The role of endoscopic imaging for an improved diagnosis of colorectal neoplasia

van den Broek, F.J.C.

Citation for published version (APA):
van den Broek, F. J. C. (2010). The role of endoscopic imaging for an improved diagnosis of colorectal neoplasia
CHAPTER 2
Systematic review of narrow band imaging for the detection and differentiation of neoplastic and non-neoplastic lesions in the colon

Frank J.C. van den Broek
Johannes B. Reitsma
Wouter L. Curvers
Paul Fockens
Evelien Dekker

Introduction

In the past decades, the perspective of diagnostic endoscopy has changed from diagnosing evident disease to the detection of subtle abnormalities. Modern endoscopic practice predominantly focuses on early detection and treatment of premalignant neoplasia, thereby interrupting the progression into unfavourable stages of malignancy. Endoscopic surveillance of high risk individuals has shown to be effective for preventing late stage esophageal and gastric cancer, as well as preventing colorectal cancer. Although the detection of gastrointestinal neoplasia at an early and curable stage is of crucial clinical importance, the endoscopic visualization of early neoplasia can be difficult, which may lead to neoplasia miss-rates and interval cancers between two successive endoscopies.

New imaging modalities, such as narrow band imaging (NBI), may allow for better detection of these early neoplastic lesions and thus improve the effectiveness of endoscopic surveillance and screening. Since NBI utilizes short wavelength (essentially blue) endoscopic light, which penetrates the mucosa only superficially and is mainly absorbed by hemoglobin, this technique highlights mucosal surface patterns and microvascular details (see figures 2-5; and video). The enhanced contrast provided by NBI can theoretically improve the detection of small and subtle mucosal lesions. In addition to a possible effect on detection, NBI has the potential for endoscopic differentiation of lesions (‘endo-pathology’) as well. Accurate differentiation during ongoing endoscopy has the advantage that on-table decisions can be made. In case of a presumed premalignant lesion, endoscopic resection may be performed immediately, whereas innocent non-neoplastic lesions may be left in situ.

Only recently, NBI has become commercially available leading to an accelerated uptake throughout the world. This systematic review summarizes the data on the performance and clinical utility of NBI during colonoscopy. The aims of this systematic review were to evaluate NBI concerning both the detection of premalignant lesions and the differentiation between neoplastic and non-neoplastic lesions.

Criteria for selecting studies

Eligible studies
Clinical trials or observational studies, assessing the performance of NBI colonoscopy with regard to the detection and/or the differentiation of lesions in the colon, were eligible for inclusion in this review. All patients undergoing colonoscopy were included, regardless of indication (screening, surveillance, symptoms).

Outcome measures
Studies reporting on the detection of neoplasia (adenomas) with NBI had to contain information on at least one of the following outcome measures: (1) the total number of detected neoplastic lesions; (2) the number or proportion of detected patients with at least one neoplastic lesion. With respect to these outcome measures of detection, the ideal comparison would be the histopathological evaluation of a complete colectomy specimen after performing NBI. Since this
comparison is unfeasible, NBI was compared to white light endoscopy (WLE) which is con-idered as current standard for the detection of neoplastic lesions.

With respect to studies reporting on *differentiation* between neoplastic and non-neoplastic lesions, the classification by NBI was compared against the reference standard which is the histopathological examination of biopsies or endoscopic resection specimen of the lesion of interest. Different classification systems co-exist for NBI and all of them were included in this review.\textsuperscript{10-14} Studies should provide sufficient data to construct 2x2 tables comparing whether the target condition (neoplasia or adenoma) was present according to NBI (index test) against the final histopathology of the same lesion (reference standard).

Information on inter- and intra-observer agreement with respect to the classification by NBI was extracted, if presented.

**Search strategy and selection of studies**

**Search strategy**

Searches were conducted on published literature without calendar year restrictions. Studies were identified through electronic searches of Pubmed and EMBASE. In addition, reference lists of included studies were scanned for additional relevant studies (citation tracking).

The following strategy was used to search PubMed [search last conducted 3 April, 2008]: Narrow band imaging OR (Narrow band AND Endoscop*) OR (Narrow band AND colonoscopy*). The following strategy was used to search EMBASE [search last conducted 3 April, 2008]: Narrow band imaging.mp. OR ((Narrow band.mp.) AND (exp INTESTINE ENDOSCOPY/ or exp ENDOSCOPY/ or exp DIGESTIVE TRACT ENDOSCOPY/ or exp GASTROINTESTINAL ENDOSCOPY/)) OR ((Narrow band.mp.) AND (colonoscopy.mp. or exp COLONOSCOPY/)).

**Selection of studies**

Two authors (FvdB and WC) independently screened the titles and abstracts of all studies identified by the abovementioned search strategy and obtained the full articles for all potentially relevant studies. Only articles published in English were included. The full text of these reports was assessed independently for eligibility. Any disagreement between the two assessors was resolved by discussion.

**Assessment of methodological quality**

The same two authors independently assessed the methodological quality of the included studies. The following quality items were scored for manuscripts reporting data on the *detection* of abnormalities: concealment of allocation, blinding of observers, patient selection and comparability of groups. A validated quality assessment tool for diagnostic accuracy studies (QUADAS) was used to assess the quality of studies reporting data on the *differentiation* of abnormalities.\textsuperscript{15} The QUADAS consist of the following quality items: patient selection (item 1-2), appropriateness of reference standard (item 3-7), possibility for replication of the index test and reference standard (item 8-9), blinding (item 10-11) and applicability of the results in daily practice (item 12-14).\textsuperscript{15} Each item of the QUADAS checklist was scored as “yes”, “no”, or “unclear”.
Data extraction
A data collection form was developed to extract relevant information from each included study. Two authors extracted the data separately and resolved differences by discussion until consensus was achieved. Data were extracted concerning study design, aim of study, patient selection, patient characteristics, allocation method and sources of bias. Concerning the detection of abnormalities, the numbers of detected neoplastic lesions and/or the numbers of patients with detected neoplasia were extracted. Concerning the differentiation of lesions, the total number of true positive (positive findings by NBI confirmed by histology), false positive (positive findings by NBI not confirmed by histology), true negative (negative findings by NBI confirmed by histology), and false negative findings (negative findings by NBI not confirmed by histology) were noted.

Data analysis

Detection
Two different outcomes were meta-analyzed for randomized studies comparing the detection capability of NBI with WLE: the proportion of patients with neoplasia (at least one adenoma) and the mean number of neoplastic lesions per patient.

The (logit transformed) odds ratio was used to compare the proportion of patients with detected neoplasia between NBI and WLE. The random effects approach of DerSimonian and Laird was used to calculate a pooled estimate with corresponding 95% confidence interval (95%-CI). This random effects model takes into account differences in sample size between studies (more weight to larger studies) and it properly incorporates any variability beyond chance in results between studies. Individual and pooled odds ratios were presented using Forrest plots.

The ratio of the mean number of detected neoplastic lesions by NBI relative to the mean number of detected neoplastic lesions by WLE was also used as outcome measure. The 95%-CI around this ratio was calculated assuming that the number of detected lesions would follow a Poisson distribution. A random effects pooling was performed if sufficient studies were available.

Differentiation
Each index test result by NBI was treated as dichotomous: suspicious (positive index test result) or unsuspicious (negative index test result) for neoplasia (adenoma). If several classification systems were used within one study for evaluating the index test (for example mucosal as well as vascular pattern), the mean number of true positive and negative index test results was used in the pooled analysis. Two-by-two tables were constructed showing the cross-classification of the index test results (NBI) versus the reference standard (histology) in order to calculate sensitivity and specificity (i.e. the underlying parameters of our analyses). To visualize data, Forrest plots were produced showing pairs of sensitivity and specificity together with 95% confidence intervals from each study. Confidence intervals were calculated using the Wilson score method. All analyses were done on a per-lesion basis, ignoring any possible correlation that might exist between multiple lesions within a patient.

If three or more studies were present, pairs of sensitivity and specificity were jointly meta-analyzed using a bivariate random effects approach.16 This approach enabled the calculation
of summary estimates of sensitivity and specificity while dealing with sources of within- and between-study variation and any correlation that might exist between sensitivity and specificity. The within-study variation or precision by which sensitivity and specificity have been measured was directly modeled based on the binomial distribution. The NLMIXED procedure of SAS was used to fit these bivariate random models.

**Results**

**Description of studies**

A total number of 34 reports on NBI for colonic use were retrieved by the search strategy. Eighteen manuscripts were reviews, commentaries or case reports. The remaining 16 studies were clinical evaluations of NBI, either concerning the detection of neoplasia (n=5), detection and differentiation of polyps (n=1) or concerning the differentiation of lesions only (n=10). Only one of those studies assessed the value of NBI for the detection of neoplasia in patients with longstanding ulcerative colitis, and two studies reported on differentiation in these patients. The characteristics of all clinical studies are represented in table 1 and the methodological quality (QUADAS scores) of the studies on differentiation is represented in table 2.

**Detection**

Three large randomized trials have been performed, which directly compared NBI to WLE concerning the detection of colorectal adenomas. In the study by Inoue et al, 243 patients were assigned to undergo either NBI or WLE. This was the only study that demonstrated a significantly improved adenoma detection rate by NBI versus WLE (mean number of adenomas per evaluated patient of 0.84 versus 0.55; p=0.046) (Table 3). When comparing the proportions of patients with at least one adenoma between NBI and WLE, no advantage for NBI could be demonstrated. In this study, however, an insufficient allocation method caused inadequate distribution of NBI procedures among all participating endoscopists. One endoscopist (who performed most of the colonoscopies) was significantly more often allocated to NBI when compared to all the others (Chi square test: p<0.001). Detected differences in this study may therefore be attributed to the experience of this single endoscopist as well as to the use of NBI.

On the contrary, Rex et al and Adler et al could not demonstrate an increased adenoma detection rate (both per-lesion and per-patient) by NBI in two large (434 and 401 subjects) randomized studies. These studies were well designed large trials with sufficient quality concerning allocation, patient selection and comparability of randomization groups; blinding of endoscopists for the procedures is not feasible in this kind of studies. In the study by Adler et al, an initial improved detection of patients with adenoma(s) was equalized in the later phase of the study when gaining more experience with NBI, thereby postulating a learning effect from NBI that resulted in improved detection with WLE.

Some differences existed between the 3 randomized studies. Rex et al used high definition monitors which may improve adenoma detection compared to standard monitors. In addition, differences in NBI-systems, inclusion criteria and experience of endoscopists between the three
systematic review of narrow band imaging studies may make pooling of the results less valid. The pooled results of the 3 randomized studies revealed a non-significant increase of patients with at least one adenoma (odds ratio 1.19; 95%-CI: 0.86-1.64) or total number of adenomas (OR 1.23; 95%-CI: 0.93-1.61) when NBI was used for detection (table 3).

In addition to the randomized trials, two cross-sectional back-to-back colonoscopy studies have been performed using WLE as primary detection technique during the first pass and additional NBI during the second pass.\textsuperscript{38, 40} No randomization was done for the sequence of the detection techniques. In the study by East \textit{et al}, patients with hereditary non polyposis colorectal cancer syndrome (including Lynch syndrome) were consecutively included, whereas patients undergoing colorectal cancer screening were included in the study by Rastogi \textit{et al}. Since no randomization was done and NBI was always performed during the second pass, the results of NBI from those studies could not reliably be compared to the abovementioned randomized trials.

In the two back-to-back colonoscopy studies, the adenoma miss-rates of WLE were 40\% (29/72) and 46\% (21/46), respectively, when an additional inspection was done with NBI (table 3).\textsuperscript{38, 40} When comparing these data to historical adenoma miss-rates of WLE from a systematic review (overall miss-rate of 22\%; 95%-confidence interval 15\% to 32\%), both studies demonstrated higher miss-rates.\textsuperscript{6} This suggests NBI to have an additional yield of adenomas when compared to WLE alone. However, since other endoscopy systems (lower quality charge coupled device chips) were used in historical studies and adenoma miss-rates do not take into account the total number of adenomas detected, a formal conclusion cannot be drawn from these studies.

Only one study by Dekker \textit{et al} reported on the use of NBI for the detection of neoplasia in patients with longstanding ulcerative colitis.\textsuperscript{36} In this study, a first prototype NBI system was compared to standard resolution WLE (CF-140 or CF-160 endoscopes; Olympus Inc.). All included patients underwent 2 colonoscopies: one with NBI for targeted biopsies only, and one with WLE for targeted plus random biopsies. Those procedures were performed with a time period of at least 3 weeks between the two examinations and the order of the procedures was randomized. This pilot study did not contain a predefined sample size and therefore included only 42 patients. In addition, a first prototype NBI system was used, having less brightness (Evis CLV-U40, Olympus Inc.) compared to current systems. In this study, NBI did not lead to an increased detection of neoplasia. On the contrary, both diagnostic strategies (NBI with targeted biopsies \textit{vs}. WLE with targeted plus random biopsies) did not detect 4 out of 12 patients with neoplasia. Therefore, the sensitivity of NBI for detecting a patient with neoplasia was only 67\%.

\textbf{Differentiation}

Nine studies reported on the use of NBI for differentiating neoplastic from non-neoplastic colonic polyps.\textsuperscript{11-14, 41-43, 45} All studies concerned post-hoc image evaluation studies, except for the study by Katagiri \textit{et al} \textsuperscript{45} in which instant diagnoses were made during ongoing endoscopy. Several classification systems were used for assessments with NBI, all encompassing either mucosal (pit pattern) or vascular pattern (vascular pattern intensity, brown hue, dense vessels, irregular vessels).

Quality assessments were made for all differentiation studies of colonic polyps based on the QUADAS checklist and are represented in table 2. For all evaluated polyps, final histopathology
was used as reference standard which was performed blinded (or was unclear for blinding) in all studies. Details of the histopathological evaluation procedure were not provided in three studies and objective information on the assessment with NBI was not provided in four and unclear in one study (e.g. assessment of ‘brown hue’ or ‘dense vessels’). In addition, blinded evaluation of NBI (without knowledge of the reference standard) was not stated, and therefore scored as unclear, in four studies. None of the studies reported on uninterpretable test results.

The two studies by Hirata et al selected patients on a retrospective basis leading to a large proportion of included colorectal cancers (>30%), and the study by Katagiri et al excluded lesions with unsuspicious vascular pattern in order to evaluate only for differentiation of low-grade and high-grade dysplastic polyps. Therefore, these 3 studies were excluded from the final pooled analysis. The 6 remaining manuscripts had a proper study design plus representative patient selection and were selected for random effects pooling.

In these 6 studies, a total of 358 neoplastic (adenomas) and 158 non-neoplastic lesions were differentiated by NBI with sensitivities, specificities and overall accuracies varying between 83-97%, 64-100% and 77-93% respectively. The bivariate random effects ‘pooled’ summary estimates of sensitivity, specificity and overall accuracy for NBI in the colon are presented in table 4. The sensitivity, specificity and overall accuracy were 92% (95%-CI: 89-94), 86% (80-91) and 89% (87-91) respectively; information on diagnostic accuracy for each individual study is represented in figure 1. Five of those studies provided figures of accuracy for chromoendoscopy as well. Corresponding sensitivity, specificity and overall accuracy of chromoendoscopy were 91% (83-96), 89% (83-93) and 91% (85-94) (table 4 and figure 1).

In addition to the abovementioned differentiation studies of colonic polyps, only two studies addressed the use of NBI for differentiation of lesions in patients with longstanding ulcerative colitis. In the study by Matsumoto et al, patients underwent white light colonoscopy for the detection of visible protruded lesions, which were subsequently inspected by NBI for mucosal classification (honeycomb appearance, villous or tortuous pattern). In addition, multiple non-suspicious flat areas were inspected with magnification NBI for mucosal classification as well. In this study, 296 colonic areas were evaluated with NBI (20 protruded lesions and 276 areas of non-suspicious flat mucosa). The sensitivity, specificity and overall accuracy of NBI for differentiating neoplasia from non-neoplastic mucosa were 80% (95%-CI: 38-96), 84.2% (80-88) and 84.1% (80-88) respectively. These rather disappointing values of accuracy may be questioned, since only one pathologist evaluated all biopsy specimens, although a second pathologist should confirm the presence of neoplasia in case of ulcerative colitis according to international guidelines. Furthermore, only 5 neoplastic areas were included in the analysis, making the sensitivity particularly unsure.

In the second study by Van den Broek et al, patients were randomized for WLE or autofluorescence imaging concerning the detection of lesions; once detected, all lesions were assessed with NBI by using the Kudo classification. Blinded histopathology (confirmed by a second expert gastrointestinal pathologist) was used as reference standard and all NBI assessments were made during ongoing endoscopy. In this study 16 neoplastic lesions and 82 non-neoplastic lesions were included with a reported sensitivity, specificity and overall accuracy for NBI of 75% (51-90), 81 (71-88) and 80% (71-86). These figures are comparable to the study by Matsumoto et al, with a wide 95%-confidence interval of the sensitivity as well.
Inter- and intraobserver agreement

Four differentiation studies reported to some extent on interobserver agreement (expressed in kappa values), but none evaluated intraobserver agreement. In only one study, an accurate description was given on the methods of measuring agreement with subsequent kappa calculation. In the study by Su et al, a kappa value of 1.0 was reached for NBI by 2 assessors who scored for brownish vascular network among 110 colorectal polyps. Although brown vascular network appears a subjective outcome measure, this led to a perfect score on interobserver agreement. Another remarkable result from this study was a kappa value of 0.981 for the assessment with conventional WLE, which was based on reddish polyp color and polyp size on the images. In the study by Chiu et al, a sub study on interobserver agreement was performed, utilizing only 10 images per diagnostic modality. The overall kappa value was 0.86, however no results were provided on each modality (including NBI) separately. In the study by Tischendorf et al 200 polyps were assessed by 2 observers, who had a perfect agreement for NBI (corresponding to a kappa of 1.0) based on mucosal as well as vascular patterns. Finally, East et al were the only authors who provided sufficient details on the methods of their interobserver study. The assessment of NBI images of 32 polyps by 2 observers led to a kappa value of 0.48 (95%-CI: 0.18-0.77) for Kudo pit pattern and 0.64 (95%-CI: 0.35-0.92) for vascular pattern intensity, which corresponds to a moderate to good agreement.

Discussion

Since the recognition of a relatively high prevalence of flat and depressed colonic lesions in western countries many efforts are being made for improved visualization of these subtle lesions, which harbor an increased risk of malignant progression. Until now, chromoendoscopy appeared exclusively to be associated with improved detection of those flat and usually small adenomas, as shown in three large randomized trials. In addition, chromoendoscopy has proven to increase the detection of neoplasia in patients with ulcerative colitis. However, methodological inadequacies in a few of these studies tend to overestimate the value of chromoendoscopy. Furthermore, since chromoendoscopy is associated with increased procedure time, higher costs and labor intensive examination, this technique has not been implemented in daily clinical practice. The use of NBI, which is also named ‘digital chromoendoscopy’, might be a more convenient and cost-effective alternative for chromoendoscopy, as NBI highlights mucosal structures without the use of dyes and visualizes vascular patterns as well.

Despite the theoretical comparability of NBI and chromoendoscopy, the results of this systematic review reveal that NBI does not definitely improve the detection of adenomas. Only one out of three large randomized studies demonstrated a significant increased detection of adenomas by NBI, although the methods of allocation raised concerns of bias in this positive study. In another randomized study, all colonoscopies were performed by only one highly experienced endoscopist who had an extraordinary high adenoma detection rate for both NBI and WLE. Moreover, the study by Adler et al revealed a significant increased adenoma detection only in the initial phase of the study, postulating a possible learning effect from the use
of NBI. Preliminary results of additional randomized studies have already been published in abstract form and hopefully will expand the already published evidence which was described in this review. Whether endoscopists with limited experience may benefit from NBI should be evaluated in studies among non-expert endoscopists in a general clinical setting.

Concerning the improved detection of neoplasia in patients with ulcerative colitis which has been demonstrated for chromoendoscopy, the only randomized study evaluating NBI for this purpose did not show an increased detection of patients with neoplasia by NBI. However, this trial was an underpowered pilot study and made use of a first generation NBI system with technical inadequacies. Future studies in ulcerative colitis should focus on new generation NBI systems with high definition or high resolution imaging and should compare NBI versus chromoendoscopy, preferably in general clinical setting. Thus far, the use of NBI for surveillance of neoplasia in ulcerative colitis cannot be recommended.

Besides the issue of improved detection, NBI also has the potential of endoscopic differentiation of neoplastic and non-neoplastic lesions by means of either mucosal pattern or vascular details. As shown by this review, NBI (using either mucosal or vascular patterns) has a comparable high sensitivity and specificity to chromoendoscopy, which has already been used for many years. With a sensitivity of 92% and specificity of 86%, NBI appears to have the potential to be used in clinical practice for differentiation of innocent hyperplastic polyps and premalignant adenomas. However, for safe clinical use the sensitivity should approach to 100%, since leaving adenomas in situ (with current sensitivity this would happen in 8% of all cases) may be harmful for patients. Future research therefore should focus on assessing the sensitivity of NBI among small versus large polyps, on defining learning curves for NBI differentiation, on interobserver variation in NBI assessments and on validation of NBI in general practice with endoscopists lacking extensive experience with this technique. Another important topic for upcoming research is the diagnostic value of NBI for differentiation of hyperplastic polyps and (sessile) serrated adenomas, which may endoscopically appear similar but have different malignant potential. The prevalence of serrated adenomas among the subgroup of patients under investigation may therefore determine whether NBI can be used for differentiation. Lastly, the use of NBI should be compared to other differentiation techniques, such as autofluorescence imaging or confocal endomicroscopy, which may further improve the diagnostic accuracy.

In conclusion, this systematic review evaluated all available evidence on the diagnostic value of NBI with respect to the detection of colonic neoplasia/adenomas as well as the differentiation between neoplastic and non-neoplastic colonic polyps. Narrow band imaging is a relatively new technique, which has extensively been studied in the last few years as it has become commercially available all over the world. Until now, NBI has failed to demonstrate an improved detection of neoplasia in the colon, and therefore its use in routine clinical practice will likely not improve the yield of neoplasia. The value of NBI for differentiating neoplastic from non-neoplastic colonic polyps has proven to be associated with high sensitivity and specificity in experienced hands. Results on differentiation with NBI seem comparable to results achieved with chromoendoscopy; however, future research should focus on defining learning curves, interobserver variation and validation in general practice.

>> For figures 2-5; see page 135-136
Reference List


31 East JE, Saunders BP. Look, remove, and discard: can narrow-band imaging replace histopathology for small colorectal polyps? It is time to push the button! Gastrointest Endosc 2007;66:953-956.


35 Rex DK, Helbig CC. High yields of small and flat adenomas with high-definition colonoscopes using either white light or narrow band imaging. Gastroenterology 2007;133:42-47.


60 Johanson JF. Practicability of high-resolution chroomoendoscopy during routine screening colonoscopy. Gastrointest Endosc 2006;63:829-830.
64 Rodriguez-Luna H, Gilani N, Ramirez FC. The role of high definition narrow band imaging endoscopy in the screening of colorectal neoplasia: initial experience with 100 consecutive patients. Gastrointest Endosc 2007;65:AB349.


69 Rubio CA, Jaramillo E. Flat serrated adenomas of the colorectal mucosa. Jpn J Cancer Res 1996;87:305-309.


Table 1: Characteristics of colon studies, subdivided into studies regarding the detection and the differentiation of abnormalities

<table>
<thead>
<tr>
<th>Detection of adenomas / neoplasia by NBI</th>
<th>Detection of adenomas / neoplasia by NBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>author, year ref study design study design</td>
<td>N inclusion criteria mean age/ male (%) number and experience of endoscopists inspection time (min.) WLE vs. NBI number of false positives WLE vs. NBI†</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>Rex 2007 35 RCT: NBI vs. WLE</td>
<td>CRC screening (n=257) - Surveillance (n=167) - Other (n=10) 434 62.5 / 231 (53%) 1 highly experienced endoscopist 7.3 vs. 7.7 -</td>
</tr>
<tr>
<td>Adler 2007 37 RCT: NBI vs. WLE</td>
<td>Surveillance (n=45) - Symptoms (n=226) - Other (n=130) 401 59.4 / 211 (53%) 7 endoscopists without previous experience with NBI 10.7 vs. 12.2 32 vs. 106</td>
</tr>
<tr>
<td>East 2007 38 Tandem design: WLE - NBI</td>
<td>HNPCC (n=62) 62 46.0 / 24 (39%) 1 experienced endoscopist (including NBI) performed most colonoscopies (&gt;90%) 6.6 vs. 7.0 33 vs. 31</td>
</tr>
<tr>
<td>Rastogi 2008 40 Tandem design: WLE - NBI</td>
<td>CRC screening (n=40) 40 62.0 / 40 (100%) 1 experienced endoscopist with unknown experience with NBI - 29 vs. 22</td>
</tr>
<tr>
<td>Inoue 2008 39 RCT: NBI vs. WLE</td>
<td>Surveillance (n=193) - Symptoms (n=50) 243 62.0 / 150 (62%) 6 endoscopists with unknown experience, of whom 1 performed &gt;60% of colonoscopies 8.5 vs. 8.8 12 vs. 24</td>
</tr>
<tr>
<td>Dekker 2007 36 RCT cross-over design: NBI vs. WLE</td>
<td>Longstanding UC (pancolitis) (n=42) 42 50.0 / 31 (74%) 3 experienced endoscopists (first evaluation of NBI) 47 vs. 50 16 vs. 43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Differentiation based on NBI patterns</th>
<th>Differentiation based on NBI patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>author, year ref study design study design</td>
<td>N index test (NBI classifications used) reference standard (histology) mean age/ male (%) number and experience of endoscopists (observer)</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Rastogi 2008 40 image evaluation</td>
<td>fine capillary network, dark dots, light rounds, tubular or gyrus-like 40 - TA (n=29) - HP (n=22) 62.0 / 40 (100%) 1 experienced observer with unknown experience with NBI</td>
</tr>
<tr>
<td>Machida 2004 41 selected patients: image evaluation</td>
<td>Kudo classification (type I-II for non-neoplastic mucosa; III-V for neoplasia) 34 - HGD (n=9) - LGD (n=25) - HP (n=9) 51.3 / 42 (54%) 2 experienced observers</td>
</tr>
<tr>
<td>Su 2006 11 consecutive image evaluation</td>
<td>brownish vascular network for predicting neoplasia 78 - CRC (n=5) - TA (n=32) - TVA (n=13) - VA (n=20) - HP (n=40)</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Chiu 2007</td>
<td>2007</td>
</tr>
<tr>
<td>East 2007</td>
<td>2007</td>
</tr>
<tr>
<td>Hirata 2007</td>
<td>2007</td>
</tr>
<tr>
<td>Hirata 2 2007</td>
<td>2007</td>
</tr>
<tr>
<td>Tischendorf 2007</td>
<td>2007</td>
</tr>
<tr>
<td>Katagiri 2008</td>
<td>2008</td>
</tr>
<tr>
<td>Matsumoto 2007*</td>
<td>2007</td>
</tr>
<tr>
<td>Van den Broek 2008*</td>
<td>2008</td>
</tr>
</tbody>
</table>

### Table 2: Methodological quality of included studies on differentiation of colonic polyps [QUADAS items]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Patient spectrum</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>unclear</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>2. Selection criteria</td>
<td>yes</td>
<td>no</td>
<td>unclear</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>3. Reference standard</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>4. Disease progression</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>5. Partial verification</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>6. Differential verification</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>7. Incorporation</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>8. Test details</td>
<td>unclear</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>9. Reference standard details</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>10. Test bias</td>
<td>yes</td>
<td>unclear</td>
<td>yes</td>
<td>unclear</td>
<td>yes</td>
<td>unclear</td>
<td>unclear</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>11. Review bias</td>
<td>yes</td>
<td>unclear</td>
<td>unclear</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>unclear</td>
<td>unclear</td>
</tr>
<tr>
<td>12. Clinical data</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>13. Uninterpretable results</td>
<td>unclear</td>
<td>no</td>
<td>no</td>
<td>unclear</td>
<td>-</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>14. Withdrawals</td>
<td>unclear</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

‘yes’ means good quality concerning the topic of interest, ‘no’ means poor quality, ‘unclear’ means unsure about quality.
Table 3: Percentage of patients with at least 1 adenoma and mean number of adenomas per examined patient for NBI versus WLE (randomized controlled trials) and adenoma miss-rates (per patient and per adenoma) for WLE when an additional inspection was done by NBI (back-to-back colonoscopy studies)

**Randomized controlled trials: NBI vs. WLE**

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>N NBI</th>
<th>N WLE</th>
<th>Patients with adenoma detected by NBI (%)</th>
<th>Patients with adenoma detected by WLE (%)</th>
<th>Odds ratio (95%-CI) of NBI vs. WLE</th>
<th>N of adenomas detected by NBI (mean per patient)</th>
<th>N of adenomas detected by WLE (mean per patient)</th>
<th>Relative ratio (95%-CI) of means for NBI vs. WLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rex, 2007</td>
<td>217</td>
<td>217</td>
<td>140 (65%)</td>
<td>145 (67%)</td>
<td>0.90 (0.61-1.34)</td>
<td>403 (1.86)</td>
<td>395 (1.82)</td>
<td>1.02 (0.89-1.17)</td>
</tr>
<tr>
<td>Adler, 2007</td>
<td>198</td>
<td>198</td>
<td>45 (23%)</td>
<td>33 (17%)</td>
<td>1.47 (0.89-2.42)</td>
<td>65 (0.33)</td>
<td>51 (0.26)</td>
<td>1.27 (0.88-1.84)</td>
</tr>
<tr>
<td>Inoue, 2008</td>
<td>122</td>
<td>121</td>
<td>51 (42%)</td>
<td>41 (34%)</td>
<td>1.40 (0.83-2.36)</td>
<td>103 (0.84)#</td>
<td>66 (0.55)#</td>
<td>1.55 (1.14-2.11)</td>
</tr>
<tr>
<td>Pooled results</td>
<td>537</td>
<td>536</td>
<td>236 (44%)</td>
<td>219 (41%)</td>
<td>1.19 (0.86-1.64)</td>
<td>571 (1.06)</td>
<td>512 (0.96)</td>
<td>1.23 (0.93-1.61)</td>
</tr>
</tbody>
</table>

**Back-to-back colonoscopy studies: WLE followed by NBI**

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>N WLE</th>
<th>N NBI</th>
<th>Patients with adenoma detected by WLE (%)</th>
<th>Added patients with adenoma by NBI (%)</th>
<th>Miss-rate of patients with adenoma</th>
<th>N of adenomas detected by WLE (mean per patient)</th>
<th>N of adenomas detected by NBI (mean per patient)</th>
<th>Adenoma miss-rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rastogi, 2008*</td>
<td>40</td>
<td>40</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>43 (1.08)</td>
<td>29 (0.73)</td>
<td>29/72 (40%)</td>
</tr>
<tr>
<td>East, 2007</td>
<td>62</td>
<td>62</td>
<td>17 (27%)</td>
<td>9 (15%)</td>
<td>17/26 (65%)†</td>
<td>25 (0.40)</td>
<td>21 (0.34)</td>
<td>21/46 (46%)</td>
</tr>
</tbody>
</table>

N number of subjects, WLE white light endoscopy (reference standard for detection), NBI narrow band imaging, *back-to-back colonoscopy studies: represented figures correspond for first inspection (reference) with WLE, † includes 2 invasive cancers, ‡ some patients with missed adenomas already had adenomas detected by WLE
Table 4: Accuracy of the index tests (classification with NBI or chromoendoscopy) for the differentiation of neoplastic (adenoma) versus non-neoplastic colonic polyps: summary estimates of sensitivity, specificity and overall accuracy from a bivariate random effects model

<table>
<thead>
<tr>
<th>Test</th>
<th>N of studies</th>
<th>N neoplastic lesions</th>
<th>N non-neoplastic lesions</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Overall accuracy (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBI *</td>
<td>6</td>
<td>358</td>
<td>158</td>
<td>92% (89% to 94%)</td>
<td>86% (80% to 91%)</td>
<td>89% (87% to 91%)</td>
</tr>
<tr>
<td>Chromoendoscopy</td>
<td>5</td>
<td>326</td>
<td>139</td>
<td>91% (83% to 96%)</td>
<td>89% (83% to 93%)</td>
<td>91% (85 to 94%)</td>
</tr>
</tbody>
</table>

* pooled data of all classification systems with NBI (e.g. Kudo classification, vascular pattern intensity etc.)
Figure 1: Forrest plots of sensitivities and specificities concerning the differentiation (adenoma versus non-neoplastic polyp) by NBI and corresponding chromoendoscopy per individual colon study as well as the pooled results (obtained with bivariate random effects method).