

Andrew Slater
Stuart A. Taylor
Emily Tam
Louise Gartner
Julia Scarth
Chand Peiris
Arun Gupta
Michele Marshall
David Burling
Steve Halligan

Reader error during CT colonography: causes and implications for training

Received: 30 December 2005
Revised: 6 March 2006
Accepted: 14 April 2006
Published online: 16 May 2006
© Springer-Verlag 2006

A. Slater · E. Tam · L. Gartner · J. Scarth ·
C. Peiris · A. Gupta · M. Marshall ·
D. Burling
Intestinal Imaging,
St. Mark's and Northwick Park Hospitals,
Harrow,
London, UK

S. A. Taylor (✉) · S. Halligan
Department of Specialist Radiology,
Level 2 Podium,
University College Hospital,
235 Euston Road,
London, NW1 2BU, UK
e-mail: csyttaylor@yahoo.co.uk
Tel.: +44-207-380-9300
Fax: +44-207-388-2147

Abstract This study investigated the variability in baseline computed tomography colonography (CTC) performance using untrained readers by documenting sources of error to guide future training requirements. Twenty CTC endoscopically validated data sets containing 32 polyps were consensus read by three unblinded radiologists experienced in CTC, creating a reference standard. Six readers without prior CTC training [four residents and two board-certified subspecialty gastrointestinal (GI) radiologists] read the 20 cases. Readers drew a region of interest (ROI) around every area they considered a potential colonic lesion, even if subsequently dismissed, before creating a final report. Using this final report, reader ROIs were classified as true positive detections, true negatives

correctly dismissed, true detections incorrectly dismissed (i.e., classification error), or perceptual errors. Detection of polyps 1–5 mm, 6–9 mm, and ≥ 10 mm ranged from 7.1% to 28.6%, 16.7% to 41.7%, and 16.7% to 83.3%, respectively. There was no significant difference between polyp detection or false positives for the GI radiologists compared with residents ($p=0.67$, $p=0.4$ respectively). Most missed polyps were due to failure of detection rather than characterization (range 82–95%). Untrained reader performance is variable but generally poor. Most missed polyps are due perceptual error rather than characterization, suggesting basic training should focus heavily on lesion detection.

Keywords Colon · CT

Introduction

Computed tomography colonography (CTC) is increasing available, largely based on encouraging performance data from studies in academic centers [1–5]. Although the largest study to date produced excellent results [6], other comparative trials have been less encouraging [7–9]. Suboptimal technical parameters may in part explain poor performance, but much attention has focused on the influence of reader experience on accuracy [10, 11]. As expected, there is a clear learning curve associated with interpretation of CTC [12], and recent data confirms superior performance characteristics in expert hands [13, 14]. Reader training is therefore now generally considered mandatory [15], and appropriate guidelines have been

provided by expert consensus [16]. However, evidence on which to base such guidelines is relatively sparse. While response to training is known to be unpredictable [17], the reasons for this remain unclear. Fidler et al. demonstrated suboptimal reader performance even after a period of formal training and suggested errors in lesion detection (i.e., nonvisualization) were a significant source of missed lesions [18]. Such data questions the duration and emphasis of current training programs. For example, it is not known to what degree untrained readers have differing innate capability, perhaps enhanced by prior experience of gastrointestinal radiology; whether training should be individualized; and what should be the differential emphasis on lesion detection or characterization. The purpose of this study was to investigate variability in

baseline observer performance using untrained readers of varying GI experience and to identify sources of error to guide future training requirements.

Materials and methods

Case selection and annotation

Twenty CTC data sets with 32 endoscopically proven polyps were selected from a local research CTC data base. Full ethical permission for the study was obtained, and patients gave informed consent for use of their data. The data base had been collated from five donor institutions all running research programs comparing CT colonography with endoscopy. All patients received full bowel preparation [90 ml sodium phosphosoda (13 patients) and 24 g magnesium citrate (7 patients), and seven received intravenous spasmolytic. Thirteen had fecal tagging (500 ml 2.1% w/v barium sulphate plus 120 ml meglumine diatrizoate), and scans were performed using multidetector-row CT (4–16 rows). CT parameters were as follows: collimation 2.5 mm, 120 kV, supine 120–233 mA, prone 25–352 mA.

Cases were loaded onto a computer workstation equipped with CTC software (ColonCAR 1.2, Medicsight plc, London, UK). The software enabled a primary two-dimensional (2D) analysis of the prone and supine images simultaneously using the original axial images, together with both multiplanar reconstructions and a three-dimensional (3D) surface-shaded cube for problem solving. No endoluminal fly-through or automated center-line function was available. All 20 cases were then read by each of the three radiologists experienced in CTC interpretation (experience of at least 200 endoscopically confirmed CTC cases) blinded to the prevalence of abnormality over a period of 1 week. Each reader noted the size and slice number of all colonic neoplasia on a report form, which was then submitted to the study coordinator.

To establish the reference standard, the 20 data sets were then reread in consensus by the same three radiologists, with full knowledge of the same-day colonoscopy report for each individual patient. Readers confirmed technical adequacy (bowel preparation and distension) in all 20 cases and sought all polyps identified on colonoscopy using primarily 2D axial scrolling with 3D for problem solving. Disagreement was resolved by face-to-face discussion. Maximum transverse diameter was used to individually match polyps if the colonoscopy had found more than one polyp in an individual colonic segment. Once detected, a free-hand region of interest (ROI) was drawn around the polyp using a mouse and drawing tool embedded in the software. This ROI was then saved as a binary image file such that the exact polyp position could be recalled during subsequent case marking (Fig. 1).

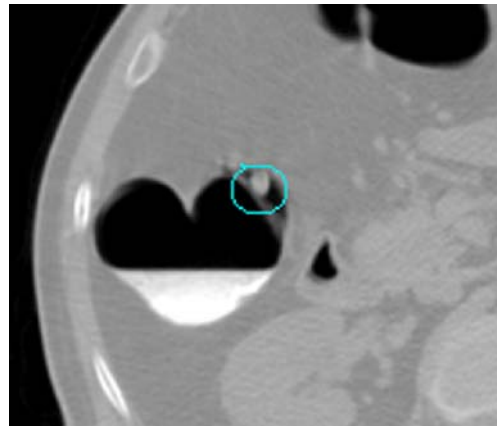


Fig. 1 Axial supine computed tomography colonographic (CTC) image showing method of polyp annotation via hand-drawn region of interest (ROI) stored as a binary image file

The three consensus readers then evaluated polyp morphology [sessile, pedunculated, or flat (less than 3 mm high)] and scored each colonic segment containing a polyp for fecal residue (based on percentage of wall circumference obscured: grade 1 0% obscured, grade 2 <25%, grade 3 25–50%, grade 4 >50%), residual fluid [based on anteroposterior (AP) diameter of colonic segment: 1 no fluid, 2 less than 25% AP diameter, 3 25–50% AP diameter, 4 >50% AP diameter], diverticular disease (absent or present), and whether the lesion was seen on one series only or both. Distension was also judged as clinically adequate or inadequate. For tagged data sets, note was made if the polyp was unsubmerged by tagged fluid on at least one data set. The study coordinator compared the blinded report forms for each of the three readers to the final annotated consensus reference standard in order to document baseline experienced reader performance for the 20 datasets.

Novice reader cohort

Six novice readers were selected to take part in the study. Four readers were residents with 1 year of experience in general abdominal CT examinations. Two readers were board-certified radiologists with a subspecialty interest in gastrointestinal radiology (with 5 and 15 years of experience, respectively, in general abdominal CT examinations). None of the readers had received any formal CTC training, and none were reporting CTC in day-to-day practice. Before the study, all readers were instructed to read a review article [19] to familiarize themselves with the basic concepts of CTC interpretation. Each reader was then shown the functionality of the viewing software (ColonCAR) during a 1-h tutorial by the study coordinator (experienced in over 100 CTC examinations) and were shown an example of a large polyp, a small polyp, and a

cancer. Functionality of the viewing software is described above.

All six readers were then instructed to independently read the 20 cases in their own time over a period of 4 weeks. Readers were blinded to the prevalence of abnormality within the data set and received no performance feedback over the course of the study. No attempt was made to record reporting times. For each case, readers were instructed to draw a free-hand ROI around every area they considered a potential colonic lesion even if they subsequently dismissed it following lesion classification. This requirement was stressed by the study coordinator, and readers were reassured that they could mark as many potentially abnormal areas as they wished. Each reader stored their ROIs, which were saved by the computer as a binary image file (so they could be recalled during case marking) before producing a final report following classification/characterization of each ROI listing the size, segment, and slice number of each lesion they would report in routine clinical practice. Readers were not instructed to ignore sub-5-mm lesions. This final report was handed to the study coordinator after each case was completed.

Case marking

The study coordinator positioned two workstations side by side. On the first, the annotated consensus reference was recalled (with true polyps circled) and on the second, the equivalent data set for each reader in turn (with saved ROIs displayed). Using the final reader written report, the coordinator classified each reader ROI as true positive (if it corresponded to a true polyp and included in the final report), false positive (not corresponding to a true polyp but was included in the final report), a false negative [error of classification, (if it corresponded to a true polyp but the

lesion was not included in the final report)]. A false negative (error of detection) was noted when no ROI was seen around a known lesion. If an ROI did not correspond to a true polyp and was not included in the final report, it was classified as “correctly dismissed”.

Analysis

Per-polyp sensitivity, both overall and according to size (1–5 mm, 6–9 mm, ≥ 10 mm), was calculated using the final reader report compared with the consensus reference standard. Detection rates, overall and according to polyp size, were compared across readers using the chi square test. Overall polyp detection for the GI subspecialist radiologist was compared with that of the residents using Fisher’s exact test. False positives and correctly dismissed ROIs were expressed as a mean per case. To investigate the influence of fecal residue, residual fluid, diverticular disease, and supine/prone visualization on detection, polyps were divided into two groups—those seen by three or more readers and those seen by two readers or less—and proportions were compared using Fisher’s exact test. For this analysis, fecal and fluid scores were divided into two groups: grades 1, 2 versus grades 3, 4.

Results

Polyp characteristics

Five data sets were normal. The other 15 cases contained between one and six lesions each. There were 14 polyps between 2 and 5 mm, 12 polyps between 6 and 9 mm, and six lesions ≥ 10 mm, one of which was a polypoid cancer measuring 30 mm. Polyps were distributed in the rectum ($n=9$), sigmoid colon ($n=9$), descending colon ($n=5$),

Table 1 Individual polyp detection for each reader according to size

Polyp size	Mean detection (experienced readers) ^a	GI radiologist (1) detected (%)	GI radiologist (2) detected (%)	Resident (1) detected (%)	Resident (2) detected (%)	Resident (3) detected (%)	Resident (4) detected (%)	Significance (<i>p</i>)*
1–5 mm ($n=14$)	6.3 (45%)	3 (21.4%)	1 (7.1%)	2 (14.3%)	3 (21.4%)	4 (28.6%)	4 (28.6%)	0.68
6–9 mm ($n=12$)	7.7 (64%)	2 (16.7%)	3 (25.0%)	5 (41.7%)	5 (41.7%)	5 (41.7%)	3 (25.0%)	0.64
≥ 10 mm ($n=6$)	6 (100%)	4 (66.7%)	1 (16.7%)	5 (83.3%)	4 (66.7%)	3 (50.0%)	4 (66.7%)	0.26
Overall ($n=32$)	23.3 (62.5%)	9 (28.1%)	5 (15.6%)	12 (37.5%)	12 (37.5%)	12 (37.5%)	11 (34.4%)	0.35

GI gastrointestinal

^aBased on individual performance of each of three experienced readers prior to unblinding to colonoscopy data and creation of reference standard

*Chi square test

transverse colon ($n=4$), ascending colon ($n=4$), and cecum ($n=1$). Four polyps were pedunculated, one was flat (7 mm), and the remaining 27 were sessile. Eight polyps were seen on one data set only; 24 were seen on both supine and prone data sets. Only two polyps had diverticular disease in the same segment. All but one polyp had no fecal residue in the segment. Segmental fluid residue was classified as grade 1 for three polyps, grade 2 for 14 polyps, grade 3 for 12 polyps, and grade 4 for three polyps. In cases that had fecal tagging, all were nonsubmerged on at least one view. All segments containing polyps were judged adequately distended for clinical purposes.

Reader performance

The mean detection for the three experienced readers (prior to the unblinded consensus read) for polyps sized 1–5 mm, 6–9 mm, and ≥ 10 mm was 45% (range 29–57%), 64% (range 50–75%), and 100% (range 100–100%), respectively. The six untrained readers highlighted an average 4.5 ROIs per case. Per polyp sensitivity for the six readers is shown in Table 1. Detection of polyps 1–5 mm, 6–9 mm, and ≥ 10 mm ranged from 7.1% to 28.6%, 16.7% to 41.7%, and 16.7% to 83.3%, respectively. There was no evidence of any interaction between individual readers and sensitivity either overall or when polyps were grouped according to size ($p>0.05$ in all cases). Furthermore, there was no significant difference between mean polyp detection for the GI radiologists compared with mean detection for residents overall ($p=0.67$) or for polyps ≥ 6 mm ($p=0.33$).

The 30-mm cancer was detected by all six readers but dismissed by one resident (Fig. 2). Only one polyp (6 mm) was seen by all six readers (Fig. 3). Conversely, one 10-mm polyp was not detected by any of the readers (Fig. 4).

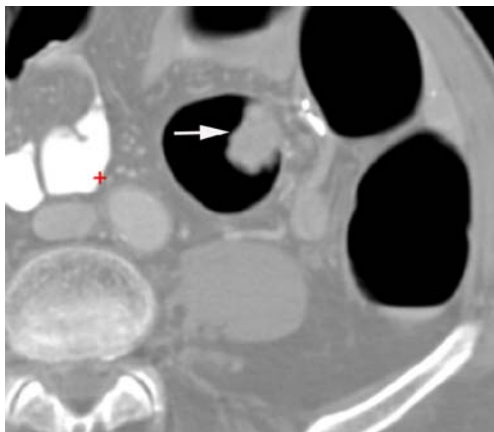


Fig. 2 Axial supine computed tomography colonographic (CTC) image demonstrating a 30-mm sigmoid cancer (*arrow*) detected but dismissed by one observer

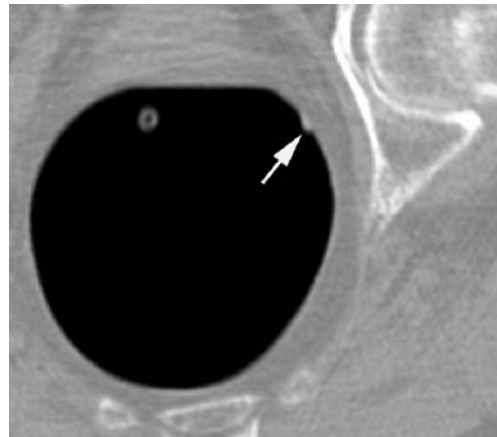


Fig. 3 Axial prone computed tomography colonographic (CTC) images demonstrating a 6-mm polyp (*arrow*) identified and correctly classified by all six observers

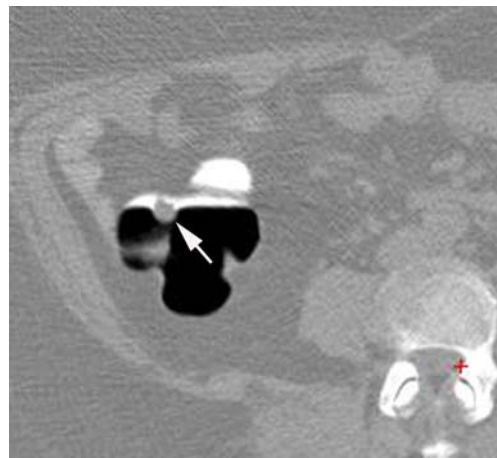


Fig. 4 Axial prone computed tomography colonographic (CTC) image demonstrates a 10-mm ascending colonic polyp (*arrow*) not detected by any of the six observers

Reader false negatives

Distribution of reader ROIs not corresponding to true polyps together with the reason for missed polyps is shown in Table 2 and the distribution of false negatives according to errors of detection or classification shown graphically in Figs. 5 and 6. For all readers, the majority of missed polyps were due to failure of detection rather than characterization, a pattern that held true for more clinically significant polyps (≥ 6 mm) (Fig. 6). For the four residents, the percentage of errors of detection were 95%, 90%, 90%, and 85%, and for the two GI radiologists, this percentage was 82% and 91% respectively. The second GI radiologist had slightly more errors of detection (40% of missed polyps ≥ 6 mm in size) compared with the other readers. Interestingly, this reader also had the lowest sensitivity for polyps ≥ 10 mm.

Table 2 Cause of reader false negatives and distribution of regions of interest (ROIs) not corresponding to true positive detections

ROI ^a	GI radiologist (1)	GI Radiologist (2)	Resident (1)	Resident (2)	Resident (3)	Resident (4)
False negative (error of detection)	21	22	17	18	18	20
False negative (error of characterization)	2	5	3	2	2	1
Correctly dismissed	55	100	73	80	36	56
False positive	7	11	10	15	7	17

ROI region of interest, GI gastrointestinal

^aFor all 20 data sets combined

Residents identified, and then correctly dismissed, a mean of 3.1 ROIs per case compared with 3.9 for the GI radiologists ($p=0.8$). The mean number of false-positive detections per case was 0.61 for residents and 0.45 for GI radiologists ($p=0.4$). Per-patient specificity (based on the five normal cases) was 22% and 38% for the two GI radiologists and 38%, 27%, 13%, and 17% for the four residents, respectively.

Data set characteristics

There was no significant interaction between the presence of fecal residue ($p=0.99$), residual fluid ($p=0.99$), fecal tagging ($p=0.99$), visualization on both supine/prone data sets ($p=0.7$), or presence of diverticular disease ($p=0.99$) on polyp detection. No reader detected the 7-mm flat polyp, but at least three detected all the pedunculated polyps.

Discussion

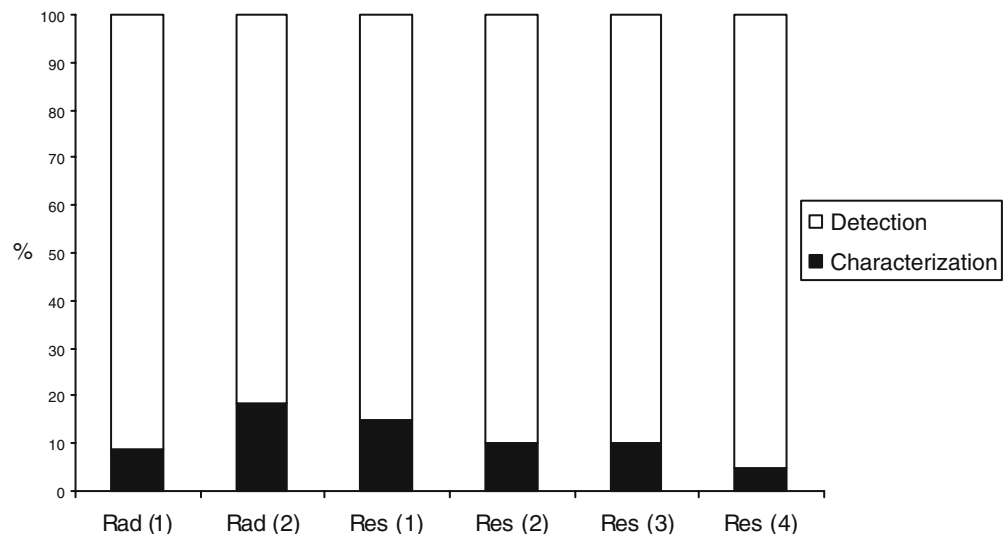
Recent disappointing results from multicenter trials [8, 9] have focused attention on reader training prior to interpretation of CTC [15]. The need for structured tuition using endoscopically validated cases is now generally accepted

throughout the CTC community although there remains little evidence on which to base the length and structure of training courses. Recent expert consensus opinion recommends training should take place over 2 days, with review of a least 40 cases that contain mixed pathology [16]. However, it is known that response to training is variable, with some individuals actually declining in performance after 50 cases [17]. Clearly, quality as well as quantity of training is therefore important, and our study was designed to investigate performance in untrained readers in an attempt to determine basic training requirements.

As expected, we found that overall untrained reader sensitivity was poor, but there was also quite marked individual variation. All readers volunteered for the study and had expressed interest in learning CTC, suggesting reader motivation did not skew the results. It also could be argued that we utilized a relatively difficult data set (mean performance of the three experienced radiologists was toward the lower limits of the published literature). However, most polyps >6 mm and all ≥ 10 mm were correctly located by the experienced radiologists during their blinded reads, and all six novice readers performed uniformly less well.

There was no evidence that experience in general GI radiology held any advantage prior to training in contrast to previous work. For example, one GI radiologist detected

Fig. 5 Distribution of reader false negatives classified according to errors of detection and characterization. Rad radiologist; Res resident



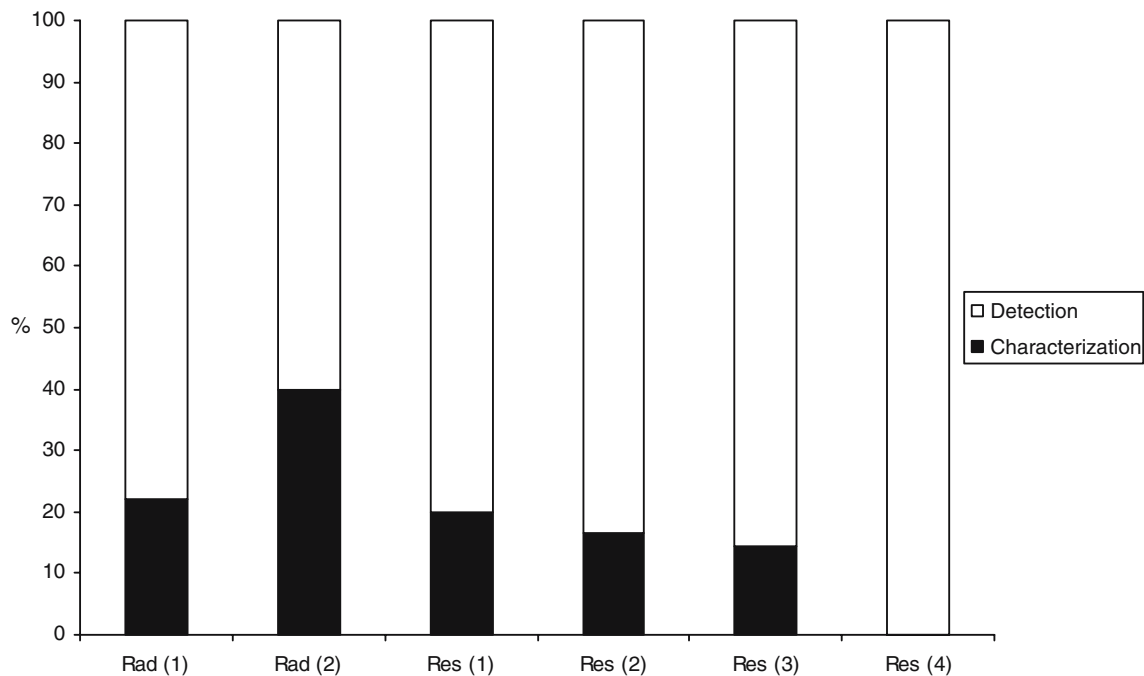


Fig. 6 Distribution of reader false negatives for polyps ≥ 6 mm, classified according to the proportion of errors of detection and characterization. *Rad* radiologist; *Res* resident

just 17% of polyps ≥ 10 mm compared with 83% for one resident. Data clearly suggest there is an innate individual capability to read CTC data sets, and this might outweigh subspecialty experience. Indeed, there was variation in the performance of the three experienced radiologists. CTC interpretation is unlike “normal” CT. When utilizing a primary 2D approach, there is a need to carefully scroll through hundreds of images, essentially focused on a black tube, something that can be time consuming and tedious. These circumstances may well suit certain individuals more than others. Indeed, previous studies comparing resident CT performance have shown very good agreement with experienced specialist radiologists for studies as diverse as abdominal trauma [20], appendicitis [21] and CT pulmonary angiography [22]. Seemingly, the same cannot be assumed for CTC.

Errors committed in CTC can be classified as failure to detect a lesion or failure to characterize a detected lesion (i.e., incorrectly dismissing a polyp as colonic fold or retained feces). We found the overwhelming majority of missed polyps were errors of detection; i.e., lesions were simply not seen. Using 15 experienced gastrointestinal radiologists, Fidler et al. demonstrated that even after training, detection errors still accounted for 55% of missed polyps [18]. The available evidence clearly suggests simple perceptual deficiencies are a serious problem in CTC. It therefore follows that initial training should focus heavily on detection technique rather than stressing lesion characterization. In keeping with most CTC data, we found specificity (in terms of false positives) was

considerably less of a problem than sensitivity. Indeed, simple rules regarding lesion characterization (lesion homogeneity, supine/prone correlation, etc.) are probably easily learned by most radiologists. Emphasis should therefore be on finding potential lesions within the colon, something that may take more than a single 2-day training course to achieve. It may also be sensible to start training using with well-cleansed and distended data sets, but clearly, diagnostic pitfalls such as retained fecal residue and poor distension must be well taught. We found no difference in reader performance using data sets with and without fecal tagging (albeit with a relatively small number of data sets utilizing a large volume of meglumine diatrizoate). The use of tagging may increase polyp conspicuity by providing greater contrast between the lesion and surrounding colonic contents. However, in our experience, interrogation of tagged data sets requires new interpretative skills (e.g., it can be easy to dismiss a real polyp because of oral contrast coating). Readers of CTC must become familiar with both tagged and nontagged data sets although when in any training program tagged data sets should be added remains unclear. It would be interesting to investigate the influence of fecal tagging on novice reader performance using a larger number of data sets.

Circulation of anonymized annotated data sets (either by mail or via online Web sites) to be reviewed prior to course attendance may help improve basic interpretation, leaving face-to-face tuition to focus on those aspects best demonstrated in a workshop environment. Interestingly, we found one reader (with the lowest sensitivity for ≥ 10 mm lesions)

had proportionally more errors of characterization than the other readers, suggesting a slightly different training emphasis would be required. Whether a test set of cases should be sent to course participants prior to attendance to assess baseline performance and perhaps allow more individualized hands-on training is an interesting proposition and may be increasingly practical with Web-based learning forums.

If CTC is to remain robust once disseminated, this problem of perceptual misses must be addressed both during training and in subsequent day-to-day clinical practice. One possible solution is use of primary 3D interpretation. Our study used primary 2D, which, although at present is likely more time efficient than 3D, may nevertheless be more challenging to untrained readers. Pickhardt et al. have strongly advocated use of primary 3D interpretation [23], and recent abstract data suggests improved performance compared with 2D for untrained readers [24]. It does, however, seem precipitant to suggest primary 2D read should be abandoned in future training programs. Basic 2D skills are essential for problem solving of 3D findings, and indeed, CTC examinations performed with reduced bowel preparation [25–27] may well demand primary 2D analysis. It therefore seems mandatory that novice readers must be trained in both 2D and 3D analysis, and perhaps both methods should be used in full as the technique is being learnt.

Double reporting may also improve CTC performance. This has obvious manpower implications, but there is increasing evidence that nonradiologists can be trained to adequate standards, potentially acting as the second reader [28]. Radiographic technicians performing the CTC studies may be best placed to fulfill this role, providing adequate supervision by a radiologist experienced in CTC.

Finally, increasingly reliable computer-aided detection (CAD) systems are now becoming available [29–33]. Preliminary data suggests a combination of 2D with CAD has equivalent sensitivity to primary 3D interpretation [34],

and workstations combining CAD with 3D rendering are imminent at the time of writing. The high prevalence of detection errors in both novice and trained readers makes CAD an intuitively attractive proposition although studies examining the actual impact on reader performance are in their infancy.

Our study does have weaknesses. We deliberately used both tagged and nontagged data sets. Although we found no evidence that tagging influenced reader performance (even when combining the six readers), it is possible the smaller number of data sets used could mask an effect. We used two GI radiologists and four residents, and these readers may not be representative of their respective groups in the wider radiological community. We did not investigate reading times, which was a deliberate decision made so readers did not feel time pressured, which may have skewed their underlying performance. It is clear that readers must spend adequate time reviewing data sets when learning the technique. We cannot be sure readers did not see polyps yet fail to circle them (i.e., incorrectly producing errors of detection rather than characterization). However, all reader ROIs were saved by the software, and readers were under strict instructions to record every potential abnormality. Unlike Gluecker et al. [35], we found no influence of data set quality on polyp detection. This may be because the study was underpowered to detect such differences, but perhaps more importantly, we chose well-cleaned and distended cases for our test data set. The purpose of the study was not really to address the number of data sets required for competent performance, and this area requires further study.

In conclusion, untrained readers perform variably when asked to interpret CTC, but overall performance is generally poor. Previous experience in general CT confers no advantage. Most missed polyps are due to errors in detection, suggesting basic training should focus heavily on lesion detection rather than characterization.

References

1. Fenlon HM, Nunes DP, Schroy PC III, Barish MA, Clarke PD, Ferrucci JT (1999) A comparison of virtual and conventional colonoscopy for the detection of colorectal polyps. *N Engl J Med* 341:1496–1503
2. Yee J, Akerkar GA, Hung RK, Steinauer-Gebauer AM, Wall SD, McQuaid KR (2001) Colorectal neoplasia: performance characteristics of CT colonography for detection in 300 patients. *Radiology* 219:685–692
3. Macari M, Bini EJ, Jacobs SL, Naik S, Lui YW, Milano A, Rajapaksa R, Megibow AJ, Babb J (2004) Colorectal polyps and cancers in asymptomatic average-risk patients: evaluation with CT colonography. *Radiology* 230:629–636
4. Dachman AH, Kuniyoshi JK, Boyle CM, Samara Y, Hoffmann KR, Rubin DT, Hanan I (1998) CT colonography with three-dimensional problem solving for detection of colonic polyps. *AJR Am J Roentgenol* 171:989–995

5. Iannaccone R, Laghi A, Catalano C, Brink JA, Mangiapane F, Trenna S, Piacentini F, Passariello R (2003) Detection of colorectal lesions: lower-dose multi-detector row helical CT colonography compared with conventional colonoscopy. *Radiology* 229:775–781
6. Pickhardt PJ, Choi JR, Hwang I, Butler JA, Puckett ML, Hildebrandt HA, Wong RK, Nugent PA, Mysliwiec PA, Schindler WR (2003) Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *N Engl J Med* 349:2191–2200
7. Johnson CD, Harmsen WS, Wilson LA, MacCarty RL, Welch TJ, Ilstrup DM, Ahlquist DA (2003) Prospective blinded evaluation of computed tomographic colonography for screen detection of colorectal polyps. *Gastroenterology* 125:311–319
8. Rockey DC, Paulson E, Niedzwiecki D, Davis W, Bosworth HB, Sanders L, Yee J, Henderson J, Hatten P, Burdick S, Sanyal A, Rubin DT, Sterling M, Akerkar G, Bhutani MS, Binmoeller K, Garvie J, Bini EJ, McQuaid K, Foster WL, Thompson WM, Dachman A, Halvorsen R (2005) Analysis of air contrast barium enema, computed tomographic colonography, and colonoscopy: prospective comparison. *Lancet* 365:305–311
9. Cotton PB, Durkalski VL, Pineau BC, Palesch YY, Mauldin PD, Hoffman B, Vining DJ, Small WC, Affronti J, Rex D, Kopecky KK, Ackerman S, Burdick JS, Brewington C, Turner MA, Zfass A, Wright AR, Iyer RB, Lynch P, Sivak MV, Butler H (2004) Computed tomographic colonography (virtual colonoscopy): a multicenter comparison with standard colonoscopy for detection of colorectal neoplasia. *JAMA* 291:1713–1719
10. Halligan S, Taylor S, Burling D (2004) Virtual colonoscopy. *JAMA* 292:432
11. Ferrucci JT (2005) Colonoscopy: virtual and optical—another look, another view. *Radiology* 235:13–16
12. Spinzi G, Belloni G, Martegani A, Sangiovanni A, Del Favero C, Minoli G (2001) Computed tomographic colonography and conventional colonoscopy for colon diseases: a prospective, blinded study. *Am J Gastroenterol* 96:394–400
13. Burling D, Halligan S, Atchley J, Dhingsa R, Maskell G, Guest P, Haywood S, Higginson A, Jobling C (2005) CT colonography Interpretative performance outside a research environment. *Eur Radiol* 13(Suppl 3):15
14. Johnson CD, Toledano AY, Herman BA, Dachman AH, McFarland EG, Barish MA, Brink JA, Ernst RD, Fletcher JG, Halvorsen RA Jr, Hara AK, Hopper KD, Koehler RE, Lu DS, Macari M, MacCarty RL, Miller FH, Morrin M, Paulson EK, Yee J, Zalis M (2003) Computerized tomographic colonography: performance evaluation in a retrospective multicenter setting. *Gastroenterology* 125:688–695
15. Soto JA, Barish MA, Yee J (2005) Reader training in CT colonography: how much is enough? *Radiology* 237:26–27
16. Soto J, Barish M, Ferrucci J. CT Colonography Interpretation: Guidelines for Training Courses (2004) Radiological Society of North America scientific assembly and annual meeting program. Oak Brook, Ill: Radiological Society of North America 618
17. Taylor SA, Halligan S, Burling D, Morley S, Bassett P, Atkin W, Bartram CI (2004) CT colonography: effect of experience and training on reader performance. *Eur Radiol* 14:1025–1033
18. Fidler JL, Fletcher JG, Johnson CD, Huprich JE, Barlow JM, Earnest F, Bartholmai BJ (2004) Understanding interpretive errors in radiologists learning computed tomography colonography. *Acad Radiol* 11:750–756
19. Macari M, Bini EJ, Jacobs SL, Lange N, Lui YW (2003) Filling defects at CT colonography: pseudo- and diminutive lesions (the good), polyps (the bad), flat lesions, masses, and carcinomas (the ugly). *Radiographics* 23:1073–1091
20. Carney E, Kempf J, DeCarvalho V, Yudd A, Noshier J (2003) Preliminary interpretations of after-hours CT and sonography by radiology residents versus final interpretations by body imaging radiologists at a level 1 trauma center. *AJR Am J Roentgenol* 181:367–373
21. Albano MC, Ross GW, Ditchek JJ, Duke GL, Teeger S, Sostman HD, Flomenbaum N, Seifert C, Brill PW (2001) Resident interpretation of emergency CT scans in the evaluation of acute appendicitis. *Acad Radiol* 8:915–918
22. Ginsberg MS, King V, Panicek DM (2004) Comparison of interpretations of CT angiograms in the evaluation of suspected pulmonary embolism by on-call radiology fellows and subsequently by radiology faculty. *AJR Am J Roentgenol* 182:61–66
23. Pickhardt PJ (2003) Three-dimensional endoluminal CT colonography (virtual colonoscopy): comparison of three commercially available systems. *AJR Am J Roentgenol* 181:1599–1606
24. Iannaccone R, Catalano C, Marin D, Guerrisi A, De Filippis G, Passariello R (2005) Primary two-dimensional (2D) versus primary three-dimensional (3D) reading: performance of abdominal radiologists with limited training in CT colonography. In: Radiological Society of North America scientific assembly and annual meeting program. Oak Brook, Ill: Radiological Society of North America: 186
25. Lefere P, Gryspeerdt S, Marrannes J, Baekelandt M, Van HB (2005) CT colonography after fecal tagging with a reduced cathartic cleansing and a reduced volume of barium. *AJR Am J Roentgenol* 184:1836–1842
26. Lefere PA, Gryspeerdt SS, Dewyspelaere J, Baekelandt M, Van Holsbeeck BG (2002) Dietary fecal tagging as a cleansing method before CT colonography: initial results polyp detection and patient acceptance. *Radiology* 224:393–403
27. Iannaccone R, Laghi A, Catalano C, Mangiapane F, Lamazza A, Schillaci A, Sinibaldi G, Murakami T, Sammartino P, Hori M, Piacentini F, Nofroni I, Stipa V, Passariello R (2004) Computed tomographic colonography without cathartic preparation for the detection of colorectal polyps. *Gastroenterology* 127:1300–1311
28. Bodily KD, Fletcher JG, Engelby T, Percival M, Christensen JA, Young B, Krych AJ, Vander K, Rodysill D, Fidler JL, Johnson CD (2005) Nonradiologists as second readers for intraluminal findings at CT colonography. *Acad Radiol* 12:67–73
29. Taylor SA, Halligan S, Burling D, Roddie ME, Honeyfield L, McQuillan J, Amin H, Dehmshki J (2006) Computer-assisted reader software versus expert reviewers for polyp detection on CT colonography. *AJR Am J Roentgenol* 186:696–702
30. Summers RM, Yao J, Pickhardt PJ, Franzaszek M, Bitter I, Brickman D, Krishna V, Choi R (2005) Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. *Gastroenterology* 129:1832–1844
31. Graser A, Bogoni L, Becker CR, Reiser M (2005) Computer-aided detection (CAD) in MDCT colonography: evaluation of the performance of a prototype system in more than 100 cases. In: Radiological Society of North America scientific assembly and annual meeting program. Oak Brook, Ill: Radiological Society of North America: 336

-
32. Halligan S, Dehmeshki J, Taylor SA, Amin H, Xujiong Y, Tsang J (2005) External validation of computer-assisted reading for CT colonography. In: Radiological Society of North America scientific assembly and annual meeting program. Oak Brook, Ill: Radiological Society of North America:336
 33. Luboldt W, Mann C, Tryon CL, Vonthein R, Stueker D, Kroll M, Luz O, Claussen CD, Vogl TJ (2002) Computer-aided diagnosis in contrast-enhanced CT colonography: an approach based on contrast. *Eur Radiol* 12:2236–2241
 34. Taylor SA, Halligan S, Slater A, Goh V, Burling D, Roddie ME, Honeyfield L, McQuillan J, Amin H, Dehmeshki J (2006) Polyp detection using CT colonography: comparison of primary 3D endoluminal analysis with primary 2D axial read supplemented by computer assisted reader (CAR) software. *Radiology* (in press)
 35. Gluecker TM, Fletcher JG, Welch TJ, MacCarty RL, Harmsen WS, Harrington JR, Ilstrup D, Wilson LA, Corcoran KE, Johnson CD (2004) Characterization of lesions missed on interpretation of CT colonography using a 2D search method. *AJR Am J Roentgenol* 182:881–889