Can Radiologist Training and Testing Ensure High Performance in CT Colonography? Lessons From the National CT Colonography Trial

Joel G. Fletcher[1], **Mei-Hsiu Chen**[2], **Benjamin A. Herman**[2,3], **C. Daniel Johnson**[1], **Alicia Toledano**[2], **Abraham H. Dachman**[4], **Amy K. Hara**[5], **Jeff L. Fidler**[1], **Christine O. Menias**[6], **Kevin J. Coakley**[7], **Mark Kuo**[8], **Karen M. Horton**[9], **Jugesh Cheema**[10], **Revathy Iyer**[11], **Bettina Siewert**[12], **Judy Yee**[13], **Richard Obregon**[14], **Peter Zimmerman**[15], **Robert Halvorsen**[16], **Giovanna Casola**[17], and **Martina Morrin**[12]

[1] Mayo Clinic, 200 First St., SW, Rochester, MN 55905

[2] Brown University, Providence, RI

[3] Statistics Collaborative, Inc. Washington, DC

[4] The University of Chicago Medical Center, Chicago, IL

[5] Mayo Clinic, Scottsdale, AZ

[6] Mallinckrodt Institute of Radiology, St. Louis, MO

[7] Clinical Radiologists, S.C., Springfield, IL

[8] Scottsdale Medical Imaging, Ltd., Scottsdale, AZ

[9] Johns Hopkins University, Baltimore, MD

[10] VA Connecticut Healthcare System, West Haven, CT

[11] University of Texas M. D. Anderson Cancer Center, Houston, TX

[12] Beth Israel Deaconess Medical Center, Boston, MA

[13] UCSF VA Medical Center, San Francisco, CA

[14] Radiology Imaging Associates, Englewood, CO

[15] University of California, Los Angeles, Los Angeles, CA

[16] Virginia Commonwealth University, Richmond, VA

[17] University of California, San Diego, San Diego, CA

Abstract

OBJECTIVE—The objective of this article is to describe the experience of the National CT Colonography Trial with radiologist training and qualification testing at CT colonography (CTC) and to correlate this experience with subsequent performance in a prospective screening study.

SUBJECTS AND METHODS—Ten inexperienced radiologists participated in a 1-day educational course, during which partial CTC examinations of 27 cases with neoplasia and full CTC examinations of 15 cases were reviewed using primary 2D and 3D search. Subsequently 15 radiologists took a qualification examination composed of 20 CTC cases. Radiologists who did not pass the first qualification examination attended a second day of focused retraining of 30 cases, which was followed by a second qualification examination. The results of the initial and subsequent qualification tests were compared with reader performance in a large prospective screening trial.

RESULTS—All radiologists took and passed the qualification examinations. Seven radiologists passed the qualification examination the first time it was offered, and eight radiologists passed after focused retraining. Significantly better sensitivities were obtained on the second versus the first examination for the retrained radiologists (difference = 16%, $p < 0.001$). There was no significant difference in sensitivities between the groups who passed the qualification examination the first time versus those who passed the second time in the prospective study (88% vs 92%, respectively; $p = 0.612$). In the prospective study, the odds of correctly identifying diseased cases increased by 1.5 fold for every 50-case increase in reader experience or formal training ($p < 0.025$).

CONCLUSION—A significant difference in performance was observed among radiologists before formalized training, but testing and focused retraining improved radiologist performance, resulting in an overall high sensitivity across radiologists in a subsequent, prospective screening study.

Keywords

ACRIN; CT; CT colonography; radiologist training; reader performance

---

Ct colonography (CTC) is becoming an established test for the detection of colorectal polyps and cancer for a variety of clinical indications. In some states, Medicare and other insurance providers reimburse for diagnostic CTC examinations after incomplete endoscopy. Screening CTC is now endorsed by the American Cancer Society [1], and screening examinations are reimbursed in some states by some private payers for normal-risk patients or those with contraindications to endoscopy [2–4]. A large screening study by Pickhardt and colleagues [5] showed that CTC can depict colorectal polyps with a sensitivity on par with that of optical colonoscopy; however, several other large clinical studies of symptomatic and surveillance populations have shown that observer performance can vary widely [6,7]. Furthermore, although a lack of reader experience may, in part, explain the suboptimal performance in the study by Cotton et al. [8], two other studies, one performed by Rockey and colleagues [6] and the other by Johnson et al. [7], showed that experience with CTC does not guarantee excellent performance.

The National CT Colonography Trial (American College of Radiology Imaging Network [ACRIN] 6664) is a multiinstitutional trial of screening CTC in an average-risk patient population funded by the National Cancer Institute. Its primary aim was to assess the multiinstitutional performance of CTC with cathartic preparation in a screening population [9]. To overcome the effects of reader inexperience and minimize interobserver variability, the National CT Colonography Trial instituted mandatory reader training and performance standards to ensure that the participating radiologists had adequate experience and aptitude in CTC. The purpose of this article is to describe our experience with training and qualification testing at CTC and to correlate results from our training methods with subsequent performance in a prospective screening trial.

Subjects and Methods

Initial Reader Training

The National CT Colonography Trial required radiologists at participating sites to pass a qualification examination before study initiation. Participating radiologists who had interpreted fewer than 500 CTC examinations with colonoscopic correlation were required to participate in a mandatory CTC training session before taking the qualification examination. Experienced readers, those who had interpreted at least 500 cases, did not need to participate in the mandatory colonography training but were required to pass the qualifying examination. All cases used in reader training had complete endoscopic correlation. During this training phase, radiologists were unblinded to the 3D location of each colorectal polyp or cancer (slice number and *x-y-z* coordinates) after reviewing each case.

CTC training was conducted in Milwaukee, WI, on November 4, 2004. Advantage Windows 4.2 workstations (GE Healthcare) using CTC and autodissection software (CTC Auto Dissection, GE Healthcare) to facilitate 3D fly-through of the colon were used for the training session. Colonic centerlines were generated for each case before the course began, with the CTC software providing interactive 2D and 3D viewing, simultaneous comparison of prone and supine images, and variable window settings. Although radiologists were permitted to view virtual dissection images, only traditional-perspective, volume-rendered endoluminal images were reviewed during instruction.

In a single day of training, radiologists received didactic instruction from an experienced team consisting of two radiologists and an applications specialist. The instructors reviewed workstation functions (i.e., "buttonology") and 42 colonography cases. The first 10 cases consisted of different morphologies of cancers: annular, flat, and polypoid (Table 1). Cases 11 through 27 showed different morphologies and features of colorectal polyps (e.g., sessile, pedunculated, flat, or villous) (Table 1). In these first 27 cases, radiologists were instructed to look for colorectal lesions in particular colonic segments, but they were not told the specific locations of the lesions. The purpose of this initial phase of training was to familiarize radiologists with the CTC workstation, to show the spectrum of appearances of colorectal neoplasia found at CTC, and to teach interactive problem-solving skills (e.g., correlation of supine and prone images, correlation of 2D multiplanar reconstructions and 3D endoluminal views, evaluation of internal attenuation characteristics). For the last 15 cases (12 positive, three negative), radiologists were required to interpret full supine and prone colonography examinations using a primary 2D or 3D search to detect colonic lesions before being unblinded to the truth. They were unaware of the number of positive and negative cases. The purpose of this latter phase of training was to teach methods of searching for colorectal lesions using 2D and 3D fly-through search techniques.

Qualification Examination

The initial qualification examination consisted of 20 CTC cases: 17 positive cases with polyps verified by colonoscopy and three negative cases without any colorectal polyps. The cases in this qualification examination were retrospectively selected from research and clinical CTC examinations performed at Mayo Clinic Rochester. All cases included in the qualification examination had complete endoscopic correlation. For the positive cases, endoscopically proven colorectal polyps were prospectively identified either by a radiologist as part of a prospective study at Mayo Clinic Rochester or by a clinical radiologist before referral for polypectomy. All cases were obtained under institutional review board (IRB)–approved protocols (i.e., either a prospective, IRB-approved study with informed consent or

an IRB-approved database of patients who had undergone CTC for clinical purposes and consented to the use of medical records for research purposes).

In the initial qualification examination, there were a total of 25 polyps ≥ 5 mm in the 20 cases, 15 of which were ≥ 10 mm. Two experienced gastrointestinal radiologists with more than 5 years of CTC experience interpreting more than 1,500 cases had individually correlated the CTC findings with the endoscopic findings in these 20 cases (17 patients with polyps or cancers, three without). In a consensus review of the 2D and 3D images, these experienced colonographers characterized the polyps as easy to detect ($n = 13$), moderately difficult to detect ($n = 7$), or difficult to detect ($n = 5$) on the basis of lesion conspicuity and their teaching and clinical experience. The size, CT morphology (as defined previously [10]), and histology of these polyps is given in Table 2.

To pass the qualification examination, radiologists at participating institutions were required to detect 90% of the easy- and moderate-to-detect polyps ≥ 5 mm. The difficult-to-detect polyps were not scored as part of the examination, but they were included to emphasize subtle polyp morphologies (subtle flat lesions, pedunculated polyps with difficult-to-visualize geometries) for educational purposes (Fig. 1). The easy-to-detect and moderate-to-detect polyps included polyps that clearly projected into the colonic lumen on 2D and 3D images (whether sessile, pedunculated, or flat) or those that were annular but may have been visible in only one position. The interpreting radiologists were blinded to the case mix (i.e., number, type, conspicuity, and size of polyps) in the qualification examination.

The CTC examinations were performed using a standard low-dose technique (≤ 85 mAs) and an MDCT scanner (8 or 16 channels) after patients underwent bowel purgation cleansing. Positive oral contrast medium was used for stool tagging in only two patients. In all patients except one, the examinations were performed using a slice thickness of ≤ 2.5 mm. In the 20 test cases, four colonic segments (as previously defined [10]) in three patients were inadequately visualized when comparing the supine and prone images because of fluid ($n = 2$) or because of collapse, spasm, or muscular hypertrophy due to diverticular disease ($n = 2$). None of the lesions was located in these four inadequately visualized segments.

The initial qualification examination consisting of 20 CTC cases was delivered to radiologists to be read on a computer workstation of their choosing at their home institution and were completed between November 21, 2004, and January 3, 2005. There was no time limit placed on the radiologists for the qualification examination, similar to the planned study design for the prospective trial. Interpretations were recorded on data forms and sent back to the principal investigator. None of the test materials had been previously seen by any of the radiologists. The two radiologists who created the qualification examination matched CT-detected lesions against the endoscopic reference standard using standard matching rules (i.e., within 50% size and one colonic segment and of similar morphology) to estimate reader performance.

Focused Retraining

After the receipt and scoring of the initial qualification examinations, it became clear that although some radiologists performed exceptionally well, many radiologists did not achieve the preset threshold for participation in the study. Radiologists who did not meet the preset threshold were asked to participate in an additional day of focused training involving more cases using primary 2D and 3D search methods. Radiologists who had passed the initial qualification examination were also allowed to participate if they desired. One day of focused retraining using primary 2D or primary 3D interpretation of full CTC cases was arranged in Rochester, MN, on January 24, 2005.

Before the second training session, an experienced radiologist reviewed all false-negative interpretations in the qualifying examination, listing polyp characteristics or colonic conditions that may have contributed to each polyp not being identified by radiologist (Table 3). The proportion of radiologists who correctly scored each lesion was calculated so that a list of problematic polyp characteristics and colonic conditions associated with lesions having been missed by ≥ 20% of the radiologists could be created. Eleven polyps were missed by ≥ 20% of the radiologists—that is, by at least three radiologists. Over 80% of the missed polyps were either sessile or flat (4/11 and 5/11, respectively), and approximately one quarter of the missed polyps were located on a fold (3/11, 27%), could be seen on only one view (3/11, 27%), had an undulating surface (4/11, 36%), or looked like a bulbous or thickened fold (3/11, 27%). Cases selected for the focused retraining session therefore consisted of a set of 30 cases containing polyps possessing these characteristics (Table 3).

Because the prospective study would use stool tagging, additional teaching cases that used oral contrast material were selected for focused retraining, even though only two cases in the qualifying examination had used oral tagging. Fourteen of the lesions in the retraining set were submerged lesions outlined by tagged fluid. The 30 CTC cases chosen for retraining were a mixture of cases from Mayo Clinic Rochester and CTC cases from the Pickhardt et al. [5] trial, now in the public domain and made available by J. Richard Choi [11], with 22 positive cases containing colorectal polyps or cancers and eight negative cases. None of the cases provided in the additional training was part of the qualifying test. The cases were interpreted using Vitrea 4.1 workstations (Vital Images) and primary 2D and 3D search techniques.

After 1 day of focused retraining on January 25, 2005, the radiologists required to undergo additional training took a second examination consisting of an individualized set of eight CTC cases. Each examination was tailored for each participant to include all of the cases he or she had missed on the first examination (up to six cases), positive and negative cases (at least one of each), and one difficult-to-detect polyp (if needed), so the second qualifying examination consisted of a total of eight cases. The radiologists were not aware of the lesions they had missed during the initial qualification examination and had not reviewed the cases after the initial qualification examination. The second examination occurred on a single morning using Vitrea 2 computer workstations (Vital Images) and had to be completed within 4 hours because of time constraints imposed by space rental. Only the originally missed cases from the first qualification examination were graded, with the added negative and difficult-to-detect cases not graded. For purposes of the qualification examination, the results of the second test replaced the results of the first test for the cases that were repeated. Qualification examination scores for each radiologist participating in this second day of training were subsequently recalculated.

Statistical Analysis

The sensitivities from the qualification examinations were calculated by dividing the number of true-positives with the number of polyps ≥ 5 mm for each reader. Unlike the sensitivity threshold for the easy- and moderate-to-detect polyps that had to be met by each radiologist to participate in the prospective study, the sensitivities reported herein are for all polyps, including the difficult-to-detect ones. All polyps were included in these analyses to mimic performance estimates in the prospective study. Sensitivities from the initial qualification examination were compared between the group of readers who passed the examination the first time (passers) and the group of readers who required focused retraining (retrainees) using generalized estimating equation (GEE) modeling. For retrainees, the sensitivities from the first and second qualifications examinations were also compared using GEE models. The Mann-Whitney test was used to compare the sensitivities from the prospective study between passers and retrainees. For all readers, the differences in sensitivities between the

final qualification examination and the prospective study were tested using the Wilcoxon's signed rank test.

Prestudy self-reported experience was first categorized and was then tested using the exact Kruskal-Wallis test for its effect on the accuracy from the prospective study. The effect of reader experience and training on accuracy was explored further by modeling the probability of a true-positive from the prospective study using logistic regression with several variables to categorize a variety of reader experience, such as the total number of cases read and the level of training, as covariates. Statistical analyses were carried out using SAS software (version 9.1.3, SAS Institute).

Results

Eleven readers were inexperienced with CTC and four were experienced. Twelve readers were radiologists working at academic institutions and three were in private practice. All had a strong interest and practice in abdominal and cross-sectional imaging. Ten inexperienced readers participated in the 1-day training course, and one inexperienced reader participated in a CTC continuing medical education course outside the formalized ACRIN training. All 15 radiologists eventually achieved the preset criteria for participation in the prospective trial—that is, $\geq$ 90% sensitivity for easy- and moderate-to-detect polyps $\geq$ 5 mm.

The initial completed qualification examinations were received from each of the participating radiologists. The median time over which test cases were reviewed was 4 days (range, 2–31 days). The median time for interpreting each case across radiologists was 20 minutes (range of medians, 15–26 minutes).

Box plots of the sensitivities from the initial qualifying examination for the easy-, moderate-, and difficult-to-detect polyps are displayed in Figure 2. As we expected, the sensitivity of the radiologists decreased as the difficulty in detecting polyps increased: The mean sensitivities for easy-, moderate-, and difficult-to-detect polyps were 92%, 76%, and 42%, respectively.

Seven of the readers (7/15, 47%) passed the initial qualification examination, but eight (8/15, 53%) did not, achieving a sensitivity of between 70% and 89% for easy- and moderate-to-detect lesions on the initial qualifying examination. A breakdown of readers' experience and participation in the focused retraining is given in Table 4. Two of the inexperienced readers who passed the initial examination opted to attend the focused retraining education, but did not take the second qualification examination.

The group of seven readers who passed the qualification examination the first time it was offered (initial passers) performed significantly better on the first examination than those requiring focused retraining (average sensitivity, 85% vs 71%, $p < 0.004$) (Table 5). However, those undergoing focused retraining achieved significantly better sensitivities on the second versus first qualification test (average sensitivity, 87% vs 71%, respectively; difference = 16%; $p < 0.001$), with their performance on the second examination being essentially identical to the initial passing group (87% vs 85%, respectively). Furthermore, in the prospective study, there was no statistically significant difference in sensitivities between the initial passers and the retrainees (average sensitivity for adenomatous polyps and cancers $\geq$ 10 mm: 88% vs 92%, respectively; $p = 0.612$) (Table 5). In the prospective study the overall per-patient sensitivity for adenomatous polyps and cancers $\geq$ 10 mm was 90% and the per-lesion sensitivity for polyps of that size was 84% [9].

When the final results of the qualification examination for each reader were compared with overall performance in the prospective study, there was no significant difference in sensitivity for the detection of polyps; the mean sensitivities for the final qualification examination and for the prospective study were 87% vs 90%, respectively ($p = 0.210$). When comparing the prestudy self-reported level of experience, there was no difference in sensitivity in the prospective trial among readers who had interpreted fewer than 100 cases, 100–200 cases, and more than 200 cases (exact Kruskal-Wallis test, $p = 0.122$).

Because readers gain experience with each case they read, we explored several variables to account for learning effect on the probability of correctly identifying a diseased case. In particular, study experience was defined as the cumulative number of cases read during the study. Reader experience was defined as the study experience plus the self-reported number of cases read with endoscopic correlation before the start of the study, and reader training was defined as the study experience plus the number of training cases reviewed. The probability of correctly identifying a diseased case increased significantly with reader experience: With every 50 cases read, a 1.5 times higher odds of correctly identifying a diseased case was observed ($p = 0.025$). If we used prestudy training to define experience instead of the self-reported prestudy experience, we found a similar effect such that the odds of correctly identifying a diseased case increased by a factor of 1.5 for every 50 cases read ($p = 0.004$). Other covariates, such as the cumulative number of positive cases or study experience alone, did not significantly affect the probability of correctly identifying disease in the models fit.

Discussion

The experience of the ACRIN 6664 National CT Colonography Trial shows the importance of rigorous training supplemented by qualification testing in CTC. We found that inexperienced radiologists can interpret CTC examinations with a high degree of accuracy after appropriate training and have validated a training and testing paradigm that produced excellent sensitivity across readers in a large prospective study. Although a significant difference in performance was observed among readers before adequate training, testing and focused retraining significantly improved the performance of readers who underwent additional training so that performance differences no longer existed. Improved and higher performance was maintained across readers throughout a large prospective study. Although there was no statistically significant difference in reader performance between reader groups, the probability that a reader could correctly identify a positive case was highly correlated with both reader experience and reader training.

In prior studies, investigators have documented the need for and benefit of both formal CTC training [6,12–14] and reader experience in interpreting CTC examinations [7,12,15–17]. Our inexperienced radiologists reviewed positive colon segments in 27 cases followed by 2D or 3D search through the entire colorectum for suspicious lesions in only 15 cases in the initial 1-day training session. Time and fiscal constraints limited this initial training period. Our goal had been to review 50 full cases, as recommended in American College of Radiology (ACR) practice guidelines [18]. The additional day of focused retraining permitted inexperienced radiologists to interpret 30 more CTC cases (for a total of 45 full cases) using a primary 2D or 3D read. Although most previous studies suggest that experience in interpreting 75 cases is required to achieve acceptable levels of sensitivity [12,15,16], we found that 45 cases was sufficient for participating radiologists when experience was supplemented by formal CTC training. This reduction in experience may have occurred because our focused retraining used cases designed to address the needs of the participating radiologists. Ideally training courses will have a large library of difficult-to-detect lesions that radiologists miss frequently, so iterative feedback can be followed by

reexamination of similar-appearing lesions. Interpreting a large volume of CTC cases using primary 2D and 3D search for colonic lesions is imperative for beginning readers, who tend to have lower sensitivity because of failures of detection (i.e., not stopping to consider a potential polyp candidate), as opposed to failures of characterization (i.e., considering a potential polyp using known problem-solving techniques and erroneously dismissing it) [14,17]. Our radiologists were abdominal radiologists, and greater experience may be needed for nonabdominal radiologists or gastroenterologists [19].

Burling et al. [13] found that experience alone—that is, without endoscopically validated cases or formal training—is insufficient for interpreting CTC with a high sensitivity. This realization from prior prospective studies [6,7] formed the basis for setting the high threshold for experience (i.e., 500 cases) needed to exempt readers in the current study from the need for formal training. Halligan and the European Society of Gastrointestinal and Abdominal Radiology CT Colonography Study Group Investigators [12] first reported on the training and testing of inexperienced radiologists and technologists in a multisite setting. Nine radiologists and 10 radiologic technologists were trained using 50 CTC cases from their institution, with the inexperienced readers taking a test of 40 cases along with an experienced radiologist from their institution. They found that the experienced radiologists performed significantly better than the inexperienced but trained radiologists and technologists even when difficult-to-detect lesions were not considered [12]. Differentiating features between their experience and that reported here include formal training and uniform selection of teaching cases known to be comparable to the spectrum of CTC findings and difficulty of cases encountered in clinical practice. Additionally, we used qualification testing and focused retraining to provide each radiologist with feedback about polyp types and morphologies that they might find difficult to detect. Similar to our experience, Dachman et al. [20] used a comprehensive training program that included a 1-day teaching course, reading assignments from a CTC atlas, a self-study computer module with 61 CTC cases, and a prospective review of 10 cases using primary search with unblinding after each case. Seven novice readers then interpreted 60 additional cases, being unblinded to the truth by an experienced radiologist after interpreting each case. This integrated teaching and training methodology yielded a mean sensitivity across readers for polyps ≥ 6 mm of 91% [20].

Although our experience cannot discriminate between the benefit of qualification testing versus the appropriate level of CTC case experience, it does show that appropriate case selection for CTC teaching is crucial. It is interesting to note that the characteristics of polyps missed by our group (Table 2) are virtually identical to those polyp and colon characteristics identified in a large review of false-negative CTC examinations [10]. It is likely that there are common characteristics of difficult-to-detect polyps; in fact, although some radiologists performed exceptionally well from the outset, most radiologists tended to miss lesions with the same imaging characteristics (e.g., flat, undulating surface, long stalk, seen on one view only, located on fold, bulbous fold, suboptimal preparation) (Table 2). We consequently selected teaching cases for focused retraining on the basis of these morphologic characteristics (Table 3).

The probability of a radiologist interpreting a case with disease correctly was affected by his or her experience as well as formal training. These results confirm those reported by Bodily et al. [21] for nonradiologists, who continued to learn and improve in performance well after 50 cases of experience. Optimal performance over time will likely be achieved with continuing training and experience, which have been advocated by the 2009 ACR Practice Guideline for CTC [19].

Colonography training should include multiple principles [19] that were neither taught in our ACRIN training nor assessed as part of our qualifying examinations. Our reader training and testing focused only on facile use of CTC computer software and interpretive ability. It did not address the appropriateness of clinical indications for CTC (e.g., screening, incomplete endoscopy, increased endoscopic risk or aversion, nonagenarians with a clinical suspicion of cancer), risks of the examination, and CT acquisition techniques and risks. We did not comprehensively teach problem-solving techniques, such as the use of IV contrast-enhanced imaging and decubitus imaging, that should be used to increase diagnostic confidence before patient dismissal. Furthermore, our training did not review the training of technologists, facility requirements, quality control processes, or documentation of results [19]. Finally, formal CTC courses should include a discussion of standardized reporting of intra- and extracolonic findings [22].

There are numerous scientific limitations to our experience in the training and qualification testing with CTC principally owing to the fact that our primary aim was to train radiologists to participate in a large prospective screening study, and we consequently had significant time and fiscal constraints. Limitations include the lack of tagging in cases included in the original qualification test; a solitary day of training devoted for initial training; CTC workstation training occurring on a commercial system often different from the one used in the subsequent prospective study; inadequate time to readminister the entire qualification test to the retrainees; and self-reported experience before training. Given the time constraints, our training was carried out on only two software platforms, and we did not take each radiologist's workstation into account when evaluating results of the qualification test. The subsequent prospective study did not show a performance difference between the workstations. Finally, 15 readers were selected to provide appropriate power to address the aims of the prospective study, and no power calculation was carried out with respect to this reader performance study. Nonetheless, we believe our experience represents a valuable, novel, and practical contribution to the CTC literature because it shows a successful model of CTC training.

In conclusion, learning CTC is time-intensive. However, formalized training and testing can result in high levels of performance across readers in interpreting screening CTC examinations. Our experience suggests that educational courses in CTC for abdominal radiologists should provide approximately 50 cases in which a full search of the colorectum for suspicious lesions is performed. These cases should have been endoscopically validated and carefully selected to include difficult-to-detect polyps and cancers (Table 3). Training courses in CTC should consequently be designed to provide experience with a large number of cases after radiologists become familiar with CTC workstations, coupled with feedback on performance and focused review of interpretive pitfalls. Testing can identify radiologists likely to benefit from additional instruction and provide radiologists useful insights into idiosyncratic strengths and weaknesses. An alternative may be the institution of outcomes assessment metrics in a radiologist's practice (i.e., calculating sensitivity of CTC for patients referred to endoscopy and false-positive rate) [23]. Finally, continuing training and experience are needed to maintain optimal levels of performance.

Acknowledgments

References

1. Levin B, Lieberman DA, McFarland B, et al. American Cancer Society Colorectal Cancer Advisory Group; US Multi-Society Task Force; American College of Radiology Colon Cancer Committee. 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 CA. Cancer J Clin 2008;58:130–160.

2. Pickhardt PJ. Virtual colonoscopy: issues related to primary screening. Eur Radiol 2005;15(suppl 4):D133–D137. [PubMed: 16479663]

3. McFarland EG, Fletcher JG, Pickhardt P, et al. American College of Radiology. ACR Colon Cancer Committee white paper: status of CT colonography 2009. J Am Coll Radiol 2009;6:756–772. e4. [PubMed: 19878883]

4. McFarland EG, Fletcher JG, Pickhardt P, et al. ACR Colon Cancer Committee white paper: status of CT colonography 2009. AJR 2009;193:1242.

5. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 2003;349:2191–2200. [PubMed: 14657426]

6. Rockey DCPE, Niedzwiecki D, Davis W, et al. Analysis of air contrast barium enema, computed tomographic colonography, and colonoscopy: prospective comparison. Lancet 2005;365:305–311. [PubMed: 15664225]

7. Johnson CD, Harmsen WS, Wilson LA, et al. Prospective blinded evaluation of computed tomographic colonography for screen detection of colorectal polyps. Gastroenterology 2003;125:311–319. [PubMed: 12891530]

8. Cotton PB, Durkalski VL, Pineau BC, et al. Computed tomographic colonography (virtual colonoscopy): a multicenter comparison with standard colonoscopy for detection of colorectal neoplasia. JAMA 2004;291:1713–1719. [PubMed: 15082698]

9. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. N Engl J Med 2008;359:1207–1217. [Erratum in *N Engl J Med* 2008; 359:2853]. [PubMed: 18799557]

10. Gluecker TM, Fletcher JG, Welch TJ, et al. Characterization of lesions missed on interpretation of CT colonography using a 2D search method. AJR 2004;182:881–889. [PubMed: 15039159]

11. Virtual Colonoscopy Screening Resource Center. [Accessed July 22, 2009]. Websitevcscreen.com. Updated December 16, 2005

12. Halligan S. European Society of Gastrointestinal and Abdominal Radiology CT Colonography Study Group Investigators. Effect of directed training on reader performance for CT colonography: multi-center study. Radiology 2007;242:152–161. [PubMed: 17185666]

13. Burling D, Halligan S, Atchley J, et al. CT colonography: interpretative performance in a non-academic environment. Clin Radiol 2007;62:424–429. discussion, 430–431. [PubMed: 17398266]

14. Fidler JL, Fletcher JG, Johnson CD, et al. Understanding interpretive errors in radiologists learning computed tomography colonography. Acad Radiol 2004;11:750–756. [PubMed: 15217592]

15. Spinzi G, Belloni G, Martegani A, Sangiovanni A, Del Favero C, Minoli G. Computed tomographic colonography and conventional colonoscopy for colon diseases: a prospective, blinded study. Am J Gastroenterol 2001;96:394–400. [PubMed: 11232681]

16. Thomeer M, Carbone I, Bosmans H, et al. Stool tagging applied in thin-slice multidetector computed tomography colonography. J Comput Assist Tomogr 2003;27:132–139. [PubMed: 12703001]

17. Slater A, Taylor SA, Tam E, et al. Reader error during CT colonography: causes and implications for training. Eur Radiol 2006;16:2275–2283. [PubMed: 16703308]

18. American College of Radiology. ACR Practice Guideline. Vol. 29. Reston, VA: American College of Radiology; 2005. ACR practice guideline for the performance of computed tomography (CT) colonography in adults; p. 295-298.

19. American College of Radiology. ACR Practice Guideline. Vol. 36. Reston, VA: American College of Radiology; 2009. ACR practice guideline for the performance of computed tomography (CT) colonography in adults; p. 1-10.

20. Dachman AH, Kelly KB, Zintsmaster MP, et al. Formative evaluation of standardized training for CT colonographic image interpretation by novice readers. Radiology 2008;249:167–177. [PubMed: 18796675]

21. Bodily KD, Fletcher JG, Engelby T, et al. Nonradiologists as second readers for intraluminal findings at CT colonography. Acad Radiol 2005;12:67–73. [PubMed: 15691727]

22. Zalis ME, Barish MA, Choi JR, et al. CT colonography reporting and data system: a consensus proposal. Radiology 2005;236:3–9. [PubMed: 15987959]

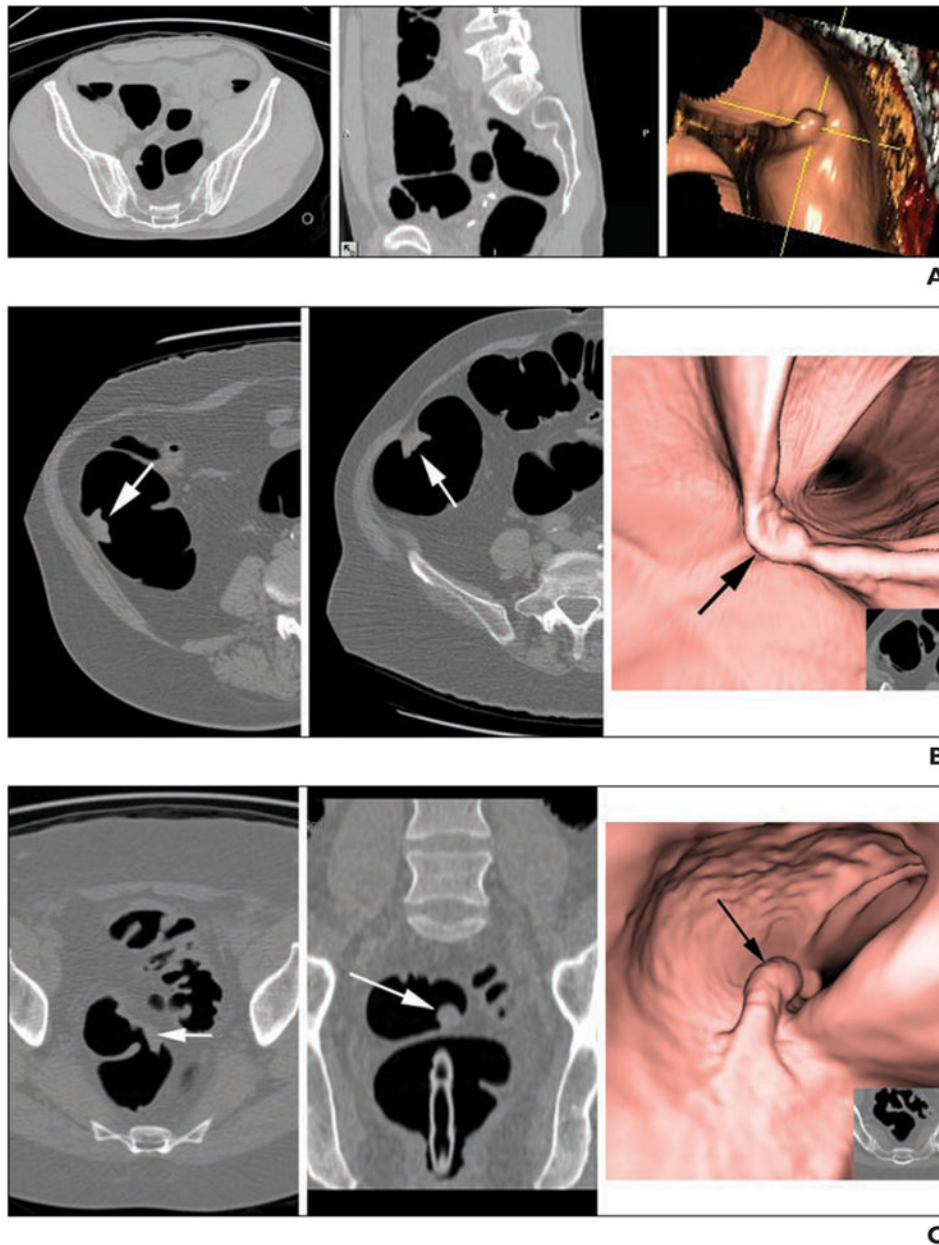23. National Radiology Data Registry. [Accessed July 22, 2009]. Website. https://nrdracrorg/portal/Nrdr/Main/pageaspx

Fig. 1.

Examples of easy-to-detect, moderate-to-detect, and difficult-to-detect polyps used at qualification testing.

A, CT colonography (CTC) example of easy-to-detect polyp in 58-year-old man. Axial (*left*), sagittal (*middle*), and 3D (*right*) views are shown. Imaging was performed with patient in prone position. Rectosigmoid colon was collapsed while patient was in supine position. This 8-mm tubular adenoma was detected by all radiologists.

B, CTC example of moderate-to-detect cancer (*arrows*) in 52-year-old man with proximal occlusive sigmoid cancer. Cancer is shown on axial prone (*left*) and supine (*middle* and *right*) CTC images. This ascending colon cancer growing along fold was missed at initial endoscopic assessment but was confirmed at surgery. At initial qualification testing, 80% (12/15) of radiologists identified this lesion.

C, CTC example of difficult-to-detect rectal polyp (*arrows*) in 50-year-old woman. Prone axial (*left*), prone coronal (*middle*), and 3D (*right*) views are shown. This polyp was detected by only 20% (3/15) of radiologists at qualification testing. Polyp was difficult to detect because only portion of lesion can be seen in any one view, with stalk lying along inferior rectal wall on prone images. Rectosigmoid was collapsed in supine view. Pathology showed 2.5-cm pedunculated tubular adenoma.
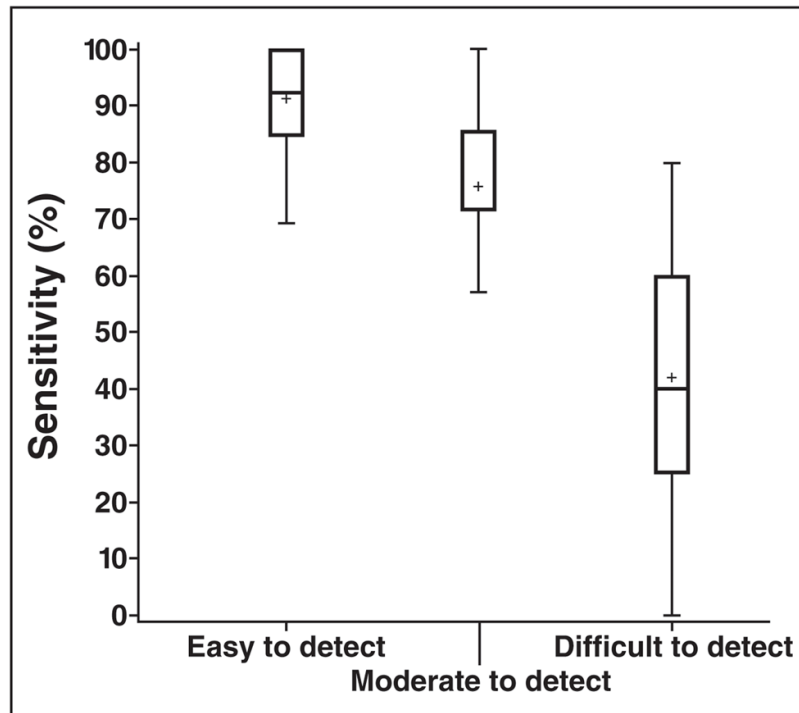
Fig. 2.
Box plots show sensitivity of 15 radiologists for easy-to-detect, moderate-to-detect, and difficult-to-detect polyps at initial qualification examination. Plus signs indicate mean value. Middle horizontal lines mark median. Top and bottom horizontal lines of boxes display interquartile range (IQR) from 25th to 75th percentiles. Whiskers are drawn to farthest observations within 1.5 IQR. Median for moderate-to-detect polyps is equal to 25th percentile because of tied values.

TABLE 1

Features of Colorectal Lesions Reviewed During Initial Day of Training

Case No.	Interpretive Techniques Used to Examine Each Set of CTC Cases	Teaching Point	No. of Cases
1–10 (Cancers, partial cases[a])	2D or 3D search, prone or supine, internal attenuation	Annular cancer	5
		Flat cancer or carpet lesion	4
		Large polypoid	1
		Shouldering	4
		Luminal constriction	4
		Position of large lesions can change because of gravity	2
11–27 (Polyps, partial cases[a])	2D or 3D search, prone or supine, internal attenuation	Sessile	9
		Pedunculated	2
		Flat	5
		Lipoma	1
		Long stalk	2
		Diminutive polyps	3
		On folds	5
		Rotation of colonic flexures with changes in position	3
		Position of large lesions can change because of gravity	3
		Irregular surface	2
		Submerged lesions (with tagging)	1
		One view only	5
28–42 (Full cases[b])	2D or 3D search, prone or supine, internal attenuation, 3D endoluminal fly-through and 2D search methods	Polypoid cancer	2
		Annular cancer	1
		Flat cancer	1
		Hemorrhoids	1
		Sessile polyps	14
		Pedunculated polyps	3
		Flat polyps	2
		Diminutive polyps	5
		On folds	4
		Long stalk	1
		Irregular surface	1
		One view only	3
		Submerged	3
		Lipoma	1
		Postoperative anastomoses	3

Note—CTC = CT colonography.

[a]Radiologists were told in which segments lesions were located. Primary 2D or 3D review of entire case not performed.

[b]Three negative cases and 12 with polyps or cancers.

TABLE 2

Breakdown of Characteristics of 25 Polyps by Conspicuity Category in Qualifying Examination Cases

	No. of Polyps		
Characteristic of Polyp	**Easy to Detect ($n = 13$)**	**Moderate to Detect ($n = 7$)**	**Difficult to Detect ($n = 5$)**
Size			
5–9 mm	5	4	1
≥ 10 mm	8	3	4
CT morphology			
Sessile	8	6	0
Pedunculated	3	0	2
Annular	2	0	0
Flat	0	1	3
Histology[a]			
Tubular adenoma	8	4	3
Tubulovillous adenoma	0	1	1
Hyperplastic polyp	1	0	0
Mixed tubular adenoma and hyperplastic polyp	0	0	1
Cancer	3	2	0

[a]One easy-to-detect polyp was seen at surgery but was not biopsied.

TABLE 3

Characteristics of Lesions Missed by 15 Radiologists Participating in the American College of Radiology Imaging Network 6664 Certifying Examination

Characteristic	No. of Lesions Showing Characteristic From Initial Test Set (*n* = 20 Cases)	% (No.) of Polyps With Characteristic Missed by ≥ 20% of Radiologists on Initial Test[a]	No. of Cases Showing Characteristic From Retraining Set (*n* = 30 Cases)
Sessile	12	33 (4/12)	20
Pedunculated	4	50 (2/4)	11
Flat	7	71 (5/7)	9
Annular	2	0 (0/2)	0
On fold	9	33 (3/9)	13
Suboptimal preparation	3	67 (2/3)	3
One view only[b]	7	43 (3/7)	4
Undulating surface	4	100 (4/4)	4
Size discrepancy[c]	2	50 (1/2)	14
Bulbous or thickened fold	3	100 (3/3)	2
Long stalk	2	50 (1/2)	0

[a] Eleven lesions were missed by at least one radiologist: six were ≥ 1 cm and five were 5–9 mm.

[b] Seen in only one position (i.e., supine or prone), but not the other.

[c] Lesion appeared smaller at CT than endoscopy using primary 2D read or vice versa.

TABLE 4

Breakdown of Reader Experience With CT Colonography and Participation in Focused Retraining

Prestudy Experience and Training	No. of Readers		
	Passed	Retrained	Total
Experienced (≥ 500 cases)	3	1	4
Inexperienced (< 500 cases)			
Two training sessions (45 cases)	2	7	9
One training session (15 cases)	2	0	2
Total	7	8	15

TABLE 5

Two-Sample Test Comparing Mean Sensitivity From First Examination for Readers Who Passed the First Examination Versus Those Who Did Not Pass Until the Second Examination and Corresponding Prospective Study Sensitivity for Those Two Groups

Reader Performance	Mean Sensitivity (SE)		*p*
	Initial Passers (*n* = 7)	Retrainees (*n* = 8)	
Qualification testing[a]			
Initial	0.85 (0.035)	0.71 (0.019)	0.004
Final	0.85 (0.035)	0.87 (0.016)[b]	0.612
Prospective study[c]			
Per patient	0.88 (0.054)	0.92 (0.036)	0.612
Per polyp or cancer	0.78 (0.069)	0.89 (0.054)	0.371

Note—SE = standard error.

[a]For qualification testing, sensitivity estimates include grading difficult-to-detect polyps and cancers ≥ 5 mm.

[b]Represents a significant improvement compared with initial qualification testing for these readers ($p < 0.001$).

[c]Prospective study results are reported per patient and per lesion for adenomatous polyps and cancers ≥ 1 cm [9].