Narrow band imaging to differentiate neoplastic and non-neoplastic colorectal polyps in real time: a meta-analysis of diagnostic operating characteristics

Sarah K McGill,1 Evangelos Evangelou,2 John P A Ioannidis,3 Roy M Soetikno,1 Tonya Kaltenbach1

ABSTRACT

Purpose Many studies have reported on the use of narrow band imaging (NBI) colonoscopy to differentiate neoplastic from non-neoplastic colorectal polyps. It has potential to replace pathological diagnosis of diminutive polyps. We aimed to perform a systematic review and meta-analysis on the real-time diagnostic operating characteristics of NBI colonoscopy.

Methods We searched PubMed, SCOPUS and Cochrane databases and abstracts. We used a two-level bivariate meta-analysis following a random effects model to summarise the data and fit hierarchical summary receiver-operating characteristic (HSROC) curves. The area under the HSROC curve serves as an indicator of the diagnostic test strength. We calculated summary sensitivity, specificity and negative predictive value (NPV). We assessed agreement of surveillance interval recommendations based on endoscopic diagnosis compared to pathology.

Results For NBI diagnosis of colorectal polyps, the area under the HSROC curve was 0.92 (95% CI 0.90 to 0.94), based on 28 studies involving 6280 polyps in 4053 patients. The overall sensitivity was 91.0% (95% CI 87.6% to 93.5%) and specificity was 82.6% (95% CI 79.0% to 85.7%). In eight studies (n=2146 polyps) that used high-confidence diagnostic predictions, sensitivity was 93.8% and specificity was 83.3%. The NPVs exceeded 90% when 60% or less of all polyps were neoplastic. Surveillance intervals based on endoscopic diagnosis agreed with those based on pathology in 92.6% of patients (95% CI 87.9% to 96.3%).

Conclusions NBI diagnosis of colorectal polyps is highly accurate—the area under the HSROC curve exceeds 0.90. High-confidence predictions provide >90% sensitivity and NPV. It shows high potential for real-time endoscopic diagnosis.

INTRODUCTION

Colonoscopy with polypectomy is considered effective at preventing colorectal cancer deaths. However, its expense makes it potentially less cost effective than other screening methods. All polyps, even diminutive polyps that rarely harbour dysplasia, are routinely sent for pathological evaluation, and this can incur major costs. Pathology diagnoses are needed to determine the patient’s interval to the next surveillance colonoscopy. Real-time endoscopic diagnosis, in which endoscopists diagnose polyp histology at the moment of identification, may have similar diagnostic operating characteristics as pathological evaluation at a significantly reduced cost.

Narrow band imaging (NBI) technology highlights the increased vasculature of neoplastic tissue and makes neoplastic polyps appear darker than surrounding mucosa. This allows for improved differentiation of polyps compared with that using white light. As such, an international classification to distinguish neoplasms such as adenomas from...
non-neoplasms such as hyperplastic polyps with the use of NBI has been developed and validated (figure 1). Real-time differentiation has been proposed as part of a ‘resect and discard’ strategy in which diminutive (≤5 mm) neoplastic polyps are resected without pathological evaluation, and diminutive rectosigmoid non-neoplastic polyps are left in situ. Such a strategy could allow for surveillance intervals to be communicated to the patient on the day of colonoscopy, and confer substantial cost savings by avoiding pathology and endoscopy fees. It could also decrease unnecessary polypectomy, a risk factor for major colonoscopy-related complications such as perforation and bleeding.

The potential for applying the ‘resect and discard’ strategy to clinical use depends on the accuracy of real-time endoscopic diagnosis. Numerous prospective studies have compared endoscopic diagnosis of polyps with NBI to a reference standard of histology. These data are valuable to allow us to understand its diagnostic operating characteristics, which, in turn, could provide insights of its potentials for clinical use. The aim of this study was to carry out a systematic review and meta-analysis on the real-time performance of endoscopic diagnosis.

METHODS
Search strategy and study selection
We (TK and SM) conducted a computerised literature search of the PubMed, SCOPUS (including EMBASE) and Cochrane Library databases up to June 2012. In addition, we searched the abstracts of conference proceedings of Digestive Diseases Week and American College of Gastroenterology. We designed the search queries with a biomedical research librarian to capture all articles related to NBI and colonoscopy. Studies in PubMed were identified with the terms narrow band or optical filter combined with the set operator AND with studies identified with the MeSH terms colonoscopy, colonic neoplasms or colonic polyps or with words beginning with colorect, adenoma, colonoscopy or polyp. Studies in SCOPUS were identified with the terms narrow band or optical filter combined with the set operator AND with words beginning with colorect, adenoma, colonoscopy or polyp. The electronic archives of conference proceedings were searched using the term narrow band, and abstracts that included this term were reviewed in detail for potential inclusion. We subsequently manually searched the citations from published reviews. Following the initial search, we identified articles for appropriateness, and performed a detailed full text assessment of potentially relevant studies.

We included studies that prospectively evaluated patients undergoing colonoscopy during which endoscopists used NBI to make a real-time prediction of polyp histology (neoplastic or non-neoplastic), and compared this with histology as the reference standard. We excluded studies that primarily analysed still images to predict polyp histology; those that included patients with inflammatory bowel disease, hereditary nonpolyposis colorectal cancer or familial polyposis syndromes; as well as those primarily designed as a retrospective study, review, editorial or meta-analysis. We translated papers that were not written in English. We discussed and resolved disagreement between investigators. Our protocol is available by request.

Data extraction
We extracted data independently onto standardised paper forms. We constructed 2×2 tables that contained the number of polyps found to be true positives (neoplastic polyps that were endoscopically predicted to be neoplastic), true negatives (non-neoplastic polyps that were endoscopically predicted to be non-neoplastic), false positives (non-neoplastic polyps that were endoscopically predicted to be neoplastic) and false negatives (neoplastic polyps that were endoscopically predicted to be non-neoplastic). When possible, additional 2×2 tables were constructed for polyps ≤5 mm, 6–9 mm and ≥10 mm, and high-confidence and low-confidence predictions. In general, a real-time diagnosis was described as high confidence when the polyp had endoscopic features strongly suggestive of its pathology, and the endoscopist could make a clinical management and surveillance decision based on his or her endoscopic diagnosis. We contacted study authors when there were insufficient published data to construct the 2×2 table for all results. In addition, the following data were extracted for each trial, when available: period of enrolment, number of endoscopists who individually performed colonoscopies on patients in the cohort, number of these who were expert endoscopists, total number of patients enrolled, number of patients with polyps examined by NBI for accuracy data, inclusion and exclusion criteria for participants; patient data including age, sex and indication for colonoscopy; size criterion for polyps examined for real-time diagnosis; polyp location, size and shape by Paris classification; and histology. We also extracted the diagnostic criteria used to differentiate neoplasms from non-neoplasms, and the features of the diagnostic modalities used, including NBI LUCERA or EXERA processor systems, high definition or high resolution, and the use of optical or digital magnification. Expert endoscopists were defined as those attending endoscopists with significant experience performing colonoscopy and using NBI.

Figure 1  Endoscopic image with narrow band imaging of diminutive colorectal polyps. (A) Adenoma and (B) non-neoplastic polyp.
Study quality assessment

To assess study quality and potential for bias, we used the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool. We rated the quality of key study design characteristics such as prespecified inclusion and exclusion criteria, defined diagnostic criteria, blinding of endoscopists to the pathological diagnosis, description of a reference standard and the inclusion of all patients in the analysis.

Statistical methods

We performed a bivariate meta-analysis using a linear mixed model approach to calculate summary estimates of sensitivity, specificity, positive likelihood ratio and negative likelihood ratio, and to fit a hierarchical summary receiver-operating characteristic (HSROC) curve. We used summary estimates of sensitivity and specificity to estimate the negative and positive predictive values when neoplasms accounted for 60%, 50%, 40% and 30% of all polyps based on prior published values. We performed prespecified subgroup analyses for published studies alone, diminutive polyps, high-confidence predictions, diminutive high-confidence predictions, diagnoses with NBI LUCERA and EXERA systems. To perform a sensitivity analysis, we performed subgroup analyses for high-quality and highest-quality studies, which were not prespecified. We calculated the area under the HSROC curve for the main analysis and subgroup analyses. We used random effects to calculate summary estimates for feasibility of high-confidence diagnoses and assessed agreement of surveillance interval recommendations based on endoscopic diagnosis compared with those based on the pathological diagnosis. We used Stata V.11 to perform the calculations.

RESULTS

Eligible studies

Twenty-eight studies consisting of 18 full published papers and 10 abstracts fulfilled the study inclusion criteria and were analysed. These 28 studies were selected from 711 screened citations. We excluded 658 citations based on the title and abstract. Of the 52 articles selected for full text review, 14 were excluded as they had predicted polyp histology based on still images, 5 were review articles, 4 were excluded because they involved a hereditary nonpolyposis colorectal cancer population, and 11 exclusions were based on other criteria (figure 2). Eventually, we identified 18 papers and 14 abstracts that met eligibility criteria. Four abstracts with a total of 541 polyps (359 patients) were subsequently excluded for missing data despite contacting the investigators.

Study characteristics

The 28 studies included a total of 6280 polyps that were diagnosed in 4053 patients. The characteristics of the included studies are listed in table 1. Fourteen studies were performed in the USA, seven in Asia, six in Europe and one in Australia. Twenty studies used the EXERA; eight used the LUCERA systems of NBI. The criteria used for prediction of polyp histology varied: six used Kudo pit pattern classification, four used the Narrow Band Imaging International Colorectal Endoscopic (NICE) classification, and others used less defined criteria of polyp colour or vascular pattern. Overall, 63% of all polyps analysed were neoplastic—the range was from 38% to 89%.

Quality assessment

In general, the included studies met most of the quality criteria (table 2). However, in many studies it was not clear whether a consecutive or random sample of patients was enrolled, and whether all patients were included in the analysis, both of which might have introduced bias.

Diagnostic performance of NBI diagnosis

Endoscopic diagnosis of colorectal polyps with NBI showed highly accurate diagnostic performance—the area under the HSROC curve was 0.92 (95% CI 0.90 to 0.94) (figure 3). For high-confidence predictions alone reported by eight studies, the area under the HSROC curve was 0.95 (95% CI 0.93 to 0.97), and for high-confidence predictions of diminutive polyps, the
<table>
<thead>
<tr>
<th>Study, year</th>
<th>Study no.</th>
<th>Country</th>
<th>Diagnostic modality</th>
<th>Diagnostic criteria</th>
<th>Endoscopists number</th>
<th>No. of experts</th>
<th>Patients analysed</th>
<th>Polyps analysed</th>
<th>No. of neoplasms/ No. of non-neoplasms</th>
<th>Polyps, n (%) with high confidence predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buchner et al (2010)</td>
<td>1</td>
<td>USA</td>
<td>Exera</td>
<td>Kudo</td>
<td>2</td>
<td>2</td>
<td>75</td>
<td>41</td>
<td>25/16</td>
<td></td>
</tr>
<tr>
<td>Henry et al (2010)</td>
<td>2</td>
<td>USA</td>
<td>Exera</td>
<td>Sano-Emura</td>
<td>1</td>
<td>1</td>
<td>52</td>
<td>126</td>
<td>67/59</td>
<td></td>
</tr>
<tr>
<td>Ignjatovic et al (2009)</td>
<td>4</td>
<td>UK</td>
<td>Lucera</td>
<td>Vascular pattern intensity</td>
<td>4</td>
<td>2</td>
<td>130</td>
<td>169</td>
<td>129/40</td>
<td></td>
</tr>
<tr>
<td>Lee et al (2011)</td>
<td>5</td>
<td>South Korea</td>
<td>Lucera</td>
<td>Mucosal pattern and vascular density</td>
<td>1</td>
<td>1</td>
<td>142</td>
<td>156</td>
<td>80/76</td>
<td>125 (80.1%)</td>
</tr>
<tr>
<td>Rastogi et al (2009)</td>
<td>7</td>
<td>USA</td>
<td>Exera</td>
<td>Type A: HP</td>
<td>1</td>
<td>1</td>
<td>101</td>
<td>236</td>
<td>143/93</td>
<td></td>
</tr>
<tr>
<td>Rastogi et al (2011)</td>
<td>8</td>
<td>USA</td>
<td>Exera</td>
<td>Type A: HP</td>
<td>6</td>
<td>6</td>
<td>134</td>
<td>384</td>
<td>147/237</td>
<td></td>
</tr>
<tr>
<td>Rex (2009)</td>
<td>9</td>
<td>USA</td>
<td>Exera</td>
<td>Pre NICE*</td>
<td>1</td>
<td>1</td>
<td>136</td>
<td>451</td>
<td>230/211</td>
<td>368 (81.6%)</td>
</tr>
<tr>
<td>Ronchino et al (2011)</td>
<td>11</td>
<td>Italy</td>
<td>Lucera</td>
<td>Kudo</td>
<td>3</td>
<td>3</td>
<td>94</td>
<td>281</td>
<td>141/140</td>
<td></td>
</tr>
<tr>
<td>Sakamoto et al (2012)</td>
<td>12</td>
<td>Japan</td>
<td>Lucera</td>
<td>Sano</td>
<td>1</td>
<td>1</td>
<td>80</td>
<td>116</td>
<td>52/42</td>
<td></td>
</tr>
<tr>
<td>Sano et al (2009)</td>
<td>13</td>
<td>Japan</td>
<td>Lucera</td>
<td>Sano</td>
<td>1</td>
<td>1</td>
<td>92</td>
<td>150</td>
<td>111/39</td>
<td></td>
</tr>
<tr>
<td>Shahid et al (2011)</td>
<td>14</td>
<td>USA</td>
<td>Exera</td>
<td>Sano</td>
<td>1</td>
<td>1</td>
<td>65</td>
<td>130</td>
<td>58/72</td>
<td></td>
</tr>
<tr>
<td>Singh et al (2011)</td>
<td>15</td>
<td>Australia</td>
<td>Lucera</td>
<td>Sano</td>
<td>1</td>
<td>1</td>
<td>32</td>
<td>50</td>
<td>30/20</td>
<td>50 (100%)</td>
</tr>
<tr>
<td>Zhou et al (2011)</td>
<td>16</td>
<td>China</td>
<td>Lucera</td>
<td>Kudo and Sano</td>
<td>1</td>
<td>1</td>
<td>118</td>
<td>109</td>
<td>67/42</td>
<td></td>
</tr>
<tr>
<td>Ren (2012)</td>
<td>17</td>
<td>China</td>
<td>Exera</td>
<td>Yoshi</td>
<td>1</td>
<td>1</td>
<td>75</td>
<td>116</td>
<td>52/42</td>
<td></td>
</tr>
<tr>
<td>Kuiper et al (2012)</td>
<td>18</td>
<td>The Netherlands</td>
<td>Exera</td>
<td>Kudo</td>
<td>3</td>
<td>0</td>
<td>108</td>
<td>281</td>
<td>141/140</td>
<td>231 (82.2%)</td>
</tr>
</tbody>
</table>

**Abstracts**

Coe et al (2012) | 19 | USA | Exera | NICE | 6 | 6 | 300 | 260 | 152/108 |
| Hewett and Rex (2010) | 20 | USA | Exera | Pre NICE* | 1 | 1 | 225 | 235 | 38/197 |
| Kaltenbach et al (2011) | 21 | USA | Exera | NICE | 2 | 2 | 220 | 236 | 146/90 | 178 (75.4%) |
| Kaltenbach et al (2012) | 22 | USA | Exera | NICE | 5 | 5 | 311 | 338 | 233/105 | 246 (72.8%) |
| Matthew et al (2008) | 23 | USA | Exera | Reported patterns | 1 | 0 | 100 | 231 | 106/125 |
| Occhipinti et al (2011) | 24 | Italy | Exera | Kudo | 5 | 5 | 93 | 220 | 120/100 |
| Pohl et al (2012) | 25 | USA | Exera | NICE | 10 | NR | 608 | 948 | 528/420 | 770 (81.3%) |
| Radaelli et al (2011) | 26 | Italy | Exera | NR | NR | all | 197 | 354 | 233/121 | 354* |
| Ringold et al (2008) | 27 | USA | Exera | ‘Cerebriform’ | 4 | 4 | 55 | 93 | 56/37 |

*Information on the number of polyps predicted with low confidence was not provided.

NICE, Narrow Band Imaging International Colorectal Endoscopic Classification. This classification system includes criteria on polyp colour, vessels and surface pattern; NR, not reported; Pre NICE, classification systems based on similar criteria that preceded NICE.
<table>
<thead>
<tr>
<th>Author</th>
<th>DOM1A1</th>
<th>DOM1A2</th>
<th>DOM1A3</th>
<th>DOM1A4</th>
<th>DOM1B</th>
<th>DOM2A1</th>
<th>DOM2A2</th>
<th>DOM2A3</th>
<th>DOM2B</th>
<th>DOM3A1</th>
<th>DOM3A2</th>
<th>DOM3A3</th>
<th>DOM3B</th>
<th>DOM4A1</th>
<th>DOM4A2</th>
<th>DOM4A3</th>
<th>DOM4A4</th>
<th>DOM4A5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buchner et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Henry et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Hirota et al</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>H</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Ignjatovic et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>H</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Lee et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Machida et al</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Rastogi et al</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Rastogi et al</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>Rex et al</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>N</td>
<td>U</td>
<td>L</td>
</tr>
<tr>
<td>Rogat et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Rotondano et al</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>H</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Sakamoto et al</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>H</td>
<td>H</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Sano et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Shahid et al</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>Singh et al</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Zhou et al</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Ren et al</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>Kuiper et al</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Abstracts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coe et al</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Hewett et al</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Kalterbach et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Kalterbach et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Matthew et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Occhipinti et al</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>Y</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>Pohl et al</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Radaelli et al</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>L</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Ringold et al</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>L</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>Yague et al</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>Y</td>
<td>Y</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
</tr>
</tbody>
</table>

DOM, domain; DOM1A1, consecutive or random sample of patients enrolled; DOM1A2, case–control design avoided; DOM1A3, study avoided inappropriate exclusions; DOM1A4, selection of patients introduced bias; DOM1B, concern that included patients do not match the review question; DOM2A1, index test results interpreted without knowledge of results of reference standard; DOM2A2, prespecified threshold used; DOM2A3, could the conduct of index test have introduced bias; DOM2B, concern that index test conduct or interpretation differ from review question; DOM3A1, reference standard correctly classified condition; DOM3A2, reference standard results interpreted independently from index test results; DOM3A3, could the reference standard, its conduct, or its interpretation have introduced bias; DOM3B, concern that target condition defined by reference standard does not match review question; DOM4A1, appropriate interval between index test and reference standard; DOM4A2, all patients received the reference standard; DOM4A3, patients received the same reference standard; DOM4A4, all patients included in the analysis; DOM4A5, could the patient flow have introduced bias; H, high; L, low; N, no; U, unclear; Y, yes.
The area under the HSROC curve was 0.92 (95% CI 0.92 to 0.96) (figure 4).

The main results are presented in table 3. The overall sensitivity of NBI diagnosis was 91.0% (95% CI 87.6% to 93.5%) and specificity was 82.6% (95% CI 79.0% to 85.7%) compared with histology. In six studies (n=1567 polyps) that provided information on high-confidence and low-confidence predictions, a high-confidence prediction was made in 77% of polyps (95% CI 73.3% to 80.7%). In eight studies (n=2146 polyps) that provided information on high-confidence predictions, the sensitivity and specificity for diagnosis made with high confidence was 93.8% and 83.3%, respectively. The sensitivity and specificity of diagnosis of diminutive polyps, made with high confidence was 93.4% (95% CI 87.4% to 96.7%) and 84.0% (95% CI 76.6% to 89.3%), respectively. The negative predictive values (NPVs) overall were 86.0%, 90.1%, 93.2% and 95.5% and for high-confidence predictions improved to 90.0%, 93.0%, 95.3% and 96.9% when neoplasms account for 60%, 50%, 40% and 30% of all polyps, respectively.

The likelihood ratio synthesis gave an overall LR+ of 5.2 (95% CI 4.3 to 6.4) and LR− of 0.11 (95% CI 0.08 to 0.15). Results were similar when limited to published papers, and improved for high-confidence and diminutive polyps.

Surveillance intervals based on NBI diagnosis

Seven studies—three papers20 22 33 and four abstracts38–41—reported surveillance intervals based on endoscopic diagnosis.

Overall, surveillance intervals based on endoscopic diagnosis were the same as those based on formal pathology in 92.6% of patients (95% CI 87.9% to 96.3%). There was statistically significant heterogeneity in the agreement rates across studies (p=4×10−8 for Q statistic). If we exclude the studies by Radaelli et al22 and Kuiper et al,33 who reported significantly less agreement at 83.2% and 81.4%, respectively, heterogeneity is not significant (p=0.28 for Q statistic) and the surveillance intervals agree in 95.5% of patients (95% CI 93.9% to 96.9%). Among the six studies that reported whether endoscopically directed intervals were shorter or longer than those directed by pathology,20 22 33 38 39 41 39 patients (3.0%) were prescribed shorter intervals and 35 patients (2.7%) were prescribed longer intervals (figure 5).

Sensitivity analysis

We performed sensitivity analyses with five studies rated of the highest quality (all risk for bias low) and 12 of high quality (all risk for bias low or low except one which was unknown). Diagnostic operating characteristics were similar compared with all results (table 3, figure 6).

DISCUSSION

The findings of this meta-analysis suggest that real-time endoscopic diagnosis of colorectal polyps performed using NBI has a high diagnostic performance, with an area under the HSROC curve exceeding 0.90. Endoscopic diagnosis correctly characterised 91% of neoplasms and 83% of non-neoplastic polyps, and these numbers improved with high-confidence predictions.
to 94% and 83%, respectively, though the improvement was not statistically significant. To our knowledge, this is the first meta-analysis to solely assess diagnosis of colorectal polyps using NBI in real time, which is the closest to simulate actual practice, rather than using still images as well. We included published papers and abstracts, since a substantial portion of this literature is recent and may not yet be fully published. Our use of a two-level statistical approach with the bivariate meta-analysis and HSROC curve analyses following a random effects model allows test stringency and test accuracy to vary across studies. The technique allows researchers to avoid misleading conclusions that can occur when analyses following a random effects model allow test stringency and test accuracy to vary across studies. The technique allows researchers to avoid misleading conclusions that can occur when analyses following a random effects model allow test stringency and test accuracy to vary across studies.

Table 3: Diagnostic accuracy of endoscopic diagnosis with NBI to distinguish between neoplastic and non-neoplastic colorectal neoplasms

<table>
<thead>
<tr>
<th>Summary estimates (95% CI)</th>
<th>Likelihood ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study characteristics</td>
<td>No. of studies (no. of polyps)</td>
</tr>
<tr>
<td>All</td>
<td>28 (6280)</td>
</tr>
<tr>
<td>Published manuscripts</td>
<td>18 (3212)</td>
</tr>
<tr>
<td>High-confidence predictions</td>
<td>20 (5148)</td>
</tr>
<tr>
<td>Exera</td>
<td>20 (5148)</td>
</tr>
<tr>
<td>Lucera</td>
<td>8 (1132)</td>
</tr>
<tr>
<td>Highest-quality studies</td>
<td>17 (29)</td>
</tr>
<tr>
<td>High-quality studies</td>
<td>17 (21)</td>
</tr>
</tbody>
</table>

HSROC, hierarchical summary receiver-operating characteristic; NBI, narrow band imaging.
estimates of test sensitivity and specificity are simply pooled or averaged. Our analysis is the largest to address the topic; a prior meta-analysis by Wu et al included only 11 prospective studies, 6 of which differentiated polyps based on still images. The only real-time diagnosis study in a meta-analysis by van den Broek et al excluded non-neoplastic polyps from its analysis.

Our study addresses the standards set forth by the American Society of Gastrointestinal Endoscopy (ASGE) for the ‘resect and discard’ strategy. The ASGE has proposed that, for endoscopically diagnosed adenomas ≤5 mm in size to be resected and discarded without pathologic assessment, endoscopic diagnosis should provide ≥90% agreement in assignment of post-polypectomy surveillance intervals compared with decisions based on pathology. The summary agreement was 92.6% in our study, supporting the clinical use of such a strategy.

Our results also relate to the ‘diagnose and leave in’ strategy of leaving endoscopically predicted diminutive non-neoplastic recto-sigmoid polyps in situ without resection. The ASGE recommended that to implement such a policy, endoscopic diagnosis should provide ≥90% NPV when used with high confidence. Our calculated NPV for high-confidence NBI diagnosis, derived using the summary sensitivity and specificity, meets this requirement when neoplastic polyps account for less than 60% of all polyps. This is true in prior studies that have reported polyp findings in the entire and recto-sigmoid colon. Lieberman et al found that 50% of 3744 diminutive polyps were neoplastic using a large database of colonoscopies performed at academic and private practices. Pickhardt et al found that 36.5% of 977 diminutive polyps were neoplastic; Hewett et al found that only 20% of 235 rectosigmoid diminutive polyps were neoplastic.

The rationale for real-time histology of colorectal polyps is based on the literature that diminutive polyps rarely harbour advanced histology such as villous features, high-grade dysplasia or cancer, and that pathologists are 85–95% accurate in polyp histology characterisation. Nonetheless, shifting colorectal polyp diagnosis from the traditional gold standard pathology to endoscopy is likely to cause much reservation across the field in regards to accountability, particularly in the strategy of forgoing resection. The importance of high-quality photo documentation of the polyp, in lieu of a pathology slide, must be underscored. The high-resolution polyp image should be permanently stored with the procedure report and accessible in the future as the reference to support the endoscopists’ assessment and clinical decision when subjected to quality review.

The agreement of surveillance intervals is important. In theory, longer surveillance intervals could allow for progression of neoplasia in high-risk individuals before follow-up while inappropriately short surveillance intervals could result in unnecessary colonoscopy, making the strategy less cost effective. In our study, when endoscopy-directed surveillance intervals disagreed with those dictated by pathology, roughly half the recommended longer follow-up and the other half, earlier follow-up. The majority of studies that reported surveillance intervals met the threshold of concordance for high-confidence predictions suggested by the ASGE. There were two exceptions that contributed to between-study heterogeneity. Kuiper and coauthors’ study, which uniquely focused on non-academic endoscopists, suggests they may have diminished accuracy compared with experts and require more or continued learning to achieve and sustain high performance. However, it is difficult to draw such a conclusion from a single study. The second study by Radaelli et al assesses expert endoscopists and it is unclear from the abstract why they had relatively worse agreement with histology than the other studies.

Recent evidence suggests that serrated polyps with features of larger size, sessile serrated polyp histology, cytologic dysplasia and location in the proximal colon may have a higher risk of colorectal cancer. As such, guidelines for colonoscopy surveillance after screening and polypectomy have recently been updated to stratify serrated lesions. The consensus recommended that a sessile serrated polyp larger than 9 mm and a sessile serrated polyp with cytological dysplasia should be managed like a high-risk adenoma; and serrated polyps that are less than 10 mm and do not have cytological dysplasia can be managed like a low-risk adenoma. Current endoscopic classification systems for the differentiation of polyp histology, however, do not address how to distinguish sessile serrated polyps from hyperplastic lesions, mainly due to the high inter-observer variability for their pathological diagnosis, and consequent lack of a reference standard. In the future, serrated polyps will need to be formally studied and integrated into the real-time histology strategy.

Our results reflect the important use of confidence levels. The highest performance of real-time NBI diagnosis of colorectal polyps was achieved when the diagnosis was made with high confidence—the area under the HSROC curve was 0.95 (95% CI 0.93 to 0.97) for polyps of any size, and 0.92 (95% CI 0.92 to 0.96) for diminutive ones. This compares to the overall area under the HSROC curve of 0.92 (95% CI 0.90 to 0.94). The use of objective validated criteria for the endoscopic differentiation of colorectal polyp histology can guide the physician in determining the confidence level. For example, a high-confidence diagnosis can be made if the polyp has one or more features as described in the classification and no features associated with the other polyp histology. If features are not present, then a low-confidence diagnosis can be made, and the polyp submitted for pathological assessment.

We analysed data of over 6000 polyps among 4053 patients who spanned four continents. Nonetheless, the main limitation to our study is that of generalisability. Most studies were performed with expert endoscopists who used variable classification methods to differentiate adenomas from non-neoplastic polyps. The results could be widely generalised if non-experts can readily learn endoscopic diagnosis; only preliminary studies exist, and results are mixed. Furthermore, the recent publication of validated international criteria specifically developed for colorectal polyps differentiation using NBI holds promise for making endoscopic diagnosis more standardised and less operator dependent. Teaching tools such as videos and computer modules may be important in the initial and continued training of endoscopists on polyp differentiation.

The exact number of NBI procedures needed to achieve proficiency in optical diagnosis is currently unknown. We identified two studies that address the potential learning curve for real-time endoscopic diagnosis. In Rogart et al’s series of 265 polyps in 131 patients, three of four endoscopists significantly improved their accuracy in the second half of the study compared with the first from 74% to 87%, respectively. In that study, endoscopists received feedback every 2 weeks about the accuracy of their endoscopic diagnosis. This suggests that with frequent feedback among experienced endoscopists, significant gains in accuracy can be attained in a relatively short amount of time. The other study was a single endoscopist study, which showed higher accuracy for polyps evaluated in the second half of study than the first. The difference however was not statistically significant.
Another potential limitation is publication bias. We searched for information included in meeting abstracts. We cannot totally exclude that some studies with poor diagnostic performance may have remained unpublished and not even presented in major meetings. However, these studies would have to be large to change substantially the results of the meta-analysis.

In conclusion, we found that endoscopic diagnosis with NBI is an accurate test to differentiate neoplastic from non-neoplastic polyps, with an area under the HSROC curve exceeding 0.90. Its sensitivity for diagnosing adenomatous and other neoplastic polyps is >90%, as is the agreement between endoscopically derived and histologically derived surveillance intervals. Its NPV is high as long as malignancies are highest with high-concentration neoplasms accounting for <60% of polyps. Endoscopic diagnostic performance is highest with high-confidence predictions.

Acknowledgements The authors would like to thank Christopher Stave, MLS, for his design of the search terms for the systematic review and Aparna Motiwala for her assistance with the database.

Contributors SM, TK, JPAI and RS contributed to the study concept and design, SM, TK and JPAI designed the protocol and data collection tools, and SM and TK performed the search and extracted data. EE, SM and TK analysed the data. SM drafted the manuscript, which was revised by all coauthors. All authors accepted the final version of the manuscript. TK is guarantor.

Competing interests RS and TK: Olympus research support and consultancy; RS: Olympus speaker’s fees. Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) licence, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/

REFERENCES

39 Kaltenbach T, Rastogi A, Rouse RV, et al. The VALID colonoscopy study—results of a multicenter prospective randomized controlled trial on real time colorectal polyp diagnosis using narrow band imaging (NBI) [abstract]. Gastrointest Endosc 2012;75 (Suppl 4):AB151 [569].
Meshed capillary vessels by use of narrow-band imaging for differential controlled study.

Endoscopy

neoplastic and nonneoplastic colonic polyps.


et al.

narrow band imaging for the diagnosis of colorectal neoplasms.


et al.


et al.

et al.

value of magnifying endoscopy in diagnostic test accuracy evaluations.

Endoscopy

et al.

et al.

et al.

et al.

et al.

et al.

et al.


et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.

et al.