestinal metaplasia (IM) within the goblet cells increases the risk of evolution to neoplasia.10 A population-based cohort study11 demonstrated a substantially lower EAC risk in subjects with columnar metaplasia without IM than those with IM (hazard ratio of 0.38% per year vs 0.07% per year).

All guidelines from major Gastroenterology societies (American College of Gastroenterology, British Society of Gastroenterology, American Society of Gastrointestinal Endoscopy [ASGE], and American Gastroenterological Association) and BOBCAT (Benign Barrett’s and Cancer Taskforce) recommend performing endoscopic surveillance in BE to detect and subsequently treat dysplasia and early neoplasia using random biopsy with the “Seattle” protocol.12 These guidelines, however, differ in surveillance intervals according to the length of BE. For non-dysplastic BE, the intervals range between 3 and 5 years. When dysplasia is detected, the interval shortens to 6–12 months, depending on the grade of dysplasia. The Seattle protocol for BE involves obtaining random, blind, four quadrant biopsies using white-light standard endoscopy (WLSE) within every 1- to 2-cm interval of a BE segment. This is in addition to taking targeted biopsies on macroscopically visible lesions. This approach is, however, fraught with problems. Sampling error can frequently occur as only an area of 3.5% in a given 2-cm BE is sampled, missing out the other 96.5%. As a result, LGD, HGD, or even early malignant lesions can easily be missed. Evidence shows that only about half of all endoscopists (41–56%) adhere to this biopsy protocol systematically.13

Narrow-band imaging

The narrow-band imaging (NBI), developed by Olympus, Tokyo, Japan, is an electronic chromoendoscopy technology, which can provide detailed contrast enhancement of the mucosal surface architecture and vasculature using a narrower bandwidth of light through specialized filters. The narrow bandwidth of light (blue and green) enables more penetration of the underlying mucosa. Given that hemoglobin has an affinity to absorb blue light, NBI accentuates the mucosa’s vascular architecture. Dysplastic tissue and cancerous cells (with increased angiogenesis) can hence be detected and characterized using NBI. Targeted biopsies instead of random biopsies could be obtained to confirm a suspected IM, dysplasia, or early EAC diagnosis. NBI endoscopy has been endorsed by ASGE recently as a diagnostic tool while performing surveillance in BE. However, NBI endoscopy has not been adopted into clinical practice or incorporated in current clinical guidelines developed by American Gastroenterological Association, American College of Gastroenterology, British Society of Gastroenterology, or BOBCAT.

White-light standard endoscopy, as opposed to NBI endoscopy, has been widely used in BE surveillance. WLSE identifies the columnar epithelium and visible macroscopic anomalies of the mucosa but cannot distinguish IM, dysplasia, or early EAC. Random quadratic biopsies obtained via WLSE represented in Seattle protocol, therefore, are required. Recently, the high-definition property has been added to white-light endoscopy, but early neoplasia remains indistinguishable. Chromoendoscopy, which involves topical application of various dyes or contrasts on esophageal mucosal to enhance its endoscopic morphology, has been used in conjunction with WLSE to improve the detection of pre-malignant lesions. However, a recent meta-analysis comparing rates for detecting neoplasia in BE with methylene blue compared with random four quadrant biopsies found no significant yield for detecting HGD or early cancer.14 Furthermore, the use of various dyes can sometimes be cumbersome, toxic, messy, and time-consuming. The technique of chromoendoscopy has, therefore, not gained widespread acceptance and has mostly been abandoned.

As a result, new approaches and techniques are increasingly advocated and utilized to detect dysplasia and early cancer in BE surveillance in real-time via targeted biopsy instead of random biopsy method. NBI endoscopy is widely available and easy to use. A push button is readily available on the device allowing the endoscopist to switch between WLSE and NBI mode. Therefore, NBI endoscopy offers a combined approach to improving lesion detection rate and allowing targeted biopsies. This is, in turn, may lead to lesions being detected, characterized, and treated earlier. This approach can potentially improve clinical outcomes in patients with BE and significantly reduce the cost and time associated with performing and interpreting random biopsies. This may also result in a paradigm shift in the surveillance of patients with BE. So far, NBI endoscopy as a targeted biopsy has been extensively evaluated, but no study has systematically assessed the available literature to provide findings that support extrapolating its use for routine community settings of practice. NBI use to date remains limited to tertiary centers.

This systematic review’s primary objective was to assess the diagnostic accuracy of NBI-targeted biopsy for dysplasia and early EAC in BE compared with random biopsy with standard WLSE protocol. The aim is to make recommendations for using NBI-guided biopsy instead of WLSE-guided random biopsy for BE surveillance.
Methods

This systematic review was developed using Preferred Reporting Items of Systematic Reviews and Meta-Analyses recommendations (Fig. 1).15 The protocol was registered on PROSPERO (ID: CRD4201707328).

**Study search.** Diagnostic search filters recommended by the National Institute of Health were used to develop the search strategy and improve the sensitivity of articles identified.16 The following search terms were combined: Barrett’s esophagus or Barrett’s esophagus and narrow-band imaging or NBI endoscopy. Language, study settings, or time restrictions were not used to develop the search strategy. All studies identified by the search process, both published and unpublished records relevant to the review question, were retrieved. However, only studies conducted exclusively in humans and published in English were retained for consideration.

The electronic search was conducted from onset until November 2018 in Cochrane DTA Register, MEDLINE/PubMed, EMBASE, and OpenGrey (Appendix B). We used SCOPUS to identify other potentially relevant papers cited in a paper that was eligible for inclusion. In addition, the reference lists of included key studies were manually screened for potentially eligible studies. Once the search was complete, we consulted experts in the field to identify any studies that may have been missed or are awaiting publication.

**Selection of studies.** In this systematic review and meta-analysis, we included all studies that evaluated targeted biopsies taken in NBI mode with random biopsies using Seattle WLSE
protocol among patients presenting for routine surveillance or were referred for further evaluation of dysplasia in BE. The reference standard test to confirm dysplasia in both methods of biopsies was histology.

Three authors (O. H., L. Z. C. T. P., and I. E.) independently performed the study selection. Articles were screened at title and abstract. Following this, two authors (O. H. and L. Z. C. T. P.) independently reviewed full texts of potentially eligible studies using the inclusion and exclusion criteria described in the succeeding text. Disagreements were resolved through consensus between authors. A fourth author (A. K.) arbitrated when consensus could not be achieved.

Inclusion criteria are as follows:

1. prospective randomized clinical trials and controlled observational cohort studies;
2. studies that evaluated dysplasia or EAC in BE as primary or secondary outcomes; and
3. studies that compared NBI against WLE random biopsy or dye-based/virtual chromoendoscopy.

Exclusion criteria are as follows:

1. any study that evaluated BE surveillance without NBI endoscopic technology;
2. studies that evaluated NBI technology with other advanced endoscopic imaging systems that did not separate the results for NBI;
3. studies that had no comparator, neither random biopsies nor chromoendoscopy;
4. studies reported as reviews, case–control, case reports, case series, and editorials; and
5. studies in non-human, animal, or laboratory samples.

**Data collection and synthesis.** Three independent researchers (O. H., L. Z. C. T. P., and I. E.) reviewed selected articles using the study eligibility criteria. The data extraction form was designed using guidelines from the Cochrane handbook for systematic reviews. It was carried out by two independent authors (C. M. and H. A.). Discrepancies were resolved through consensus. A third blinded reviewer (A. L.) was consulted to resolve the disagreement when consensus could not be achieved. The characteristics of the included studies and quality assessment tables were collated and presented in a tabular format. When information regarding any of the data collected in the succeeding text was unclear, we contacted the corresponding authors to request further details.

All endoscopic results obtained via NBI-guided targeted biopsies were reported as categorical data (positive or negative for dysplasia). This was obtained in real time based on the NBI-enhanced mucosal appearance. Dysplasia is suspected when the mucosal villous arrangement is distorted or the vascular pattern is irregular. We anticipated a high inter-observer disagreement in the diagnostic threshold of dysplasia detection among eligible studies. To account for this discrepancy, we checked if studies used standard classification to diagnose dysplasia optically. We also used Barrett’s International NBI Group classification system as a reference.

**Assessment of methodological quality.** We assessed the methodological quality of included studies using a modified version of the Quality Assessment of Diagnostic Accuracy Studies-2 tool. Each study was assessed for quality by one author (I. E.) using four key domains to assess the risk of bias and concerns regarding application to the research question (patient selection, index test, reference standard, and flow–timing domains). The patient selection domain, we omitted the signaling question “Was a case-control design avoided?” as we already excluded case–control studies as a criterion for study selection. For the index test domain, we assessed the risk of bias for NBI versus Seattle protocol of random biopsies and image-to-pathology detection separately.

Moreover, we added the following signaling question to judge concerns regarding the applicability of the index test: “Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?”

Each paper was scored as having “low,” “high,” or “unclear” risk of bias for each of the four domains, patient selection, index test, and reference standard applicability. Studies classified as high or unclear risk of bias or high concern regarding applicability in at least one domain would be regarded as having a low methodological design. In the case of disagreements, differences were resolved through consensus by consultation. The results were presented in graphics as a summary of all studies (Figs 2 and 3).

Figure 2  Risk of bias and applicability concerns graph: review authors’ judgments about each domain presented as percentages across included studies.  

- **High**  
- **Unclear**  
- **Low**
**Statistical analysis.** Data from included studies were extracted and analyzed using RevMan and R software version 3.5.1 to assess the primary outcome—the diagnostic accuracy of NBI in dysplasia detection in BE. Pooled estimates were combined using the “metafor” and “lme4” packages. Receiver-operating characteristic (ROC) curves were produced using the “sROC” package. The total number of patients was obtained, and the total number of detected lesions, then NBI yield assessment was carried out on both per-patient and per-lesion bases. The primary outcome was the overall accuracy of NBI against Seattle protocol reported as pooled sensitivity, specificity, and positive and negative predictive values (PPV and NPV). To calculate pooled estimates from eligible studies, data from each study were summarized using 2 × 2 matrices to describe the proportion of true positives, false positives, true negatives, and false negatives (Table 2). When only summary statistics are presented instead of raw counts, these were converted to raw counts using the information provided in the paper and RevMan. The additive value of NBI and WLSE over WLSE alone was calculated using the number needed to detect (NND) method.

Significance was calculated using both a fixed-effects or random-effects meta-analysis. The Mantel–Haenszel (fixed-effects) or Der Simonian and Laird (random-effects) method was used when heterogeneity was identified. Results were displayed using ROC curves. For studies with a common threshold, the model considered within-study variation and between-study variance.
variation and focused on estimating a summary operating point (i.e. a summary value for sensitivity and specificity). In addition, we estimated the 95% confidence region and the 95% prediction region around the summary operating point. We performed these analyses using the command R3.4.4. We summarized finding using a GRADE (Grades of Recommendation, Assessment, Development, and Evaluation) summary of findings table approach to diagnostic test accuracy.20

**Investigations of heterogeneity.** We explored heterogeneity by performing a visual assessment of study results in forest plots and ROC space in the prediction region. We formally assessed the source of heterogeneity by examining differences in diagnostic accuracy between subgroups of studies. We then used the bivariate method to analyze how the summary estimate of sensitivity and specificity varied according to study-level covariates. For this purpose, we created a factor variable with $N$ categories.
and generated an N of 1 dummy entered into the bivariate model to test the effect of covariates on both sensitivity and specific.\textsuperscript{21} We defined sources of heterogeneity a priori as the following factors:

1. overall prevalence of BE in a study population;
2. median length of BE in each study;
3. percentage of male in the study population;
4. blinding of the pathologist on the pooled primary outcomes; and
5. variation in the diagnostic threshold of NBI across all included studies.

**Results**

**Results of the search.** Our search strategy identified 9528 studies from 1947 to November 2018. After duplicates were removed, 7826 studies were screened at title and abstract, resulting in 137 studies eligible for inclusion. After full-text screening, 14 studies of the 137 potentially eligible studies were included for qualitative synthesis, while only six fulfilled the criteria for quantitative synthesis (Fig. 1). The ROC curves (Figs 4–6) summarize the NBI model performance for detecting dysplasia. The area under the curve was 0.8550 for all grades of dysplasia (AGD) and 0.8536 for HGD.

**Diagnostic accuracy of narrow-band imaging in detecting low-grade dysplasia.** We found only three eligible studies\textsuperscript{22–24} that analyzed the utility of NBI for detecting LGD.\textsuperscript{22,23,25} The sensitivity of NBI for detecting LGD ranged from 0.10 (95% CI: 0.00–0.36) to 0.87 (95% CI: 0.69–1.00), whereas the specificity ranged from 0.93 (95% CI: 0.87–0.98) to 1.00 (95% CI: 0.99–1.00). The pooled sensitivity and specificity of NBI for detecting LGD were 0.60 (95% CI: 0.11–1.00) and 0.98 (95% CI: 0.95–1.00), respectively. The PPV of NBI for detecting LGD ranged from 0.45 (95% CI: 0.16–0.75) to 0.96 (95% CI: 0.87–1.00), with a pooled PPV of 0.71 (95% CI: 0.30–1.00). Similarly, the NPV of NBI for detecting LGD ranged from 0.96 (95% CI: 0.91–1.00) to 0.99 (95% CI: 0.96–1.00), with a pooled NPV of 0.98 (95% CI: 0.96–1.00). The number needed to detect additional patients with LGD using NBI was 3.95. No study reported data for the median prevalence of LGD.

**Diagnostic accuracy of narrow-band imaging in detecting high-grade dysplasia.** Five studies reported the utility of NBI for detecting HGD.\textsuperscript{23,25–27} The sensitivity of NBI for detecting HGD ranged from 0.54 (95% CI: 0.27–0.81) to 0.89 (95% CI: 0.67–1.00), whereas the specificity ranged from 0.81 (95% CI: 0.67–0.95) to 1.00 (95% CI: 0.99–1.00). The pooled sensitivity and specificity of NBI for detecting HGD were 0.83 (95% CI: 0.73–0.93) and 0.99 (95% CI: 0.99–1.00), respectively. The PPV of NBI for the detecting HGD ranged from 0.73

---

**Figure 6** The receiver-operating characteristic (ROC) curves summarize the performance of narrow-band imaging models for detection of high-grade dysplasia (HGD). --- SROC; ----, conf. region; △ data; ○ summary estimate.
| Study name | Wolfsen | Sharma | Jayasekera | Kara | Lee | Pascarenco |
|------------|---------|--------|------------|------|-----|------------|
| Study title | Prospective, controlled tandem endoscopy study of narrow band imaging for dysplasia detection in Barrett’s esophagus | Standard endoscopy with random biopsies versus narrow band imaging targeted biopsies in Barrett’s esophagus: a prospective, international, randomized controlled trial | Added-value of narrow-band imaging and confocal laser endomicroscopy in detecting Barrett’s esophagus neoplasia | High-resolution endoscopy plus chromoendoscopy or narrow-band imaging in Barrett’s esophagus: a prospective randomized crossover study | Benefits of the Seattle biopsy protocol in the diagnosis of Barrett’s esophagus in a Chinese population | A preliminary feasibility study: narrow-band imaging targeted versus standard white light endoscopy non-targeted biopsies in a surveillance Barrett’s population |
| Country | USA | USA, Netherlands | Australia | Netherlands | Taiwan | Romania |
| Study design | A prospective, blinded, tandem endoscopy study | Randomized, crossover trial | Cross-sectional | Randomized crossover | Single-arm, crossover | Randomized, crossover trial |
| Study duration | NR | October 2005 to April 2009 | February 2010 to September 2011 | NR | October 2012 to December 2014 | January 2013 to December 2014 |
| Study setting | Multicenter | Multicenter | Single center | Single center | Single center | Single center |
| Total patients N | 65 | 123 | 50 | 28 | 143 | 84 |
| Mean age | 65 | 61 | 93 | 79 | NR | NR |
| Male, % | 82 | 1.8 | 1.8 | 1.8 | 1.8 | 1.8 |
| BE average length, cm | 4 | 4.7 mean | 4.7 mean | 4.7 mean | 4.7 mean | 4.7 mean |
| NBI biopsies, N | 1.2 mean | 1.2 mean | 1.2 mean | 1.2 mean | 1.2 mean | 1.2 mean |
| WLESE biopsies, N | 4.7 mean | 4.7 mean | 4.7 mean | 4.7 mean | 4.7 mean | 4.7 mean |
| Study outcomes | HGD, IM, HGD, LGD, EAC | HGD, LGD, EAC | HGD, IM | HGD, IM | LGD, LGD, EAC | LGD, LGD, EAC |
| Real-time imaging | Yes | Yes | Yes | Yes | Yes | Yes |
| Endoscopy type | Olympus GIF-Q160 (SRE) versus Olympus Inc., Tokyo, Japan (NBI) | Olympus GIF-H180 endoscope (PCFQ180AL/I; Olympus, Tokyo, Japan) | Olympus GIF-Q240Z, Tokyo, Japan | Olympus GIF-QZ40Z, Tokyo, Japan | Olympus GIF-QZ40Z, Tokyo, Japan | Olympus GIF-QZ40Z, Tokyo, Japan |
| Endoscopist, N | 8 for SRE, 2 for NBI | NR | 2 | 2 | NR | NR |
| Blinded pathologist | Yes | Yes | No | Yes | NR | Yes |
| NBI magnification | No | Yes | Yes | Yes | Yes | Yes |
See Figures 2 and 3 to summarize each findings of methodological quality assessment of prevalence of IM. Exclusions.23,25,26,28 Three studies (25%) were judged at high risk or consecutive participant recruitment and avoided inappropriate bias. Eleven studies (91.6%) were judged to have low concerns for applicability, as the same clinical data were available when test results were interpreted as available when the test is used in practice.

Diagnostic accuracy of narrow-band imaging in intestinal metaplasia. Three studies reported the diagnostic accuracy of NBI for detecting IM.22–24 The pooled sensitivity of NBI for detecting IM was 0.83 (95% CI: 0.63–1.00), with individual study sensitivity ranging from 0.61 (95% CI: 0.47–0.74) to 0.94 (95% CI: 0.88–0.99). The specificity of NBI for detecting IM ranged from 0.85 (95% CI: 0.77–0.92) to 0.95 (95% CI: 0.83–1.00), with a pooled specificity of 0.89 (95% CI: 0.81–0.97). The pooled PPV of NBI for the detection of IM was 0.90 (95% CI: 0.72–1.00), with individual study PPV ranging from 0.69 (95% CI: 0.55–0.82) to 0.99 (95% CI: 0.98–1.00). The NPV of NBI for the detection of IM ranged from 0.53 (95% CI: 0.30–0.75) to 0.80 (95% CI: 0.46–0.85), with a pooled NPV of 0.66 (95% CI: 0.46–0.85). No study reported data for the median prevalence of IM.

Findings of methodological quality assessment of included studies. See Figures 2 and 3 to summarize each domain’s judgment for each included study.

Patient selection. We judged four studies (33.3%) to be at low risk of bias for participant selection regarding reported random or consecutive participant recruitment and avoided inappropriate exclusions.23,25,26,28 Three studies (25%) were judged at high risk of bias.22,24,27 Patients were not randomly selected or enrolled in two studies, and important exclusion of patients with erosive esophagitis was not avoided in one study. For the remaining five studies,29–33 no details were reported on patients’ sampling, and two studies did not provide enough details about exclusion criteria; hence, we assigned them to have an unclear risk of bias. Eleven studies (91.6%) were judged to have low concerns for applicability that the included patients and setting matched the review question. Only one study involved four phases, with only the second phase related to the review question, and thus judged with serious applicability concerns.

Index test
Narrow-band imaging to white-light standard endoscopy comparison. Nine studies compared NBI to WLSE, four of them (44%) were judged at low risk of bias for this domain in terms of interpreting the results of NBI without knowledge of the results of the reference standard and whether a threshold was pre-specified or not, if used.22,23,25,27 One study did not provide details about the blinded assessment of NBI and WLSE and judged at unclear risk of bias.24 The rest of the four studies (44%) were judged at high risk of bias due to un-blinded assessment of NBI about WLSE,29–33 and one study30 did not report the rationale for calculating sensitivity and specificity at the cut-off point (pit pattern 4 or 5) for detecting specialized IM by NBI endoscopy. These studies had low concerns for applicability, as the same clinical data were available when test results were interpreted as available when the test is used in practice.

Image to pathology comparison. Ten studies compared NBI endoscopic examination to histopathological findings. Eight (80%) were judged at low risk of bias for this domain in interpreting the results of NBI without knowledge of histopathological findings and whether a threshold was pre-specified or not if used.22,23,27,29,31–33 One study26 (10%) was judged at high risk of bias due to un-blinded assessment of NBI concerning histopathology, and one study30 did not report the rationale for calculating sensitivity and specificity at cut-off point (pit pattern 4 or 5) for detecting specialized IM by NBI endoscopy. All of these studies except one had low concerns for applicability, as the same clinical data were available when test results were interpreted as available when the test is used in practice.

Reference standard. Eight studies were assessed at low risk of bias.22,23,25,27,29,31–33 The pathologist was blinded to the endoscopic findings in seven studies. One study reported that all study patients first underwent standard resolution white-light endoscopy by experienced gastrointestinal endoscopists.25 Authors of another study23 stated that patients were evaluated using HD-WLE during the entire examination with no examination with NBI, and endoscopists kept blinded to the patient’s previous endoscopy and biopsy results. Nine studies showed low concerns for applicability. However, one study30 showed that the conventional endoscopy’s sensitivity and specificity were 24% and 67%, respectively. Thus, it did not provide a definitive indication of the presence or absence of specialized IM as a gold standard.
Moreover, independent blinded assessment by NBI endoscopy and conventional endoscopy and by the pathologist was unclear.

**Flow and timing.** Eight studies illustrated an appropriate interval between NBI endoscopic examination and the performance of the reference standard. In addition, all patients received the same reference standard, and all were included in the final analysis. One study reported that one patient was lost to follow-up without further details. Some patients were not included in the final analysis of the two studies. One study showed that patients were asked to repeat another open-access trans-oral upper gastrointestinal endoscopy over a long period, together with the Seattle protocol-guided biopsy, which may change the histopathological finding.

**Sensitivity and subgroup analyses.** We ran a sensitivity analysis for aspects that might affect the results of the meta-analysis, such as the risk of bias associated with the quality of included trials based on overall risk of bias assessment (low vs uncertain and high risk of bias) according to Quality Assessment of Diagnostic Accuracy Studies-2 patient selection, index test, reference standard, and flow–timing domains. We checked the effect of the following factors on the pooled sensitivity, specificity, PPV, and NPV. Firstly, we explored real-time imaging versus studies that use still images to report dysplasia and early EAC. Secondly, we explored studies that have used a validated endoscopic classification for detection and characterization of dysplasia and early EAC versus those that did not. Thirdly, we explored NBI with magnification versus NBI without magnification. None of these factors has significantly changed the overall sensitivity, specificity, PPV, or NPV of dysplasia detection.

**Discussion**

Effective endoscopic surveillance of BE is key in managing EAC and the only tool we have to reduce its rising incidence and high mortality when diagnosed in later stages. A random quadratic biopsy is the current standard of care but has several limitations. Targeted biopsy guided by advanced endoscopic imaging could improve the yield of detecting dysplasia and early EAC. These advanced imaging techniques’ ultimate goal is to improve on endoscopic detection of BE while reducing procedure time, expense, and sampling error. Electronic chromoendoscopy is increasingly replacing dye-based chromoendoscopy techniques in clinical practice because of the advantages of ease of use and safety. Targeted biopsies taken via NBI endoscopy are the most studied BE surveillance tool, with results showing high accuracy and precision. However, the random biopsy protocol, although outdated, continues to be recommended by all gastroenterology guidelines. In this meta-analysis, we researched to see if NBI could replace the current standard of care.

We pooled data from six studies (Table 1) that compared the NBI with random biopsy protocol using histology as a reference standard test. We reported the per-patient analysis as it is more tangible than per-lesion analysis. The overall sensitivity, specificity, NPV, and PPV for NBI-targeted biopsy detecting all dysplasia were 76%, 99%, 97%, and 84%, respectively. For HGD alone, the results were 83%, 99%, 97%, and 92%, with an area under the curve of 85%. However, the sensitivity of LGD detection by NBI was low. The pooled sensitivity across three studies that reported NBI detection of LGD was 0.60 (95% CI: 0.11–1.00), and the pooled PPV was 0.71 (95% CI: 0.30–1.00). The specificity and NPV were both 98%, similar to the performance of NBI in HGD.

We evaluated the incremental yield of improved detection of dysplasia using the NBI method to illustrate NBI endoscopy’s superiority over the Seattle WLSE protocol. Nonetheless, two studies provided sufficient data on granularity to calculate the NND LGD and HGD. The NND for these outcomes were 3.95 and 1.95, respectively. This indicates that for every four patients tested where no LGD was observed under WLSE, an additional patient would be identified as having LGD using NBI. For HGD, an additional two patients would need to be screened with NBI to identify one more patient with low-grade or high-grade dysplasia. This reduction in the NND for HGD is likely due to individuals with severe forms of LGD being reclassified under NBI as having HGD, which indicates additional gains using NBI.

The magnification function was added to the NBI mode to improve the ability to detect lesions. Of the six studies included, four used NBI with magnification. We pooled data from these studies in sensitivity analysis and observed no change in the overall sensitivity and specificity of detecting dysplasia.

The gold standard for assessing dysplasia in Barrett’s segment is the histology of a surgically resected segment. Due to its limitations in research and clinical contexts, multiple random biopsies coupled with histology represent a reasonable substitute as a reference standard test to calibrate the BE surveillance. We have included all studies that compared targeted NBI biopsy against the random biopsy. As such, we believe that studies included for quantitative meta-analysis have minimal differences.

The calculation of the diagnostic odds ratio between the NBI method and random biopsy as a primary outcome would have been a more robust output than the ROC of pooled data of NBI. However, we could not present pooled data from a meta-analysis as data on sensitivity, specificity, NPV, or PPV from the random biopsy method were not available in all included studies except one. Nonetheless, in this meta-analysis, the primary result of the pooled sensitivity and specificity of NBI-targeted biopsy should represent the possible diagnostic accuracy given the fact that in each included study, the raw values that were used to classify true/false positive and negative were produced based on comparisons against the method of random quadratic biopsies.

We found eight other studies that evaluated NBI-guided-targeted biopsy in BE. The authors showed high reproducibility and accuracy in detecting and characterizing dysplasia and IM in BE. However, these studies were not included in our meta-analysis, as they did not involve taking a random biopsy, which may have influenced two main aspects of validity. Firstly, the cross-classification of true and false negatives used to calculate the overall sensitivity and specificity might not have been accurate as biopsies were only taken from areas deemed an endoscopist to have unusual patterns indicating dysplasia or IM. Secondly, the absence of direct head-to-head comparisons between random and targeted biopsy methods makes it difficult to compare the outcomes.

Two meta-analyses have evaluated NBI-guided-targeted biopsy in BE and reported similar outcomes. Song et al. pooled
data across studies evaluating NBI in BE up to 2014. No study included in both meta-analyses used random biopsy as a comparator; hence, their results might not have been reflective of the accurate discriminate function of NBI. In another systematic review by ASGE, the yield of targeted biopsy method taken by NBI and chromoendoscopy and confocal laser endomicroscopy was pooled to calculate the overall sensitivity and specificity. Only four studies compared the targeted biopsies approach against random biopsies. We have identified and included two studies using the NBI approach that were not included in any published meta-analyses and used a different methodology.

Some of the limitations of our study include significant heterogeneity among studies meta-analyzed. This is quite common in meta-analyses of diagnostic test accuracy due to variations in diagnostic thresholds, inclusion, and exclusion criteria. No single classification system has been universally adopted in NBI endoscopy in BE surveillance, and each study used a different classification system. To address this discrepancy, where possible, we have used the Barrett’s International NBI Group classification system to determine the threshold for positive detection of dysplasia and MI.

In conclusion, we found 137 published clinical trials that evaluated NBI performance in BE surveillance, all of which showed better performance. We quantitatively analyzed six trials and narratively synthesized evidence from eight others that met our inclusion criteria. NBI-targeted endoscopic biopsy compared with random biopsy demonstrated higher diagnostic accuracy and reproducibility for detecting HGD and, to a lesser extent, LGD in BE surveillance. These findings, therefore, show that it may indeed be time to substitute random biopsies during endoscopic surveillance with targeted biopsies using NBI for detecting dysplasia in BE.

Acknowledgments

This work is part of a master’s thesis of the Master’s Program in Clinical Research, Center for Clinical Research and Management Education, Division of Healthcare Sciences, Dresden International University, Dresden, Germany. The authors thank Denny John (Campbell Collaboration, India) for writing a paragraph in the first draft proposal. The authors thank Simone Krautz (DIU) for providing access to some of the needed full-text journal articles.

References

1 Ferlay J, Soerjomataram I, Dikshit R et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBECAIN 2012. Int. J. Cancer 2015; 136: E359–86.

2 Herszényi L, Tulassay Z. Epidemiology of gastrointestinal and liver tumors. Eur. Rev. Med. Pharmacol. Sci. 2010; 14: 249–58.

3 Sundelöf M, Ye W, Dickman PW, Lagergren J. Improved survival in both histologic types of esophageal cancer in Sweden. Int. J. Cancer 2002; 99: 751–4.

4 Fitzgerald RC, di Pietro M, Ragunath K et al. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett’s esophagus. Gut 2014; 63: 7–42.

5 Bennett C, Moayyedi P, Corley DA et al. BEB CAT: a large-scale review and Delphi consensus for management of Barrett’s esophagus with no dysplasia, indefinite for, or low-grade dysplasia. Am. J. Gastroenterol. 2015; 110: 662–82 quiz 683.

6 Hvid-Jensen F, Pedersen L, Drewes AM, Sørensen HT, Funch-Jensen P. Incidence of adenocarcinoma among patients with Barrett’s esophagus. N. Engl. J. Med. 2011; 365: 1375–83.

7 Edgren G, Adami H-O, Weiderpass E, Nyrén O. A global assessment of the esophageal adenocarcinoma epidemic. Gut 2013; 62: 1406–14.

8 Napier KJ, Scheerer M, Misra S. Esophageal cancer: a review of epidemiology, pathogenesis, staging workup and treatment modalities. World J Gastrointest Oncol. 2014; 6: 112–20.

9 Sharma P, Dent J, Armstrong D et al. The development and validation of an endoscopic grading system for Barrett’s esophagus: the Prague C & M criteria. Gastroenterology 2006; 131: 1392–9.

10 Shaheen NJ, Falk GW, Iyer PG, Gerson LB. ACG clinical guideline: diagnosis and management of Barrett’s esophagus [Internet]. Am. J. Gastroenterol. 2016; 111: 30–50. Available from: https://doi.org/10.1038/ajg.2015.322

11 Bhat S, Coleman HG, Yousef F et al. Risk of malignant progression in Barrett’s esophagus patients: results from a large population-based study. J. Natl. Cancer Inst. 2011; 103: 1049–57.

12 Levine DS, Blount PL, Rudolph RE, Reid BJ. Safety of a systematic endoscopic biopsy protocol in patients with Barrett’s esophagus. Am. J. Gastroenterol. 2000; 95: 1152–7.

13 Abrams JA, Kapel RC, Lindberg GM et al. Adherence to biopsy guidelines for Barrett’s esophagus surveillance in the community setting in the United States. Clin. Gastroenterol. Hepatol. 2009; 7: 736–42 quiz 710.

14 Ngamuraengphong S, Sharma VK, Das A. Diagnostic yield of methylene blue chromoendoscopy for detecting specialized intestinal metaplasia and dysplasia in Barrett’s esophagus: a meta-analysis. Gastrointest. Endosc. 2009; 69: 1021–8.

15 Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. PLoS Med. 2009; 6: e1000097. https://doi.org/10.1371/journal.pmed.1000097 PRISMA [Internet], http://www.prisma-statement.org Available from:. Accessed 2020 Apr 22

16 Health Information Research Unit (HIRU). Search strategies for MEDLINE in Ovid Syntax and the PubMed translation [Internet]. [cited 2020 Apr 22]. Available from: https://hiru.mcmaster.ca/hiru/hiru_hedges_medline_strategies.aspx

17 Sharma P, Bergman J, Goda K et al. Development and validation of a classification system to identify high-grade dysplasia and esophageal adenocarcinoma in Barrett’s esophagus using narrow-band imaging. Gastroenterology 2016; 150: 591–8.

18 Whiting PF, Rutjes AW, Westwood ME et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann. Intern. Med. 2011; 155: 529–36.

19 Grunau G, Grunau P, Linn S, Leipsic J. How much net gain does a diagnostic imaging test provide? J. Magn. Reson. Imaging 2014; 39: 1339–40.

20 Hsu J, Brozek JL, Terracciano L et al. Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. Implement Sci. 2011; 6: 62.

21 Naakgeborne CA, van Enst WA, Ochodo EA et al. Systematic overview finds variation in approaches to investigating and reporting on sources of heterogeneity in systematic reviews of diagnostic studies. J. Clin. Epidemiol. 2014; 67: 1200–9. https://doi.org/10.1016/j.jclinepi.2014.05.018

22 Pascareno OD, Coroş MF, Pascareno G et al. A preliminary feasibility study: narrow-band imaging targeted versus standard white light endoscopy non-targeted biopsies in a surveillance Barrett’s population. Dig. Liver Dis. 2016; 48: 1048–53.

23 Sharma P, Hawes RH, Bansal A et al. Standard endoscopy with random biopsies versus narrow band imaging targeted biopsies in Barrett’s esophagus: a prospective, international, randomized controlled trial. Gut 2013; 62: 15–21.
24 Lee S-W, Lien H-C, Chang C-S, Lin M-X, Chang C-H, Ko C-W. Benefits of the Seattle biopsy protocol in the diagnosis of Barrett’s esophagus in a Chinese population [Internet]. World J. Clin. Cases 2018; 6: 753–8. Available from:. https://doi.org/10.12998/wjcc.v6.i14.753
25 Wolsøn HC, Crook JE, Krishna M et al. Prospective, controlled tandem endoscopy study of narrow band imaging for dysplasia detection in Barrett’s Esophagus. Gastroenterology 2008; 135: 24–31.
26 Jayasekera C, Taylor A, Desmond P, Macrae F, Williams R. Added value of narrow band imaging and confocal laser endomicroscopy in detecting Barrett’s esophagus neoplasia [Internet]. Endoscopy 2012; 44: 1089–95. Available from:. https://doi.org/10.1055/s-0032-1325734
27 Kara MA, Peters FP, Rosmolen WD et al. High-resolution endoscopy plus chromoendoscopy or narrow-band imaging in Barrett’s esophagus: a prospective randomized crossover study. Endoscopy 2005; 37: 929–36.
28 Nagales O, Caballero-Marcos A, Clemente-Sánchez A et al. Usefulness of non-magnifying narrow band imaging in EVIS EXERA III video systems and high-definition endoscopes to diagnose dysplasia in Barrett’s esophagus using the International NBI Group (BING) classification [Internet]. Dig. Dis. Sci. 2017; 62: 2840–6. Available from:. https://doi.org/10.1007/s10620-017-4581-3
29 Anagnostopoulos GK, Yao K, Kaye P, Hawkey CJ, Ragunath K. Novel endoscopic observation in Barrett’s esophagus using high resolution magnification endoscopy and narrow band imaging [Internet]. Aliment. Pharmacol. Ther. 2007; 26: 501–7. Available from:. https://doi.org/10.1111/j.1365-2036.2007.03774.x
30 Hamamoto Y, Endo T, Nosho K, Arinuma Y, Sato M, Imai K. Usefulness of narrow-band imaging for endoscopic diagnosis of Barrett’s esophagus. J. Gastroenterol. 2004; 39: 14–20.
31 Singh R, Karageorgiou H, Owen V et al. Comparison of high-resolution magnification narrow-band imaging and white-light endoscopy in the prediction of histology in Barrett’s esophagus. Scand. J. Gastroenterol. 2009; 44: 85–92.
32 Sharma P, Bergman J, Goda K et al. Development and validation of a classification system to identify high-grade dysplasia and esophageal adenocarcinoma in Barrett’s esophagus using narrow-band imaging [Internet]. Gastroenterology 2016; 150: 591–8. Available from:. https://doi.org/10.1053/j.gastro.2015.11.037
33 Singh R, Anagnostopoulos GK, Yao K et al. Narrow-band imaging with magnification in Barrett’s esophagus: validation of a simplified grading system of mucosal morphology patterns against histology. Endoscopy 2008; 40: 457–63.
34 Curvers W, Bohner C, Mallant-Hent R et al. Mucosal morphology in Barrett’s esophagus: interobserver agreement and role of narrow band imaging. Endoscopy 2008; 40: 799–805.
35 Herrero L, Curvers W, Bansal A et al. Zooming in on Barrett esophagus using narrow-band imaging: an international observer agreement study. Eur. J. Gastroenterol. Hepatol. 2009; 21: 1068–75.
36 Mannath J, Subramanian V, Hawkey CJ, Ragunath K. Narrow band imaging for characterisation of high grade dysplasia and specialized intestinal metaplasia in Barrett’s esophagus: a meta-analysis. Endoscopy 2010 May; 42: 351–9.
37 Song J, Zhang J, Wang J et al. Meta-analysis: narrow band imaging for diagnosis of gastric intestinal metaplasia. PLoS ONE 2014; 9: e94869.
38 ASGE Technology Committee, Thosani N, Abu Dayyeh BK et al. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE Preservation and Incorporation of Valuable Endoscopic Innovations thresholds for adopting real-time imaging-assisted endoscopic targeted biopsy during endoscopic surveillance of Barrett’s esophagus. Gastrointest. Endosc. 2016 Apr; 83: 684–98.e7.

Appendix A

Diagnosis of Barrett’s esophagus

Recommendations of American College of Gastroenterology Clinical Guideline 2015.

1 BE should be diagnosed when there is an extension of salmon-colored mucosa into the tubular esophagus extending ≥ 1 cm proximal to the gastro-esophageal junction (GEJ) with biopsy confirmation of IM (strong recommendation, low level of evidence).
2 Endoscopic biopsy should not be performed in the presence of a normal Z line or a Z line with < 1 cm of variability (strong recommendation, low level of evidence).
3 In the presence of BE, the endoscopist should describe the extent of metaplastic change, including circumferential and maximal segment length using the Prague classification (conditional recommendation, low level of evidence).
4 The location of the diaphragmatic hiatus, GEJ, and squamocolumnar junction should be reported in the endoscopy report (conditional recommendation, low level of evidence).
5 In patients with suspected BE, at least eight random biopsies should be obtained to maximize IM’s yield on histology.
6 In patients with short (1–2 cm) segments of suspected BE, eight biopsies may be unobtainable, at least four biopsies per cm of circumferential BE, and one biopsy per cm in tongues of BE should be obtained (conditional recommendation, low level of evidence).
7 In patients with suspected BE and lack of IM on histology, a repeat endoscopy should be considered in 1–2 years to rule out BE (conditional recommendation, very low level of evidence).

Appendix B

CENTRAL search strategy

#1 “Barrett Esophagus”
#2 “Barrett Esophagus”
#3 MeSH descriptor: [Barrett Esophagus] explode all trees and with qualifier(s): [diagnostic imaging - DG]
#4 MeSH descriptor: [] explode all trees and with qualifier(s): [diagnostic imaging - DG, diagnosis - DI]
#5 #1 OR #2 OR #3 OR #4
#6 “Narrow Band Imaging Endoscopy”
#7 MeSH descriptor: [] explode all trees
#8 (NBI Endoscopy):ti,ab,kw
#9 #6 OR #8
#10 Dysplasia
11 Adenocarcinoma
12 MeSH descriptor: [] explode all trees and with qualifier(s):
[diagnostic imaging - DG, diagnosis - DI]
13 #10 OR # 11 OR # 12
14 #5 AND #9 AND #13

**EMBASE search strategy via OVID (search date: 1974 to November 2018)**

1 exp. Barrett esophagus/
2 Barrett Osophagus.mp.
3 (barrett adj esophagus).ti,ab.
4 (barrett adj osophagus).ti,ab.
5 Endoscop$.ti,ab.
6 (Endoscopy adj Narrow Band adj Imag$).ti,ab.
7 exp. endoscopy/
8 prognos:.tw. OR survival.tw.
9 (cohort adj stud$).ti,ab.
10 (longitudinal adj stud$).ti,ab.
11 (prospective adj stud$).ti,ab.
12 Cohort analysis/
13 1 OR 2 OR 3 OR 4
14 OR 5 OR 6 OR 7
15 OR 8 OR 9 OR 10 11 OR 12
16 13 AND 14 AND 8 (this returns 710 papers);
17 13 AND 14 AND 15 (this returns 1457 papers)