Published in final edited form as:

Gastroenterology. 2013 January ; 144(1): 81–91. doi:10.1053/j.gastro.2012.09.054.

Real-Time Optical Biopsy of Colon Polyps With Narrow Band Imaging in Community Practice Does Not Yet Meet Key Thresholds for Clinical Decisions

URI LADABAUM^{1,2}, ANN FIORITTO³, AYA MITANI^{2,4}, MANISHA DESAI^{2,4}, JANE P. KIM^{2,4}, DOUGLAS K. REX⁵, THOMAS IMPERIALE⁵, and NARESH GUNARATNAM³

¹Division of Gastroenterology and Hepatology, Stanford University School of Medicine, Stanford, California

²Department of Medicine, Stanford University School of Medicine, Stanford, California

³Huron Gastroenterology Associates, Ann Arbor, Michigan

⁴Quantitative Sciences Unit, Stanford University School of Medicine, Stanford, California

⁵Division of Gastroenterology, Indiana University, Indianapolis, Indiana

Abstract

BACKGROUND & AIMS—Accurate optical analysis of colorectal polyps (optical biopsy) could prevent unnecessary polypectomies or allow a “resect and discard” strategy with surveillance intervals determined based on the results of the optical biopsy; this could be less expensive than histopathologic analysis of polyps. We prospectively evaluated real-time optical biopsy analysis of polyps with narrow band imaging (NBI) by community-based gastroenterologists.

METHODS—We first analyzed a computerized module to train gastroenterologists (N = 13) in optical biopsy skills using photographs of polyps. Then we evaluated a practice-based learning program for these gastroenterologists (n = 12) that included real-time optical analysis of polyps in vivo, comparison of optical biopsy predictions to histopathologic analysis, and ongoing feedback on performance.

RESULTS—Twelve of 13 subjects identified adenomas with >90% accuracy at the end of the computer study, and 3 of 12 subjects did so with accuracy 90% in the in vivo study. Learning curves showed considerable variation among batches of polyps. For diminutive rectosigmoid polyps assessed with high confidence at the end of the study, adenomas were identified with mean (95% confidence interval [CI]) accuracy, sensitivity, specificity, and negative predictive values of 81% (73%–89%), 85% (74%–96%), 78% (66%–92%), and 91% (86%–97%), respectively. The

Address requests for reprints to: Uri Ladabaum, MD, MS, Division of Gastroenterology and Hepatology, Stanford University School of Medicine, 300 Pasteur Drive, Always Building, Room M211, Stanford, California 94305-5187. uri.ladabaum@stanford.edu; fax: (650) 723-5488.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at <http://dx.doi.org/10.1053/j.gastro.2012.09.054>.

Conflicts of interest

The authors disclose the following: D. K. Rex has received research support and serves on the speaker's bureau for Olympus Corp. The remaining authors disclose no conflicts.

adjusted odds ratio for high confidence as a predictor of accuracy was 1.8 (95% CI, 1.3–2.5). The agreement between surveillance recommendations informed by high-confidence NBI analysis of diminutive polyps and results from histopathologic analysis of all polyps was 80% (95% CI, 77%–82%).

CONCLUSIONS—In an evaluation of real-time optical biopsy analysis of polyps with NBI, only 25% of gastroenterologists assessed polyps with 90% accuracy. The negative predictive value for identification of adenomas, but not the surveillance interval agreement, met the American Society for Gastrointestinal Endoscopy–recommended thresholds for optical biopsy. Better results in community practice must be achieved before NBI-based optical biopsy methods can be used routinely to evaluate polyps; ClinicalTrials.gov number, NCT01638091.

Keywords

Optical Diagnosis; Endoscopic Diagnosis; Polyp Histology

Accurate endoscopic determination of the histology of colorectal polyps could prevent unnecessary polypectomies or allow a strategy in which all diminutive polyps are resected but “optical biopsy” informs surveillance recommendations, which could substantially decrease the costs related to histopathologic assessment of polyps. The American Society for Gastrointestinal Endoscopy (ASGE) Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI) statement on this topic proposes that a “resect and discard” strategy for diminutive colorectal polyps should yield 90% agreement with surveillance interval recommendations when compared with decisions based on pathologic assessment of all polyps and that, in order to not resect suspected diminutive rectosigmoid hyperplastic polyps, there should be 90% negative predictive value for adenomatous histology.¹ These thresholds reflect expert opinion informed by the reported degree of agreement on polyp histology between pathologists.¹

Narrow band imaging (NBI) uses narrow band light filters to highlight mucosal architecture and vasculature. An international expert group has proposed the NBI International Colorectal Endoscopic (NICE) classification to distinguish adenomas from hyperplastic polyps.² Experts have achieved very high performance levels in optical diagnosis of polyp histology with NBI.^{3–5} Several studies suggest that nonexperts can learn optical diagnosis with NBI ex vivo.^{6–9} It remains to be shown whether the high levels of performance achieved by experts can be replicated in vivo in routine clinical practice by community-based gastroenterologists.

We performed a prospective evaluation of real-time optical biopsy of colorectal polyps with NBI by community-based gastroenterologists. An initial ex vivo study phase tested the impact of a computerized self-learning module on participants’ optical biopsy skills based on photographs. A subsequent in vivo study phase evaluated prospectively a practice-based learning program. The primary outcome of the in vivo study phase was the proportion of participants achieving 90% accuracy in differentiating independent diminutive (< 5 mm) adenomas from nonadenomas. Secondary outcomes included the nature of learning curves, test performance characteristics, and agreement between surveillance recommendations with versus without the use of NBI.

Subjects and Methods

Study Site, Participants, Endoscopic Equipment, and General Study Design

Gastroenterologists from a single-specialty practice in Ann Arbor, Michigan, participated in the study. A research coordinator in Ann Arbor entered data into a REDCap (Vanderbilt University) database designed and administered at a data coordination and analysis center at Stanford University (Stanford, CA).

The institutional review boards of St Joseph Mercy Hospital in Ann Arbor, Michigan, and Stanford University approved the study. All authors had access to the study data and reviewed and approved the final manuscript. Participants provided written informed consent. Both institutional review boards agreed that informed consent for research was not required from patients, because their care was not affected and the research subjects were the endoscopists. No participant had formal training or significant experience with NBI. Participants' annual colonoscopy volume and adenoma detection rate were measured in the year before study entry.

Endoscopy suites were equipped with Evis Exera II CV-180 processors, CF-H180AL and PCF-H180AL colonoscopes (Olympus America, Center Valley, PA), high-definition monitors, and posters showing the NICE classification and photo examples (Appendix). Three community-based, fellowship-trained gastrointestinal pathologists interpreted all specimens as part of routine practice and were blinded to optical diagnosis.

Our aim was to design a practical learning program that could be instituted in a busy practice. The study took place from March 2011 to March 2012.

Ex Vivo Phase of the Study

The ex vivo study phase consisted of 3 self-administered, computerized components that participants completed at their own pace during the first study week: a pretest, a learning module on the NICE classification, and a posttest (Rex et al., Olympus America). The pretest and posttest consisted of 2 different sets of 25 endoscopic NBI images of adenomatous or hyperplastic polyps, with possible answers of "adenoma," "hyperplastic," and "I don't know."

The primary ex vivo outcome was change in score (percent correct) between tests, with potential predictors defined a priori as years of practice experience, annual colonoscopy volume, and adenoma detection rate.

In Vivo Phase of the Study

The in vivo study phase consisted of a practice-based learning program that included real-time optical diagnosis with NBI, comparison to final pathology diagnosis, and confidential feedback on individual performance every 1 to 2 weeks. Participants were asked to include in the study any colonoscopy in which at least one polyp was removed, including nonscreening examinations.

Participants completed a clinical research form for each study colonoscopy. For each study polyp, participants recorded the location, size estimated by comparison to open biopsy forceps or snare catheter, and morphology; confirmation of photo acquisition under white light and NBI; NICE classification features observed and predicted histology using NBI (hyperplastic or adenoma; or other, with explanation); and the level of confidence (“high” if polyps had one or more features associated with one histology and no features associated with the other; “low” if there was uncertainty regarding features or if there were features of both histologies).⁴ Participants were encouraged to call out these items in real time to a nurse and to check the form for accuracy after the procedure. Each study polyp was resected and then sent in an individual jar for pathologic analysis. Research forms were filed with a procedure report copy with photos for later reference.

When pathology reports became available, participants were required to enter the actual pathologic diagnosis for each study polyp on the study colonoscopy research form. Participants were encouraged to compare the pathology and their optical diagnosis and to study the images in a process of ongoing practice-based learning. The study coordinator entered all data from fully completed forms into the REDCap database once a week.

The data coordinating center analyzed each participant’s sensitivity, specificity, and accuracy for adenoma by polyp size on an ongoing basis. Polyps ≤ 5 mm were defined as diminutive, 6 to 9 mm as small, and ≥ 10 mm as large. Reports of cumulative test performance and test performance in the last batch of 10 independent polyps were provided to each participant in a confidential fashion every 1 to 2 weeks.

The primary analysis was based on diminutive-sized independent polyps. One diminutive polyp was selected randomly within each study colonoscopy to avoid possible correlation in outcome among polyps from the same patient.

Success of In Vivo Training and Early Completion

The in vivo phase had a 3-stage design with 2 opportunities for early training completion, which were governed by 2 prespecified decision rules. Success for a participant was defined as achieving $\geq 90\%$ accuracy in optical diagnosis of diminutive polyps. This was based on the last 30 consecutive independent diminutive polyps per participant at one of the 3 prespecified points.

Based on expert opinion that assessment of 250 polyps might be needed to achieve proficiency, we estimated that success with training could be assessed after predictions had been made for 90 independent diminutive polyps (ie, polyps 61–90), at which point many additional diminutive, small, and large polyps would also have been evaluated. The first decision rule was applied after assessment of 50 independent diminutive polyps and the second after assessment of 70 independent diminutive polyps (ie, polyps 21–50 and 41–70, respectively). Participants were encouraged to remain in the study even if their training was deemed a success to provide information on sustaining performance levels. The final evaluation was performed after assessment of 90 independent diminutive polyps. We performed separate analyses in which all diminutive polyps were considered.

Statistical Analysis

Descriptive statistics were computed for endoscopists' characteristics and for the in vivo study colonoscopies and polyps.

In the ex vivo phase, the change in score from pretest to posttest was analyzed using a 2-sided paired *t* test. The change in score was analyzed first by considering "I don't know" a wrong answer and second by excluding all "I don't know" answers. Potential predictors of change in score were explored by the 2-sided Fisher exact test after dichotomizing the predictors and the change in score (absolute 5% vs >5% improvement).

In the in vivo phase, test performance characteristics were based on the diagnoses of "adenoma" (including tubular, tubulovillous, and villous adenomas) or "not adenoma." Sessile serrated adenomas and traditional serrated adenomas were considered "not adenoma" based on preliminary reports of NBI appearance.^{10,11} The 14 polyps with missing size were excluded.

Accuracy, sensitivity, specificity, positive predictive value, and negative predictive value for each endoscopist were computed. Group means and 95% confidence intervals were calculated by polyp size and by other prespecified features of interest, including timing of assessment (eg, first vs last batches of 20 diminutive polyps to explore a learning effect), location (rectosigmoid or proximal to rectosigmoid), confidence level (low vs high), and the last batches of 20 diminutive rectosigmoid polyps per endoscopist assessed with high confidence. Formal comparisons between features (eg, timing of assessment) on change in outcome (eg, change in accuracy) were made using 2-sided paired *t* tests.

Learning curves for accuracy, sensitivity, and specificity were constructed for individuals and the group mean overall as a function of nonoverlapping consecutive batches of 20 polyps per endoscopist for diminutive polyps, diminutive and small polyps combined, and high confidence assessments or all assessments.

The associations between accuracy of optical diagnosis of diminutive polyps and timing of assessment, confidence level, polyp location, endoscopist's characteristics or ex vivo performance, and an indicator variable for true adenoma by histopathology were explored using generalized estimating equations models to account for the correlation expected among outcomes for an endoscopist. We assumed an exchangeable correlation structure and made use of robust standard errors when drawing inference. First, unadjusted associations between accuracy and each potential predictor were estimated. Next, to select a subset of features that jointly predicted accuracy, we used the following model selection approach. Unadjusted predictors of statistical significance at the 0.10 level were included in a multivariable model. Variables with corresponding significance levels of 0.05 were retained. Those that did not meet the significance level of 0.10 in the initial step were reconsidered in the presence of other variables. Finally, every 2-way interaction was evaluated for all covariates in the final multivariable model.

We calculated the percent agreement and κ statistic between surveillance recommendations that would follow from the use of NBI optical diagnosis for diminutive polyps combined

with pathologic assessment of all other polyps versus pathologic assessment of all polyps. Analyses were performed with inclusion of all diminutive polyps or only those assessed with high confidence; all study colonoscopies, only those with at least one high-confidence diminutive polyp, or only the last 20 study colonoscopies per endoscopist with at least one high-confidence diminutive polyp; and based on the Multi-Society Task Force surveillance intervals¹² or a modified scenario in which surveillance for 1 to 2 adenomas was at 10 years instead of 5 to 10 years.¹³ We determined how tubulovillous and villous adenomas were characterized with NBI and compared the surveillance intervals assigned with and without use of NBI. We also considered differences in surveillance intervals when the presence of diminutive sessile serrated adenomas and traditional serrated adenomas informed surveillance intervals.¹⁴

All statistical analyses were performed using the R package (<http://cran.r-project.org/>) and SAS 9.3 (SAS Institute, Cary, NC).

Sample Size Considerations

As a group, 4 academic endoscopists reached 80% accuracy with NBI after 131 examinations with 265 polyps.¹⁵ Assuming a participant achieved 60% to 70% accuracy at the first or second rules for early completion, the probability of prematurely considering this successful training was 0.01. Assuming a participant achieved 80% accuracy at the first or second rules for early completion, the probability of prematurely considering this successful training was 0.11. Our design had 88% power to correctly conclude that a participant achieved success if accuracy increased from 0.7 to 0.9 from the first to second rule for early completion. We calculated a priori that with 12 participants, our study design provided 79% power to detect an 80% success rate, based on a one-sided exact binomial test with an 8% type I error rate.

Results

Participant Demographics, Study Colonoscopies, and Polyps

Fourteen participants enrolled and completed the ex vivo study phase, and 12 of them entered the in vivo study phase (Table 1). The in vivo study phase included a total of 1673 study colonoscopies and 2596 study polyps (1858 diminutive, 547 small, 177 large, 14 size missing), with adenomas accounting for 62% of diminutive, 72% of small, and 78% of large study polyps (Table 1).

Ex Vivo Pretest and Posttest

Posttest results were missing for 1 subject, and pretest results were missing for 2 subjects. Test scores improved for 10 of the 11 subjects who submitted complete pretest and posttest results data, and 12 of 13 subjects who submitted posttest results scored >90% accuracy on the posttest (Figure 1). When “I don’t know” was considered an incorrect answer, the mean change in score was 16 (SD, 15) ($P = .006$); when items answered “I don’t know” were excluded, the mean change in score was 10 (SD, 8) ($P = .002$) (Table 2).

In Vivo Practice-Based Learning Program Success Rate

Three of the 12 participants (25%) who entered the in vivo study phase achieved accuracy 90% at one of the 3 prespecified points. Of the 3 participants who achieved 90% accuracy, one achieved it at the first and sustained it through the second and third prespecified evaluation points, and 2 achieved it at the second but not the third point.

Ten participants assessed at least 21 independent diminutive polyps. Eight participants assessed at least 50 independent diminutive polyps, and one achieved 90% accuracy at the first prespecified point. These 8 participants all assessed at least 20 more independent diminutive polyps, and 2 additional participants achieved 90% and 93% accuracies, respectively, at the second prespecified point. Seven participants, including the 3 who had already achieved 90% accuracy, assessed at least 90 independent diminutive polyps, but no additional participants achieved 90% accuracy at this point.

When all, not only independent, diminutive polyps were considered, 4 participants achieved 90% accuracy, 2 each at the second and third prespecified points.

In Vivo Learning Curves

Test performance characteristics over time as a function of batches of 20 consecutive polyps showed considerable batch-to-batch variation (Figure 2 for high confidence assessments, and Figure 3 for all assessments). This variation was more pronounced for specificity compared with sensitivity (Figures 2 and 3). There was no clear pattern of early learning with later stabilization of performance at a higher level.

In Vivo Test Performance Characteristics

Test performance characteristics overall and for the prespecified polyp subgroups are shown in Table 3. For diminutive polyps compared with small polyps overall, sensitivity was lower ($P = .019$) but specificity was higher ($P = .001$) and accuracy was not significantly different ($P = .600$) (Table 3). For diminutive polyps, sensitivity improved between the first and last batches of polyps per endoscopist ($P = .044$), whereas specificity decreased ($P = .005$) (Table 3). High confidence was associated with improved test performance for diminutive polyps ($P = .1$ for accuracy, $P = .049$ for sensitivity, and $P = .017$ for specificity) (Table 3). At the level of the endoscopist, the overall fraction of diminutive polyps assessed with high confidence was not associated with accuracy ($P = .95$) or sensitivity ($P = .54$).

Predictors of Accuracy

In unadjusted analyses, high confidence assessment with NBI and a polyp being a true adenoma were the only predictors associated with accuracy (both $P < .001$). The last 20 colonoscopies per endoscopist showed a marginal positive association with increased accuracy ($P = .086$). Polyp location and endoscopist's years in practice, colonoscopy volume, adenoma detection rate, ex vivo pretest score, and change in score were not associated with accuracy (all P values $> .49$).

Confidence level, location, and a polyp being a true adenoma jointly predicted accuracy. In the final multivariable model, the odds ratio for accurate assessment was 1.8 (95%

confidence interval [CI], 1.3–2.5; $P < .001$) for high versus low confidence, and the odds ratios for being a true adenoma varied by location ($P < .001$): 7.5 (95% CI, 3.6–15.3) proximal to the rectosigmoid and 1.6 (95% CI, 0.7–3.3) in the rectosigmoid.

PIVI Thresholds: In Vivo Negative Predictive Value and Surveillance Intervals

In the last batches of 20 diminutive rectosigmoid polyps per endoscopist assessed with high confidence, the negative predictive value for adenoma was 91% (95% CI, 86%–97%). Six subjects achieved 90% negative predictive value for all diminutive rectosigmoid polyps assessed with high confidence.

The agreement between surveillance recommendations based on high-confidence NBI optical diagnosis for diminutive polyps combined with pathologic assessment of all other polyps versus pathologic assessment of all polyps is shown in Table 4. For colonoscopies with at least one high-confidence diminutive polyp, the agreement based on traditional surveillance intervals was 80% (95% CI, 77%–82%), with NBI use leading to 136 (13%) shorter and 78 (7%) longer recommended intervals than with histopathology alone; the agreement based on modified intervals (10 years instead of 5–10 years for 1–2 diminutive or small adenomas) was 97% (95% CI, 96%–98%), with NBI use leading to 24 (2%) shorter and 10 (1%) longer recommended intervals than with histopathology alone. When the presence of diminutive sessile serrated adenomas and traditional serrated adenomas informed surveillance intervals, the agreement between strategies was only minimally affected: 79% (95% CI, 77%–82%) with traditional intervals and 96% (95% CI, 95%–97%) with modified intervals.

For colonoscopies with at least one high-confidence diminutive polyp, based on traditional surveillance intervals, 3 subjects achieved 90% agreement and 2 met both this and the negative predictive value PIVI thresholds. Based on modified intervals, 10 subjects achieved 90% agreement and 5 met both PIVI thresholds.

There were 29 polyps (9 diminutive, 6 small, 14 large) diagnosed at pathology as tubulovillous or villous adenomas, and none had high-grade dysplasia. All were classified as adenomas under NBI. Among the 8 patients with 9 diminutive tubulovillous or villous adenomas, the recommended surveillance intervals based on NBI optical diagnosis for diminutive polyps combined with pathologic assessment of all other polyps were 3 years in 3 patients and 5 years in 5 patients, compared with 3 years in all 8 patients based on histopathology.

Discussion

We assessed whether the high performance levels in optical diagnosis of polyp histology by experts using NBI can be replicated in real-time practice by community gastroenterologists using commercially available equipment. Ex vivo, after completion of a computerized self-learning module, 12 of 13 participants scored >90% correct on a photograph-based posttest. In vivo, 3 of 12 participants achieved accuracy 90% in assessing diminutive polyp histology. Learning curves showed considerable variability as a function of consecutive polyp batches, with some improvement over time in sensitivity but at the expense of a

decrease in specificity. Multivariable modeling confirmed level of confidence as a predictor of accuracy. The negative predictive value for adenoma in the last batches of 20 diminutive rectosigmoid polyps assessed with high confidence, but not in less selected polyp subgroups, reached the minimum 90% threshold recommended by the recent ASGE PIVI statement.¹ Agreement in surveillance intervals with and without the use of high-confidence NBI assessments for diminutive polyps fell below the 90% threshold recommended by the recent ASGE PIVI statement.¹ The results achieved in our study were not as good as the best results of experts,^{3–5} but they were in the range of previous results from academic centers.^{15,16}

Several groups have investigated ex vivo training on NBI features of polyps. Rastogi et al¹⁷ reported accuracies of 80% to 86% in the first and 85% to 91% in the second reading of 65 images performed 2 months after training.¹⁸ In a study of trimodal imaging, the accuracies achieved ex vivo with NBI by 3 experienced and 4 nonexperienced endoscopists from a university hospital and 6 nonexperienced endoscopists from a nonuniversity hospital were 70%, 60%, and 70%, respectively, after a short training session.⁹ Higashi et al reported improvements in accuracy on the same set of NBI images before and after a 1-hour training session from 53% to 66% by 4 nonendoscopists and from 65% to 88% by 4 endoscopists without NBI experience compared with 81% accuracy for 4 endoscopists highly experienced with NBI.⁶ A short didactic session by an expert in NBI led to improvements in ex vivo test scores from 38% to 87% for 12 residents, from 60% to 94% for 12 gastroenterology fellows, and from 45% to 91% for 13 gastroenterology faculty.⁸ A computerized training module led to improvements in ex vivo accuracy from 62% to 84% for 7 novices, from 75% to 90% for 7 trainees, and from 68% to 84% for 7 experienced endoscopists without NBI experience.⁷ Our results extend previous observations by showing that high accuracy can be achieved ex vivo by community gastroenterologists after self-paced training with a computerized module.

Little is known about the learning curves for in vivo optical biopsy of polyps. Four experienced endoscopists who were trained for 1 hour on modified Kudo pit patterns and vascular color intensity attained a pooled in vivo accuracy of 74% on an initial 133 polyps, with improvement to 87% on a subsequent 132 polyps.¹⁵ East et al did not find significant differences in test performance characteristics for 3 endoscopists already experienced with NBI during the first and second parts of a study using high-magnification NBI.¹⁹ In a study of a single endoscopist experienced with NBI, numerical but not statistically significant improvements in test performance characteristics were found between the first and second half of a 100-patient study, but initial performance was already good.³ In this study, we constructed detailed in vivo learning curves for 12 community-based gastroenterologists without prior experience with NBI and found considerable variation between batches of polyps, without a clear pattern of early learning or later stabilization of performance at a higher level. It is not known if participants manifested their ultimate level of performance early on or whether more training could improve performance. We can only speculate whether the improvement over time in sensitivity at the expense of specificity, as well as greater specificity for diminutive than small polyps, reflect an inclination to minimize the possibility of missing an adenoma.

A range of performance levels has been reported for post-hoc differentiation of adenoma versus nonadenoma with NBI. In a systematic review from 2009, the pooled mean test performance estimates (and 95% CIs) based on 6 studies of post-hoc image evaluation were 92% (89%–94%) for sensitivity, 86% (80%–91%) for specificity, and 89% (87%–91%) for accuracy.²⁰ Additional studies have reported similar post-hoc results,^{21–23} but lower performance levels have also been reported for both NBI expert and nonexpert endoscopists.²⁴

In vivo, using NBI without magnification in real time, Rex achieved sensitivity of 96%, specificity of 92%, and accuracy of 94% for high-confidence diagnoses,⁴ and Rastogi et al achieved sensitivity of 99%, specificity of 90%, and accuracy of 96%.³ Sano achieved similarly impressive results with magnifying NBI.⁵ The 3 endoscopists experienced with NBI in the study by East et al achieved sensitivity of 88%, specificity of 91%, and accuracy of 90% based on NBI pit pattern and sensitivity of 94%, specificity of 89%, and accuracy of 91% based on vascular pattern intensity.¹⁹ However, not all studies from expert centers have reported 90% sensitivity and specificity with NBI.^{16,25–27} The participants in the study by Rogart et al achieved sensitivity of 80%, specificity of 81%, and accuracy of 80%.¹⁵ In the DISCARD trial, using the Lucera system (Olympus, Tokyo, Japan), the 2 experienced endoscopists achieved sensitivity of 96%, specificity of 90%, and accuracy of 95% compared with sensitivity of 88%, specificity of 85%, and accuracy of 87% for the 2 inexperienced participants.¹¹ In a study of 8 endoscopists from 6 nonacademic centers in The Netherlands, NBI displayed a sensitivity of 87%, specificity of 63%, and accuracy of 75%.²⁸

In our study, the test performance characteristics achieved were not as good as those in the studies by experts reporting the best in vivo results, but our results are comparable to those of Rogart et al¹⁵ and Rastogi et al¹⁶ at academic centers and the Dutch study of nonacademic centers.²⁸ The levels of performance achieved in vivo toward the end of our study and those for high-confidence assessments were comparable to previously reported post-hoc results.^{20,21,23,24} It is not known whether a community versus an academic setting, the role of endoscopists as study subjects versus authors, or the self-guided nature of our training could explain some of the differences between our results and those of some previous studies.

In our study, the agreement in surveillance intervals with the use of high-confidence NBI assessments for diminutive polyps versus pathologic assessment of all polyps fell below the recommended 90% threshold.¹ The negative predictive value for adenoma in the last batches of 20 diminutive rectosigmoid polyps assessed with high confidence reached the recommended minimum 90% threshold.¹ If surveillance after removal of 1 to 2 diminutive adenomas was at 10 years instead of 5 to 10 years,¹³ then the agreement in surveillance intervals in our study exceeded the recommended 90% threshold.

Our study provides lessons for the potential implementation of practice-based learning programs. There was greater participation in the ex vivo than in the in vivo phase, and not all participants remained engaged in the in vivo phase. Incentives may need to be developed for busy clinicians to learn and use optical biopsy techniques. We found evidence of operator

dependence. It remains to be determined how much of an endoscopist's ultimate proficiency in NBI optical diagnosis is determined by technical issues, dedication, motivation, or innate skill.

The strengths of our study include its setting, size, design, and scope. Most previous studies of NBI optical biopsy were from expert and/or academic centers, and many performed ex vivo assessments. Our study setting reflects real-world practice in the United States. We enrolled 14 endoscopists, and the number of study colonoscopies and polyps were among the largest in the literature on NBI. Our design included both an ex vivo and an in vivo phase designed to be applicable in general practice.

Our study has limitations. We did not test different training program designs. Although participants had to review the final pathology diagnoses and they received continuous feedback, we could not ensure self-critical review. The goal was to include every eligible consecutive colonoscopy, but we did not set out to enforce this given our objective to make the protocol acceptable in clinical practice. Not all participants reached the third prespecified assessment point in the in vivo phase. The in vivo learning phase did not include immediate feedback, but relied instead on review of pathology reports days after a given procedure, and there was no formal supplementary training for participants who did not achieve success at the prespecified points. In addition, there was no opportunity for live feedback by experts, but we chose to study self-learning because of the potential for broad applicability.

The United States may be entering an era of bundled payments for services, with an emphasis on quality and outcomes. It has been estimated that a "resect and discard" strategy for diminutive polyps detected by screening colonoscopy could result in annual savings of \$33 million in the United States.²⁹ For such a strategy to be accepted, it will be important to show that individual practitioners can implement it with high levels of performance. To justify leaving polyps in place, the negative predictive value must be very high; the minimum 90% threshold recommended by the ASGE PIVI statement was achieved in the last batches of 20 diminutive rectosigmoid polyps assessed with high confidence in this study, but not in less selected polyp subgroups. Given the erratic nature of the observed learning curves, it may be necessary to perform individualized assessments and periodic reassessments of endoscopists' skills.

In conclusion, our results suggest that the prospects of applying NBI-based optical diagnosis in community practice are promising. However, the levels of proficiency achieved in this study were not as high as the best results published by experts. It remains to be explored whether refinements in training program design, including video clips for training or use of sharp frozen images, endoscopic technology, optical magnification, or the use of novel tools such as computer-aided diagnosis can improve the performance of NBI-based optical diagnosis in routine practice.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

Supported by a grant from the Division of Gastroenterology at Stanford University School of Medicine.

Abbreviations used in this paper

ASGE	American Society for Gastrointestinal Endoscopy
CI	confidence interval
NBI	narrow band imaging
NICE	NBI International Colorectal Endoscopic (classification)
PIVI	Preservation and Incorporation of Valuable Endoscopic Innovations

References

1. Rex DK, Kahi C, O'Brien M, et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointest Endosc.* 2011; 73:419–422. [PubMed: 21353837]
2. Oba S, Tanaka S, Sano Y, et al. Current status of narrow-band imaging magnifying colonoscopy for colorectal neoplasia in Japan. *Digestion.* 2011; 83:167–172. [PubMed: 21266811]
3. Rastogi A, Keighley J, Singh V, et al. High accuracy of narrow band imaging without magnification for the real-time characterization of polyp histology and its comparison with high-definition white light colonoscopy: a prospective study. *Am J Gastroenterol.* 2009; 104:2422–2430. [PubMed: 19584829]
4. Rex DK. Narrow-band imaging without optical magnification for histologic analysis of colorectal polyps. *Gastroenterology.* 2009; 136:1174–1181. [PubMed: 19187781]
5. Sano Y, Ikematsu H, Fu KI, et al. Meshed capillary vessels by use of narrow-band imaging for differential diagnosis of small colorectal polyps. *Gastrointest Endosc.* 2009; 69:278–283. [PubMed: 18951131]
6. Higashi R, Uraoka T, Kato J, et al. Diagnostic accuracy of narrow-band imaging and pit pattern analysis significantly improved for less-experienced endoscopists after an expanded training program. *Gastrointest Endosc.* 2010; 72:127–135. [PubMed: 20493482]
7. Ignjatovic A, Thomas-Gibson S, East JE, et al. Development and validation of a training module on the use of narrow-band imaging in differentiation of small adenomas from hyperplastic colorectal polyps. *Gastrointest Endosc.* 2011; 73:128–133. [PubMed: 21184878]
8. Raghavendra M, Hewett DG, Rex DK. Differentiating adenomas from hyperplastic colorectal polyps: narrow-band imaging can be learned in 20 minutes. *Gastrointest Endosc.* 2010; 72:572–576. [PubMed: 20561618]
9. van den Broek FJ, van Soest EJ, Naber AH, et al. Combining autofluorescence imaging and narrow-band imaging for the differentiation of adenomas from non-neoplastic colonic polyps among experienced and non-experienced endoscopists. *Am J Gastroenterol.* 2009; 104:1498–1507. [PubMed: 19491863]
10. Boparai KS, van den Broek FJ, van Eeden S, et al. Hyperplastic polyposis syndrome: a pilot study for the differentiation of polyps by using high-resolution endoscopy, autofluorescence imaging, and narrow-band imaging. *Gastrointest Endosc.* 2009; 70:947–955. [PubMed: 19595313]
11. Ignjatovic A, East JE, Suzuki N, et al. Optical diagnosis of small colorectal polyps at routine colonoscopy (Detect InSpect ChAracterise Resect and Discard; DISCARD trial): a prospective cohort study. *Lancet Oncol.* 2009; 10:1171–1178. [PubMed: 19910250]

12. Levin B, Lieberman DA, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology*. 2008; 134:1570–1595. [PubMed: 18384785]
13. Gupta N, Bansal A, Rao D, et al. Accuracy of in vivo optical diagnosis of colon polyp histology by narrow-band imaging in predicting colonoscopy surveillance intervals. *Gastrointest Endosc*. 2012; 75:494–502. [PubMed: 22032847]
14. Lieberman DA, Rex DK, Winawer SJ, et al. Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology*. 2012; 143:844–857. [PubMed: 22763141]
15. Rogart JN, Jain D, Siddiqui UD, et al. Narrow-band imaging without high magnification to differentiate polyps during real-time colonoscopy: improvement with experience. *Gastrointest Endosc*. 2008; 68:1136–1145. [PubMed: 18691708]
16. Rastogi A, Early DS, Gupta N, et al. Randomized, controlled trial of standard-definition white-light, high-definition white-light, and narrow-band imaging colonoscopy for the detection of colon polyps and prediction of polyp histology. *Gastrointest Endosc*. 2011; 74:593–602. [PubMed: 21802078]
17. Rastogi A, Pondugula K, Bansal A, et al. Recognition of surface mucosal and vascular patterns of colon polyps by using narrow-band imaging: interobserver and intraobserver agreement and prediction of polyp histology. *Gastrointest Endosc*. 2009; 69:716–722. [PubMed: 19251016]
18. Rastogi A, Bansal A, Wani S, et al. Narrow-band imaging colonoscopy—a pilot feasibility study for the detection of polyps and correlation of surface patterns with polyp histologic diagnosis. *Gastrointest Endosc*. 2008; 67:280–286. [PubMed: 18155210]
19. East JE, Suzuki N, Bassett P, et al. Narrow band imaging with magnification for the characterization of small and diminutive colonic polyps: pit pattern and vascular pattern intensity. *Endoscopy*. 2008; 40:811–817. [PubMed: 18828077]
20. van den Broek FJ, Reitsma JB, Curvers WL, et al. Systematic review of narrow-band imaging for the detection and differentiation of neoplastic and nonneoplastic lesions in the colon (with videos). *Gastrointest Endosc*. 2009; 69:124–135. [PubMed: 19111693]
21. Henry ZH, Yeaton P, Shami VM, et al. Meshed capillary vessels found on narrow-band imaging without optical magnification effectively identifies colorectal neoplasia: a North American validation of the Japanese experience. *Gastrointest Endosc*. 2010; 72:118–126. [PubMed: 20381799]
22. Sikka S, Ringold DA, Jonnalagadda S, et al. Comparison of white light and narrow band high definition images in predicting colon polyp histology, using standard colonoscopes without optical magnification. *Endoscopy*. 2008; 40:818–822. [PubMed: 18668472]
23. Tischendorf JJ, Schirin-Sokhan R, Streetz K, et al. Value of magnifying endoscopy in classifying colorectal polyps based on vascular pattern. *Endoscopy*. 2010; 42:22–27. [PubMed: 19899031]
24. Ignjatovic A, East JE, Guenther T, et al. What is the most reliable imaging modality for small colonic polyp characterization? Study of white-light, autofluorescence, and narrow-band imaging. *Endoscopy*. 2011; 43:94–99. [PubMed: 21271465]
25. Buchner AM, Shahid MW, Heckman MG, et al. Comparison of probe-based confocal laser endomicroscopy with virtual chromoendoscopy for classification of colon polyps. *Gastroenterology*. 2010; 138:834–842. [PubMed: 19909747]
26. Lee CK, Lee SH, Hwangbo Y. Narrow-band imaging versus I-Scan for the real-time histological prediction of diminutive colonic polyps: a prospective comparative study by using the simple unified endoscopic classification. *Gastrointest Endosc*. 2011; 74:603–609. [PubMed: 21762907]
27. van den Broek FJ, Fockens P, Van Eeden S, et al. Clinical evaluation of endoscopic trimodal imaging for the detection and differentiation of colonic polyps. *Clin Gastroenterol Hepatol*. 2009; 7:288–295. [PubMed: 19168154]
28. Kuiper T, van den Broek FJ, Naber AH, et al. Endoscopic trimodal imaging detects colonic neoplasia as well as standard video endoscopy. *Gastroenterology*. 2011; 140:1887–1894. [PubMed: 21419769]

29. Hassan C, Pickhardt PJ, Rex DK. A resect and discard strategy would improve cost-effectiveness of colorectal cancer screening. *Clin Gastroenterol Hepatol*. 2010; 8:865–869. 869 e1–3. [PubMed: 20621680]

Figure 1.

Ex vivo test scores before and after completion of a computerized learning module. Scores improved for 10 of the 11 subjects who submitted complete pretest and posttest results data, and 12 of 13 subjects who submitted posttest results scored >90% accuracy on the posttest. The *top line* represents 2 subjects who had a pretest score of 96% and a posttest score of 100%.

Figure 2.

Learning curves as a function of performance on nonoverlapping batches of 20 consecutive polyps per endoscopist assessed with high confidence. There was considerable batch-to-batch variation, which was more pronounced for individual participants than for the group mean overall and for specificity compared with sensitivity. There was no clear pattern of early learning with later stabilization of performance at a higher level. (A) Accuracy, (B) sensitivity, and (C) specificity for diminutive polyps. (D) Accuracy, (E) sensitivity, and (F) specificity for diminutive and small polyps combined. Each participant is represented by a different color. The group average is superimposed in *red*.

Figure 3.

Learning curves as a function of performance on nonoverlapping batches of 20 consecutive polyps per endoscopist assessed with high or low confidence. There was considerable batch-to-batch variation, which was more pronounced for individual participants than for the group mean overall and for specificity compared with sensitivity. There was no clear pattern of early learning with later stabilization of performance at a higher level. (A) Accuracy, (B) sensitivity, and (C) specificity for diminutive polyps. (D) Accuracy, (E) sensitivity, and (F) specificity for diminutive and small polyps combined. Each participant is represented by a different color. The group average is superimposed in red.

Table 1

Characteristics of Participants, Study Colonoscopies, and Study Polyps

	Ex vivo phase	In vivo phase
Participants, n	14	12
Endoscopy practice experience (y), median (interquartile range)	16 (7–22)	12 (6–21)
Colonoscopy volume per year, median (interquartile range)	922 (839–1056)	901 (803–1105)
Adenoma detection rate, median (interquartile range)	34% (29%–37%)	35% (30%–38%)
Study colonoscopies, n	—	1673
With 1 polyp	—	1107 (66%)
With 2 polyps	—	355 (21%)
With 3 polyps	—	116 (7%)
With >3 polyps	—	95 (6%)
Study colonoscopies per endoscopist, median (interquartile range)	—	121 (53–180)
Study polyps, n ^a	—	2596
Diminutive polyps (< 5 mm)	—	1858 (72%)
Small polyps (6–9 mm)	—	547 (21%)
Large polyps (≥ 10 mm)	—	177 (7%)
Study polyps per endoscopist, median (interquartile range)	—	174 (96–278)
Diminutive polyps (< 5 mm)	—	135 (59–200)
Small polyps (6–9 mm)	—	29 (18–83)
Large polyps (≥ 10 mm)	—	9 (4–15)

^a Adenomas accounted for 62% of diminutive, 72% of small, and 78% of large polyps.

Table 2

Ex Vivo Test Results

	Pretest	Posttest	Change	P value
Participants who submitted complete score sheets, n	12 of 14	13 of 14		
Items answered "I don't know"	7%	0%	-7%	.134
Score (% correct) with items answered "I don't know" counted as wrong, mean (SD)	79 (15)	94 (5)	16 (15)	.006
Score (% correct) with items answered "I don't know" excluded, mean (SD)	85 (9)	94 (5)	10 (8)	.002

Gastroenterology. Author manuscript; available in PMC 2017 July 20.

Diminutive polyps (< 5 mm)						Small polyps (6–9 mm)		
No.	1858					547		
Adenoma (% of polyps)	1150 (61.9)					391 (71.5)		
Endoscopists, n	12					11		
Sensitivity, mean (95% CI)	86.5 (80.9–92.1)					95.5 (92.8–98.2)		
Specificity, mean (95% CI)	64.7 (54.9–74.6)					28.1 (13.7–42.4)		
Positive predictive value, mean (95% CI)	79.8 (74.3–85.3)					79.8 (74.3–85.3)		
Negative predictive value, mean (95% CI)	75.9 (69.1–82.7)					61.0 (36.6–85.3)		
Accuracy, mean (95% CI)	78.1 (73.7–82.5)					78.5 (73.6–83.4)		
First vs last batches to explore learning effect								
Diminutive polyps (< 5 mm)						Small polyps (6–9 mm)		
	First 20 per endoscopist	Last 20 per endoscopist	Mean (SD) difference	P value	First 10 per endoscopist	Last 10 per endoscopist	Mean (SD) difference	P value
No.	231	231	—	—	104	104	—	—
Adenoma (% of polyps)	134 (58.0)	154 (66.7)	—	—	86 (82.7)	74 (71.2)	—	—
Endoscopists, n	12	12	—	—	11	11	—	—
Sensitivity, mean (95% CI)	76.6 (65.3–88.0)	90.2 (81.5–98.9)	−13.5 (20.6)	.044	91.7 (83.4–100.0)	95.2 (90.6–99.8)	−3.5 (13.1)	.403
Specificity, mean (95% CI)	76.6 (63.3–89.9)	55.3 (36.9–73.8)	21.2 (20.8)	.005	30.3 (13.8–46.8)	24.5 (3.7–45.4)	5.8 (42.8)	.665
Positive predictive value, mean (95% CI)	79.7 (67.2–92.1)	80.8 (74.1–87.6)	−1.2 (21.7)	.854	88.0 (83.9–92.0)	78.4 (65.2–91.5)	9.6 (19.0)	.123

	Diminutive polyps (< 5 mm)			Small polyps (6–9 mm)		
	Rectosigmoid	Proximal to rectosigmoid	Mean (SD) difference	Rectosigmoid	Proximal to rectosigmoid	Mean (SD) difference
No.	640	1214	—	180	366	—
Adenoma (% of polyps)	234 (36.6)	914 (75.3)	—	109 (60.6)	282 (77.0)	—
Endoscopists, n	12	12	—	9	11	—
Sensitivity, mean (95% CI)	79.4 (67.9–90.9)	88.2 (82.2–94.2)	–8.8 (18.0)	98.7 (95.8–100.0)	94.2 (90.7–97.7)	5.2 (5.8)
Specificity, mean (95% CI)	76.3 (66.1–86.6)	49.7 (34.7–64.6)	26.7 (22.8)	45.6 (16.2–75.0)	21.1 (5.6–36.6)	19.8 (51.9)
Positive predictive value, mean (95% CI)	66.3 (50.7–82.0)	85.0 (81.5–88.5)	–18.7 (24.6)	80.9 (69.8–91.9)	81.0 (75.5–86.5)	–0.8 (15.3)
Negative predictive value, mean (95% CI)	87.4 (82.5–92.4)	57.3 (38.4–76.2)	30.1 (30.7)	75.7 (42.4–100.0)	37.4 (11.2–63.6)	30.0 (78.5)
Accuracy, mean (95% CI)	77.4 (69.1–85.3)	79.3 (74.7–83.9)	31.9 (13.5)	83.6 (74.3–92.9)	77.8 (73.2–82.4)	5.3 (12.5)
	Diminutive polyps (< 5 mm)			Small polyps (6–9 mm)		
	Low confidence	High confidence	Mean (SD) difference	Low confidence	High confidence	Mean (SD) difference
No.	368	1481	—	57	485	—
Adenoma (% of polyps)	210 (57.1)	934 (63.1)	—	29 (50.9)	360 (74.2)	—
Endoscopists, n	12	12	—	10	11	—
Sensitivity, mean (95% CI)	80.0 (72.7–87.4)	88.4 (82.2–94.7)	–8.4 (13.1)	56.1 (29.5–82.7)	97.8 (96.1–99.6)	–41.5 (38.3)

	Diminutive polyps (≤ 5 mm)			Small polyps (6–9 mm)		
Specificity, mean (95% CI)	88.4 (82.2–94.7)	44.1 (26.5–61.6)	–24.2 (13.1)	.017	45.7 (14.7–76.6)	20.2 (4.6–35.7)
Positive predictive value, mean (95% CI)	72.1 (59.0–85.3)	82.8 (77.0–88.6)	–10.7 (21.3)	.111	57.9 (30.6–85.3)	84.3 (77.9–90.7)
Negative predictive value, mean (95% CI)	51.8 (35.3–68.3)	78.3 (69.6–87.0)	–26.5 (32.0)	.015	47.1 (15.0–79.2)	52.2 (22.1–82.3)
Accuracy, mean (95% CI)	70.4 (58.9–82.0)	81.1 (75.8–86.3)	–10.6 (20.5)	.100	66.5 (47.6–85.3)	83.8 (77.7–90.0)
	Subgroups of interest ^b					
	Diminutive polyps (≤ 5 mm)				Small polyps (6–9 mm)	
	Last 20 per endoscopist, all locations, high confidence	Last 20 per endoscopist, rectosigmoid, high confidence			Last 15 per endoscopist, all locations, high confidence	Last 15 per endoscopist, rectosigmoid, high confidence
Number	240	219	–	–	150	98
Adenoma (% of polyps)	162 (67.5)	88 (40.2)	–	–	113 (75.3)	37 (37.8)
Endoscopists, n	12	12	–	–	11	9
Sensitivity, mean (95% CI)	92.4 (86.5–98.4)	85.0 (74.3–95.7)			97.9 (93.2–100.0)	99.1 (96.9–100.0)
Specificity, mean (95% CI)	57.7 (36.9–78.5)	78.4 (66.0–91.7)			20.0 (1.9–38.1)	41.5 (9.1–73.9)
Positive predictive value, mean (95% CI)	84.4 (77.7–91.1)	71.5 (55.2–87.8)			88.3 (83.0–93.5)	82.7 (70.8–94.5)
Negative predictive value, mean (95% CI)	69.3 (45.7–93.0)	91.4 (86.3–96.5)			38.6 (5.6–71.7)	63.0 (25.7–100.0)
Accuracy, mean (95% CI)	83.3 (76.7–90.0)	81.2 (73.3–89.0)			87.2 (81.4–93.0)	84.3 (74.5–94.2)

^aThe overall fraction of high-confidence characterizations of diminutive polyps was 76% for the first 50 polyps per participant and 81% for the subsequent polyps per participant.

^bBecause polyps may belong to both categories, statistical comparison is not appropriate.

Table 4
In Vivo Surveillance Intervals Recommendations Based on High-Confidence Characterizations

	Agreement				
	Recommended surveillance interval No. (%)				
	10 y	5–10 y	3 y	Percent agreement (95% CI)	κ value
All study colonoscopies					
Using the Multi-Society Task Force recommendations					
Diminutive polyps assessed by NBI	466 (28.3)	957 (58.1)	223 (13.6)	88.4 (86.8–89.9)	0.795
All polyps assessed by pathology	507 (30.8)	931 (56.6)	208 (12.6)		<.001
Using modified recommendations (10 y for 1–2 small adenomas)					
Diminutive polyps assessed by NBI	1423 (86.5)	—	223 (13.6)	98.4 (97.6–98.9)	0.928
All polyps assessed by pathology	1438 (87.4)	—	208 (12.6)		<.001
Study colonoscopies with at least one diminutive polyp characterized with high confidence					
Using the Multi-Society Task Force recommendations					
Diminutive polyps assessed by NBI	357 (33.5)	578 (54.3)	130 (12.2)	79.9 (77.4–82.3)	0.654
All polyps assessed by pathology	402 (37.8)	547 (51.4)	116 (10.9)		<.001
Using modified recommendations (10 y for 1–2 small adenomas)					
Diminutive polyps assessed by NBI	935 (87.8)	—	130 (12.2)	96.8 (95.6–97.8)	0.844
All polyps assessed by pathology	949 (89.1)	—	116 (10.9)		<.001
Only the last 20 colonoscopies per endoscopist with at least one diminutive polyp characterized with high confidence					
Using the Multi-Society Task Force recommendations					
Diminutive polyps assessed by NBI	50 (21.7)	156 (67.5)	25 (10.8)	82.7 (77.2–87.3)	0.649
All polyps assessed by pathology	65 (28.1)	150 (64.9)	16 (6.9)		<.001
Using modified recommendations (10 y for 1–2 small adenomas)					
Diminutive polyps assessed by NBI	206 (89.2)	—	25 (10.8)	97.0 (93.9–98.8)	0.784
All polyps assessed by pathology	215 (93.1)	—	16 (6.9)		<.001