Radiology

Ronald M. Summers, MD, PhD
Anna K. Jerebko, PhD
Marek Franaszek, PhD
James D. Malley, PhD
C. Daniel Johnson, MD

Index terms:
Colon, CT, 75.12115, 75.12117
Colon neoplasms, 75.311
Colonoscopy, 75.1299
Computers, diagnostic aid

Published online before print
10.1148/radiol.2252011619
Radiology 2002; 225:391–399

Abbreviations:
CAD = computer-aided detection
FROC = free-response receiver operating characteristic

[1] From the Department of Diagnostic Radiology, Warren Grant Magnuson Clinical Center, National Institutes of Health, 10 Center Dr, MSC 1182, Bldg 10, Rm 1C660, Bethesda, MD 20892-1182 (R.M.S., A.K.J., M.F., J.D.M.); and Department of Radiology, Mayo Clinic, Rochester, Minn (C.D.J.). Received October 2, 2001; revision requested December 18; revision received February 4, 2002; accepted March 14. Supported by the intramural research programs of the Department of Diagnostic Radiology, Warren Grant Magnuson Clinical Center. Supported in part by NIH grant RO1CA75333. **Address correspondence to** R.M.S. (e-mail: *rms@nih.gov*).

R.M.S., J.D.M., and C.D.J. have patents pending and/or have been awarded patents for the subject matter described in the manuscript.

Author contributions:
Guarantors of integrity of entire study, R.M.S., C.D.J.; study concepts, R.M.S., C.D.J.; study design, R.M.S.; literature research, R.M.S.; clinical studies, C.D.J.; experimental studies, R.M.S.; data acquisition, C.D.J.; data analysis/interpretation, R.M.S., A.K.J., M.F.; statistical analysis, R.M.S., J.D.M.; manuscript preparation, R.M.S.; manuscript definition of intellectual content and editing, R.M.S., C.D.J.; manuscript revision/review and final version approval, all authors.

Colonic Polyps: Complementary Role of Computer-aided Detection in CT Colonography[1]

PURPOSE: To apply a computer-aided detection (CAD) algorithm to supine and prone multisection helical computed tomographic (CT) colonographic images to confirm if there is any added benefit provided by CAD over that of standard clinical interpretation.

MATERIALS AND METHODS: CT colonography (with patients in both supine and prone positions) was performed with a multisection helical CT scanner in 40 asymptomatic high-risk patients. There were two consecutive series of patients, 20 of whom had at least one polyp 1.0 cm in size or larger and 20 of whom had normal colons at conventional colonoscopy performed the same day. The CT colonographic images were interpreted with an automated CAD algorithm and by two radiologists who were blinded to colonoscopy findings.

RESULTS: For 25 polyps at least 1.0 cm in size ("large" polyps), sensitivity for detection by at least one radiologist was 48% (12 of 25). The sensitivity of CAD for detecting large polyps was also 48% (12 of 25), but the CAD algorithm detected four of 13 large polyps that were not detected by either radiologist (31%, 95% two-sided CI: 9, 61), increasing the potential sensitivity to 64% (16 of 25). For polyps identifiable retrospectively, sensitivity of CAD was 67% (12 of 18), and sensitivity of the combination of detection with the CAD algorithm or by at least one radiologist was 89% (16 of 18). There were an average of 11 false-positive detections per patient for CAD.

CONCLUSION: In this series of patients in whom radiologists had difficulties detecting polyps (compared with sensitivities of 75%–90% reported in the literature), this CAD algorithm played a complementary role to conventional interpretation of CT colonographic images by detecting a number of large polyps missed by trained observers.

Ongoing research at a number of academic centers supports the notion that computed tomographic (CT) colonography is a sensitive and specific method for detecting colonic polyps and cancers (1–4). Although the results are promising, concerns exist as to whether CT colonography will be equally effective when placed into general use. Two areas of concern are interobserver variability and excessive image interpretation time. For example, to compensate for interobserver variability, researchers in the larger clinical trials have reported results from consensus readings to boost sensitivity (1,3). In routine clinical practice, however, it is likely that consensus reading will be the exception rather than the rule. In addition, interpretation times of 10–60 minutes have been reported. Computer-aided detection (CAD) is a potential solution to these concerns (5–7). The purpose of this study was to apply our CAD algorithm to images obtained by means of CT colonography performed with a multisection helical CT scanner, with patients in both supine and prone positions, to determine if there is any added benefit provided by CAD over that provided by standard clinical interpretation.

MATERIALS AND METHODS

Patient Population

The study cohort (19 men, 21 women; mean age, 63 years; age range, 51–75 years) consisted of two groups of asymptomatic patients at high risk for colorectal cancer. The first group consisted of a consecutive series of 20 patients who had at least one polyp 1.0 cm or larger at conventional colonoscopy performed the same day that CT colonography was performed. The second group consisted of a consecutive series of 20 patients who had normal results at colonoscopy performed the same day that CT colonography was performed. Both groups were selected from a larger group of asymptomatic high-risk patients who underwent 719 CT colonographic examinations. During the period of data collection required to achieve consecutiveness for this study, 410 CT colonographic examinations had to be performed before the 20 patients who had polyps 1.0 cm or larger were identified. *High risk* indicates that a patient had a family or personal history of colorectal cancer; *high-risk patients* therefore includes patients in a surveillance population.

Patients were excluded if they had melena or hematochezia (ie, symptoms of colorectal cancer), inflammatory bowel disease, a hereditary polyposis syndrome, or a large amount of retained stool or residual colonic fluid. All patients underwent complete colonoscopy to the cecum. This study was approved by the institutional review board, and informed consent was obtained from the patients.

Six patients (four of whom had polyps) had previously undergone partial colonic resection for colon cancer. Ten patients (four of whom had polyps) had diverticulosis.

CT Scanning

Each patient ingested 1 gallon of a standard oral colonoscopy preparation (polyethylene glycol electrolyte solution, Colyte; Reed and Carnrick, Jersey City, NJ) and two 5-mg tablets of bisacodyl (Dulcolax; CIBA Consumer, Edison, NJ). The colon was insufflated with CO_2 to the limit of patient tolerance. All patients received 1 mg of glucagon (Eli Lilly, Indianapolis, Ind), which was administered subcutaneously 10 minutes prior to the examination.

CT scanning was performed with a Lightspeed QX/i (72 scans) or Lightspeed Plus (six scans) multisection helical CT scanner (GE Medical Systems, Milwaukee, Wis) in all patients except one (two scans), who was imaged with a single-section helical CT scanner (HiSpeed Advantage; GE Medical Systems) because the multisection scanner was unavailable. This patient's data were kept in the study data set to maintain consecutiveness.

CT scanning parameters used were 120 kVp, 50 mAs (mean), field of view to fit (38–46 cm), 5-mm collimation, HQ mode, and a 3-mm reconstruction interval with a 2-mm overlap (8). The protocol required only one 20-second breath hold for all sequences except one, which was performed with overlapping sections and multiple breath holds. Patients were imaged in both prone and supine positions. Scanning with the patient in the supine position was performed first, followed by scanning with the patient in the prone position. The size of the typical CT colonographic data set (supine or prone) was approximately 80 megabytes (160 images each). Therefore, there were a total of approximately 320 images per patient.

Colonoscopy

Conventional colonoscopy was performed after CT colonography on the same day. Colonoscopies were performed by experienced colonoscopists with more than 3 years of training. The colonoscopists were not aware of the CT colonography results. The size and location of any polyps were identified in the colonoscopy report. Polyp sizes were determined by the colonoscopist at the time of the examination; a probe or forceps was used for reference. The pull-back of the colonoscope was videotaped for later review. The sizes of 11 polyps were given qualitatively (eg, "diminutive" or "large") in the colonoscopy reports. The quantitative sizes of these polyps were determined, on the basis of our best estimates, from review of the pathology report or videotape. For purposes of brevity, polyps 1.0 cm or larger and polyps smaller than 1.0 cm are hereinafter referred to as *large* and *small* polyps, respectively.

Human Observers

The CT colonographic images obtained with the patients in the supine position were interpreted independently on the day of the examination by two of three radiologists (C.D.J., R.L.M., T.J.W., randomly assigned) who were blinded to the results of conventional colonoscopy. The observers recorded the presence and location of lesions they suspected to be polyps. All three radiologists were board certified and had at least 10 years of experience. They each had interpreted images from a minimum of 50 CT colonographic examinations before they interpreted the images in the present study. In this study two radiologists evaluated the images because an evaluation of interobserver variability was contemplated and so that we could assess whether double reading was necessary.

In contrast to our previous study (6), in which there was a strong bias that patients had polyps, in this study the radiologists knew that the patients were asymptomatic (ie, were a screening population) and probably had a lower prevalence of disease. The CT colonography examinations were chosen from among those performed in a cohort of patients in whom imaging had predominantly yielded normal results. Approximately 8% of this patient group was found to have colorectal polyps 1.0 cm or larger. Therefore, the radiologists had a strong bias toward the idea that the patients had normal colons.

The radiologists used research software developed at the Mayo Clinic for image display and interpretation (9). This software has been used in multiple published studies, including a recent large clinical trial (2). Preliminary evidence suggests that the performance of this software is comparable to that of commercial software (10). The primary mode of interpretation was analysis of two-dimensional transverse CT colonographic images, supplemented when needed with analysis of three-dimensional perspective endoluminal displays and reformatted two-dimensional images. The average interpretation time was approximately 15 minutes per case.

Matching Polyps at CT Colonography and Colonoscopy

An important step in determining the sensitivity of polyp detection by the computer algorithm was to locate precisely each polyp found at colonoscopy on the CT colonographic images. To make this assessment, two radiologists (R.M.S. and C.D.J.) evaluated transverse source images, endoluminal three-dimensional images, and, when needed, reformatted coronal and sagittal images. If more than one polyp was present in a segment, we used reported sizes to determine which was which to the best of our abilities. Once a match was made be-

tween a polyp described in the colonoscopy report and one visible in retrospect on CT colonographic images, the coordinates (ie, row, column, and section) of the center of the polyp as it appeared on the CT colonographic image were recorded in a computer database (Microsoft Access 2000; Microsoft, Redmond, Wash).

If the radiologists (R.M.S. and C.D.J.) could not find the polyp on CT colonographic images and match it to the colonoscopy findings despite a diligent search, the CAD software interpretation was assigned a false-negative result. Such false-negative results reduced the maximum sensitivity achievable with the CAD algorithm even before the algorithm was applied.

Computer-aided Polyp Detection Algorithm

We transferred the CT colonographic images to a personal computer (Dell Precision 620 Workstation; Dell, Austin, Tex) and analyzed the images with our computer-aided polyp detection software package (5,6,11–14). In brief, the CAD algorithm operates by identifying voxels along the wall of the colon and measuring the shape and CT numbers of the wall to classify the wall locally into polypoid and nonpolypoid (ie, normal) areas. The software analyzed CT colonographic data sets (supine or prone) at the rate of one every 2 minutes (4 minutes total for each set of supine and prone CT colonographic images).

On the basis of results of our earlier study (6), a window width (1,050 HU) and level (−475 HU) for converting the data from 12 to 8 bits, a region-growing threshold for seeding the colonic lumen (−475 HU), and a threshold for generating the isosurface (−800 HU) were chosen. Region-growing and isosurface extraction were performed to identify the wall of the colon.

Polyp detection was performed with a prototype automated polyp detector software with criteria ("filter 7") developed in an earlier study (6). Filter 7 was chosen on the basis of its high sensitivity and relatively low number of false-positive results per colon. The specifications for filter 7 are as follows: elliptic curvature of the peak subtype, a mean curvature range of −4.0 to −0.5 cm^{-1}, 10 or more vertices, a diameter greater than or equal to 0.5 cm, and sphericity less than or equal to one. Curvature is a measure of shape, elliptic curvature is a polypoid shape, vertices and diameter are measures of size, and sphericity is a measure of roundness.

In a previous study, filter 7 had 79% sensitivity for detection of polyps 1.0 cm or larger in well-distended segments of the colon (6). The patient population was different in that study: Most patients were known or highly suspected to have a polyp 1.0 cm or larger on the basis of results of barium enema examination or flexible sigmoidoscopy; there was therefore a higher pretest probability of disease.

A polyp was considered to be "detected" by the computer algorithm when the center of a computer-detected polyp was within 1 cm of the center of a polyp as recorded in the database, thereby matching a polyp in the colonoscopy report. Perspective volume-rendered endoluminal views of the polyps were generated by custom research software designed at our institution (12,15).

Polyp Densitometry

A further reduction in the number of false-positive results was achievable by sampling the CT numbers of each voxel within a possible polyp along a ray directed through the polyp. The purpose of this method was to identify voxels within the possible polyp that had soft-tissue attenuation. If any voxel along the ray exceeded a threshold, the detection was retained; conversely, if all voxels along the ray were below the threshold, the detection was discarded. Because of volume averaging, a threshold lower than 0 HU had to be selected; based on results from an earlier study (6), use of a threshold of −124 HU eliminated a number of false-positive results and no true-positive results for the optimal filter (filter 6) in that study. This method was applied only to detections of possible polyps that measured less than 1.0 cm. All detections of possible polyps 1.0 cm or greater were considered to be of potential clinical importance and were retained. Details about polyp densitometry measurements can be found in reference 6.

Assessment of False-Positive Detections

False-positive detections were assessed to understand how many were due to the presence of what could plausibly be considered polyps and how many could readily be discarded and eliminated from further evaluations. Because there were hundreds of false-positive detections, a subset was analyzed. Twenty false-positive detections of polyps measuring 1.0 cm or larger were randomly selected for further analysis.

Characterization of Colonic Distention and Colonic Segmentation

To evaluate the influence of adequate colonic distention on polyp detection, each segment of each colon was given a distention score (6). Scores were assigned by a single radiologist observer (R.M.S.) who inspected an anteroposterior exoscopic CT colonographic image of the surface-rendered colon. The scoring scale was as follows: 0, collapsed; 1, partially distended; and 2, well distended. A score was given for each colonic segment on the basis of the amount of distention that predominated in that segment. Differentiation between partially distended and well-distended segments was based on expected colonic diameters for each segment, according to the observer's experience at performing air-contrast barium enema examinations and expected distributions of anatomic size (eg, the knowledge that the cecum can distend more than the sigmoid).

The colon was divided into eight segments. Polyps were located in segments on the basis of the colonoscopy report, and, as described above, each colonic segment was scored for distention. The segments scored were the rectum, the sigmoid colon, the descending colon, the splenic flexure, the transverse colon, the hepatic flexure, the ascending colon, and the cecum.

Statistical Analysis

P values less than or equal to .05 were considered to represent significant differences. Unpaired *t* tests were used to compare polyp size distributions and colonic distention scores. The Fisher exact test was used to compare sensitivity values and polyp spatial distributions for unmatched data. CIs for proportions were computed from the binomial distribution.

Sensitivity for polyp detection was computed as a function of polyp size and pathologic type and on a per-polyp and per-patient basis. Because more than one blinded radiologist analyzed each set of CT colonographic images, we computed sensitivity for polyp detection in two ways: by using *(a)* the number of true-positive polyp detections that were made by both radiologists and *(b)* the number of true-positive detections made by only one of the two radiologists.

We report the average number of false-positive detections per patient for CAD rather than specificity. As is typical for other types of CAD used in radiology (eg, CAD for breast microcalcifications and masses, CAD for lung nodules at helical CT), CAD for colonic polyps usually results in one or more false-positive detections for each CT colonographic examination (ie, specificity approaches zero).

A free-response receiver operating characteristic (FROC) analysis of CAD results was performed by plotting the mean number of lesion-localized true-positive detections against the mean number of false-positive detections per CT colonographic image set (which includes both supine and prone CT colonographic images of each patient) for various CAD diagnostic criteria (16). The criteria for an irregularity to be considered a lesion were as follows: mean curvature between -4.0 and -0.5 cm^{-1}, at least two vertices, and a size of at least 0.2 cm. (These criteria constituted "filter 8" in our previous study [6].)

The following variables were linked and varied in equal increments over the following ranges to produce each point on the FROC curve: sphericity (a measure of roundness, where 0 is perfectly spherical and 2 is a surface curved in one direction and almost flat in the perpendicular direction, approaching so-called cylindrical curvature), 0.8–1.2; number of vertices (a measure of area), 0–20; and minimum size, 0–1.0 cm. These ranges were chosen on the basis of histogrammatic analysis of each parameter considered individually for the true- and false-positive results. The midpoints of the ranges were identical to the criteria for filter 7 used in this study.

Summary of Sensitivity Values for Polyp Detection by Radiologists, CAD, and Radiologists and CAD

Detecting Agent	Colonoscopy*			Retrospective Review of CT Colonographic Data†		
	Large (≥1.0-cm) Polyps	Small (<1.0-cm) Polyps	All Polyps	Large (≥1.0-cm) Polyps	Small (<1.0-cm) Polyps	All Polyps
Radiologists‡	48 (12/25)	22 (9/40)	32 (21/65)	67 (12/18)	75 (9/12)	70 (21/30)
CAD§	48 (12/25)	15 (6/40)	28 (18/65)	67 (12/18)	50 (6/12)	60 (18/30)
Radiologists and CAD‖	64 (16/25)	25 (10/40)	40 (26/65)	89 (16/18)	83 (10/12)	87 (26/30)

Note.—Data are sensitivity values expressed as percentages. Data in parentheses are numbers used to calculate the percentages.

* Denominators used to calculate sensitivity values were total numbers of polyps found with colonoscopy.

† Denominators used to calculate sensitivity values were total numbers of polyps found at retrospective review of CT colonographic data, which were evaluated in an unblinded fashion, after the precise size and location of polyps had been determined at colonoscopy.

‡ For this group, a polyp was considered to be detected at CT colonography if either of two radiologists, blinded to colonoscopy results, identified the polyp.

§ For this group, a polyp was considered to be detected at CT colonography if the computer program found a polyp that was confirmed at colonoscopy.

‖ For this group, a polyp was considered to be detected at CT colonography if either of the two radiologists, blinded to colonoscopy results, or the computer identified the polyp.

RESULTS

Twenty patients had 65 polyps, of which 25 were at least 1.0 cm (large polyps) and 40 were smaller than 1.0 cm (small polyps). Results of pathologic examination revealed that there were 49 adenomas, two cancers, and 13 hyperplastic polyps. No pathologic data were available for one small polyp. This distribution was more heavily weighted with adenomas than would be expected on the basis of known prevalences of polyps in these size ranges. The majority (14 of 25, 56%) of the large polyps were located in the transverse colon, ascending colon, or cecum. There was no significant difference between the sizes of the large polyps in this study (mean ± SD, 1.8 cm ± 0.7) and the sizes of the large polyps observed in our earlier study (1.7 cm ± 1.2) ($P = .84$).

Eighteen (72%) of 25 large polyps and 12 (30%) of 40 small polyps could be identified in retrospect on the CT colonographic images. Thus, these were the maximum possible sensitivities for the CAD algorithm. To avoid bias, we did not use CAD to help locate the remaining polyps. Observers had 70 opportunities (ie, images from one supine examination and one prone examination were available) to identify the 35 missed polyps. Potential causes for the false-negative diagnoses include the fact that four polyps were located in collapsed segments, 20 were buried in fluid or stool, one was obscured by motion artifact, two were obscured by streak artifact, eight coexisted with diverticular disease, and two were obscured by a rectal balloon. No cause could be identified for 33 false-negative diagnoses.

The mean colonic distention score was 1.6 ± 0.7; 423 (66%) of 640 colonic segments were well distended, 130 (20%) were partially distended, 73 (11%) were collapsed, and 14 (2%) had been resected. (Percentages do not add up to 100% due to rounding.) There was no significant difference between the distention scores in the patients with polyps and the scores in the patients without polyps (1.6 ± 0.7 for both groups; $P = .84$, unpaired t test). The colons were clinically considered to be well distended and were better distended than in our previous study (6); this was thought to be due to the use of the single-breath-hold technique, multisection CT scanner, and improved insufflation technique that resulted from greater clinical experience. (In this study, nurses were encouraged to inflate the colon more vigorously, and all patients in the current study received glucagon, whereas all but one patient in our previous study did not.)

On a per-colon basis, there was no significant difference between insufflated colonic volume as observed on supine images (1,193 cm^3 ± 319) and that observed on prone images (1,157 cm^3 ± 337) (difference, 36 cm^3 ± 185; $P = .26$, two-tailed paired t test). Partially resected colons were excluded from this analysis because their volumes are smaller.

The CAD algorithm detected 12 of the large polyps and six of the small polyps, for sensitivities of 48% and 15%, respectively, and an overall sensitivity of 28% (Table). Of 18 polyps detected with CAD, 10 were detected on supine CT colonographic images only; five, on prone CT colonographic images only; and three, on both supine and prone CT colonographic images. Per-patient sensitivity was 65% (13 of 20). Sensitivity for adenoma detection was 27% (13 of 49) for all adenomas and 53% (10 of 19) for large adenomas. Per-lesion sensitivity for detecting polyps proximal to the splenic flexure (50% [eight of 16]) was not significantly different from per-lesion sensitivity for detecting distal polyps (44% [four of nine]).

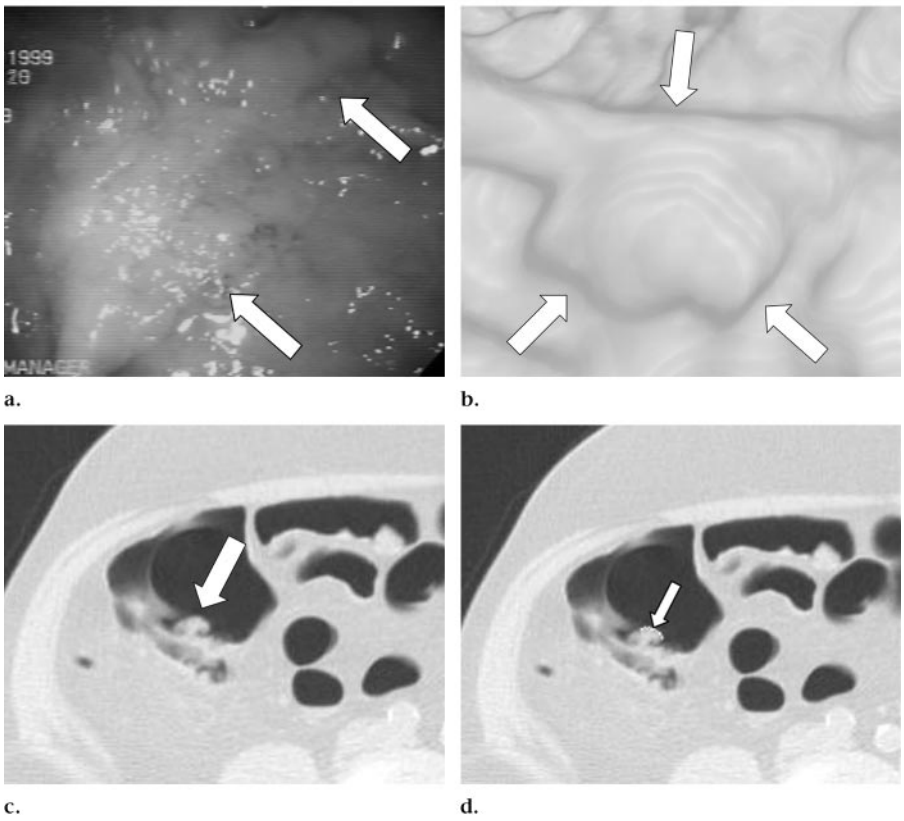

Figure 1. Images of a flat, sessile, tubular 2-cm adenoma (arrows) with high-grade dysplasia that was detected in the cecum of a 75-year-old man with our CAD algorithm but not by either radiologist. **(a)** Image from conventional colonoscopy. **(b)** Perspective endoluminal CT colonographic image. The raised portion of the mass is visible. **(c, d)** Transverse CT colonographic images obtained with the patient in the supine position **(c)** without and **(d)** with automated detection marks. In **d**, an automated detection mark (small white line on edge of mass) identifies one portion of the polyp. A substantial amount of retained solid stool can be observed.

At least one radiologist detected 12 of the large polyps and nine of the small polyps, for sensitivities of 48% and 23%, respectively, and an overall sensitivity of 32% (Table). Only a minority of polyps (six of the large polyps and three of the small polyps) were detected by both radiologists. On a per-patient basis, sensitivity was 60% (12 of 20). Per-lesion sensitivity for adenomas was 31% (15 of 49) for all adenomas and 53% (10 of 19) for large adenomas. Per-lesion sensitivity for polyps proximal to the splenic flexure (50% [eight of 16]) was not significantly different from per-lesion sensitivity for distal polyps (44% [four of nine]).

The CAD algorithm detected four large polyps and one small polyp that were not detected by either radiologist and seven large and four small polyps that were not detected by both radiologists. Examples are shown in Figures 1–3. The four large polyps not detected by either radiologist but found with CAD ranged in size from 2.0 to 3.5 cm, were all adenomas, and were scattered throughout the colon (in the cecum, ascending colon, sigmoid colon, and rectum). One was called a flat lesion (ie, its height was less than or equal to half the size of its base) by the colonoscopist. The additional three large polyps not detected by both radiologists (ie, detected by only one of two radiologists) but found with CAD ranged in size from 1.0 to 3.3 cm. Two were adenomas, and one was a carcinoma; they were located in the ascending colon, hepatic flexure, and transverse colon, respectively. Two of these were flat lesions. The radiologists found four large polyps and four small polyps that were missed with the polyp detector.

For the combination of detection with CAD and by at least one radiologist, 16 large polyps and 10 small polyps were found, for a combined sensitivity of 64% (16 of 25) and 25% (10 of 40) for large and small polyps, respectively; overall sensitivity was 40% (26 of 65) (Table). The addition of CAD markedly improved sensitivity for detecting the most clinically important large polyps (a 16% increase; sensitivities without CAD and with CAD were 48% and 64%, respectively). The fraction of polyps missed by radiologists but detected with CAD was four of 13 (31%; 95% two-sided CI: 9, 61). On a per-patient basis, sensitivity was 75% (15 of 20). Sensitivity for adenoma detection was 39% (19 of 49) for all adenomas and 74% (14 of 19) for large adenomas. The fraction of patients with polyps missed by radiologists but detected with CAD was three of eight (38%; 95% two-sided CI: 9, 76). Of large polyps identifiable in retrospect, 89% (16 of 18) were detected with the addition of CAD (Table).

There were an average of 15 false-positive detections per patient. These could be reduced by 29% to an average of 11 per patient with the use of polyp densitometry. No true-positive results were excluded when polyp densitometry was applied. A random sample of 20 false-positive classifications of polyps 1.0 cm or larger revealed that only 20% (four of 20) were of entities that could plausibly have been considered polyps, and 80% (16 of 20) were of entities that could easily be defined and discarded from the classification (ie, six haustral folds, four ileocecal valves, two rectal tubes, two rectal folds, one mobile stool, and one case of extrinsic compression by the uterus were misclassified). Of the four false-positive lesions that could plausibly have been considered polyps, one was on a fold, one was at a confluence of folds, one was related to a thick fold, and one may have been due to an indentation in the colon by the liver.

Results of FROC analysis (Fig 4) show the trade-offs between sensitivity and false-positive detections. To produce the graph in Figure 4, polyp size and sphericity thresholds were incrementally varied between more-inclusive and less-inclusive values. The figure shows how greater sensitivity for detecting large polyps (more-inclusive case) can be achieved at the expense of a greater number of false-positive detections. For example, 72% sensitivity (the maximal achievable value for this CT colonographic data set) can be attained at the expense of 31 false-positive results per CT colonographic examination or an average of 62 false-positive results per patient (when each patient undergoes one CT colonographic examination in the prone position and one in the supine position).

DISCUSSION

In a previous work (6), we showed that our CAD algorithm had a sensitivity of 71%, with an average of 3.5 false-positive detections per patient for detecting polyps 1.0 cm or larger. In that study, we examined multiple-breath-hold single-section helical CT colonographic data obtained with overlapping sections. Only results of CT colonography performed with patients in the supine position were examined to simplify the analysis. On the basis of the results of that study, we designed an experiment to validate the earlier results and to improve upon them, namely by using both the supine and prone data obtained at single-breath-hold multisection helical CT colonography that did not require section overlap and by examining a consecutive series of patients to avoid selection bias.

In the present study, our computer algorithm boosted sensitivity for polyp detection at CT colonography by 16% (from 48% to 64%) and increased the number of true-positive detections of polyps 1.0 cm or larger by 33% (from 12 to 16), compared with the sensitivity of radiologists' interpretations alone. Statistical testing of observed detections revealed that the increase in the number of new polyps detected was significantly different from zero. That is, the CI for the proportion of polyps missed by both radiologists but found with CAD (four of 13) did not include zero. This is the basis for our claim that the 16% increase in sensitivity is significant.

This is a dramatic improvement, which, if confirmed in a larger clinical trial, could have substantial benefit for patients. The size of such a trial can be estimated from our results. To obtain an increase in sensitivity of 16%, starting from 48% with an α of 0.05 and a power of 0.80, a sample size of 59 would be required (as calculated with the one-sided binomial statistic). This calculation assumes the same prevalence of polyps in the patient population as in the present study (ie, it assumes an enriched population).

In the present study, the CAD algorithm detected four large (≥1.0 cm) polyps missed by both radiologists and seven missed by one of two radiologists. Missed polyps ranged in size from 1.0 to 3.5 cm. Three were observed to be flat lesions at colonoscopy. Six of the missed lesions were adenomas, and one was a carcinoma. According to our present understanding of the adenoma-to-carcinoma sequence, adenomas are neoplasms with malignant potential (17). These missed lesions were therefore clearly clinically important ones that needed to be found.

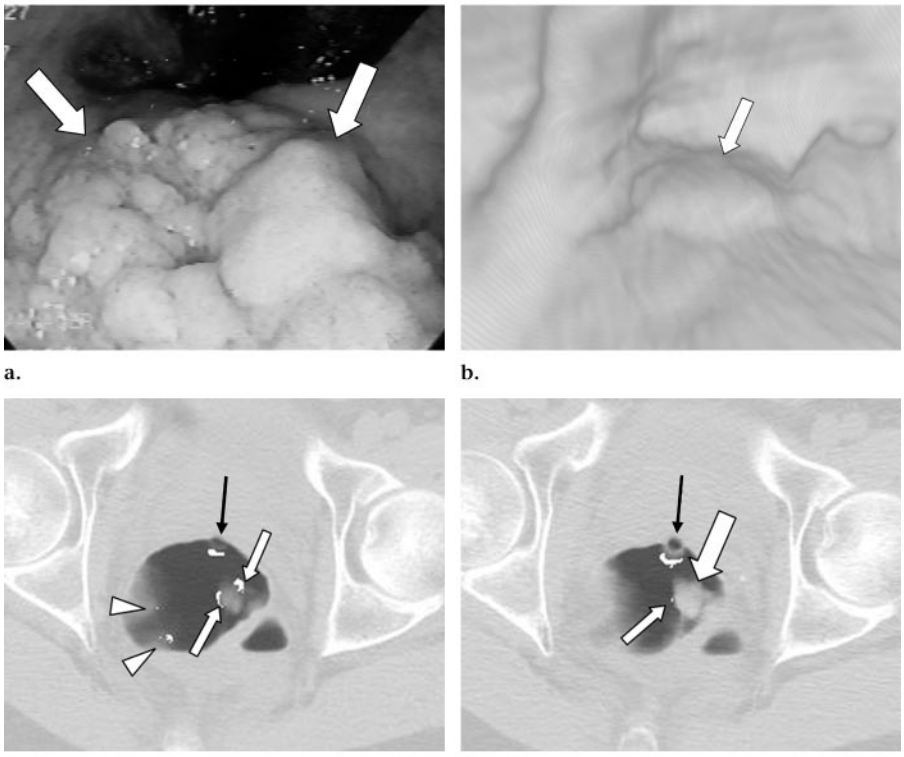

Figure 2. Images of a 3.5-cm villous adenoma, detected with our CAD algorithm but not by either radiologist, in the rectum of a 54-year-old man. (Large white arrows indicate a soft-tissue mass; small white arrows indicate an automated detection mark.) **(a)** Image from conventional colonoscopy. **(b)** Perspective endoluminal CT colonographic image. **(c, d)** Transverse CT colonographic images of two adjacent sections obtained with the patient in the supine position. In **d**, the automated detection mark (small white line on edge of mass) appears small but is larger on adjacent section (**c**) of the top of the polyp. The tip of the rectal tube (black arrows) and false-positive locations (arrowheads in **c**) on rectal folds were also marked by the computer algorithm.

While in retrospect, missed polyps such as those shown in the figures may appear obvious, perceptual error may be unavoidable when hundreds of images are being interpreted. In fact, it has been shown that approximately 17%–21% of false-negative diagnoses are due to perceptive error (2). Additional training may help to improve sensitivity; nevertheless, trained observers will miss lesions that are visible in retrospect. CAD may help to reduce such misses.

To place in perspective the importance of locating four to seven of 28 large polyps, consider the recent article on CAD of breast cancer during screening mammography (18). In that study, a commercially available Food and Drug Administration–approved CAD system helped radiologists identify an additional eight cancers in a population of 12,860 patients. While it is not known how many cancers were actually present in that population, the percentage improvement in the number of lesions detected (19.5%) is comparable to our results (a 33% increase in the number of true-positive detections), despite the early stage of research on CAD in the colon.

To place the number of false-positive results observed in this study in perspective, in the study by Freer and Ulissey (18), fewer than 3% of detections were deemed "actionable"; the remaining 97% were false-positive detections. Therefore, it may be that a large number of false-positive detections are tolerable if most of them can easily be recognized and ignored. For CT colonography with CAD, the clinical effect of false-positive detections has yet to be assessed.

Our results showed that, rather than usurping the diagnostic role of the radiologist, CAD has a complementary role in polyp detection. Both the radiologists

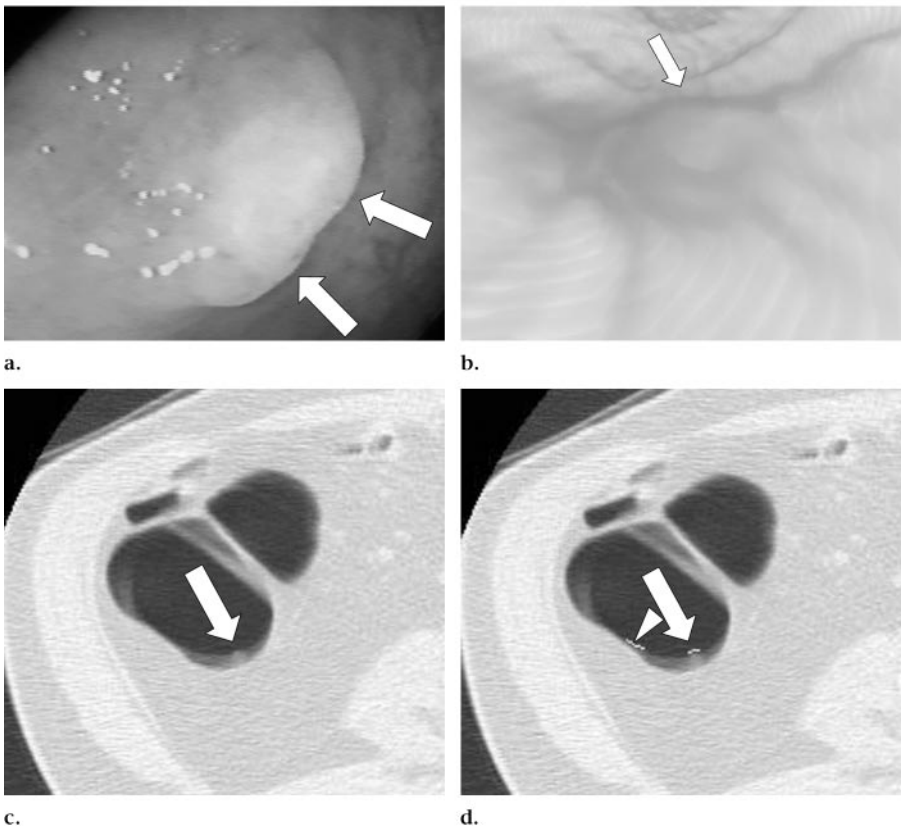

Figure 3. Images of a small (described as "diminutive" by the colonoscopist) tubular adenoma (arrows) that was detected in the hepatic flexure of a 66-year-old man by our CAD algorithm but not by one of two radiologists. **(a)** Image from conventional colonoscopy. **(b)** Perspective endoluminal CT colonographic image obtained with the patient in the prone position. **(c)** Unmarked and **(d)** marked transverse CT colonographic images obtained with the patient in the prone position. A false-positive detection mark (arrowhead in **d**) is also seen.

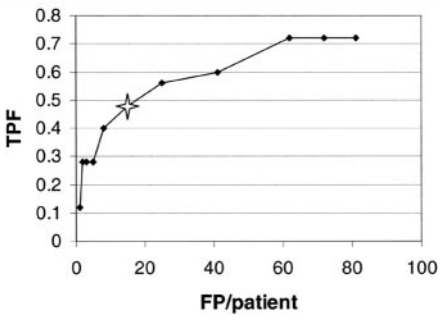

Figure 4. FROC curve for the computer-aided polyp detection algorithm used in this study shows results that correspond to a sensitivity value of 48% (15 false-positive detections per patient) (✧). The maximum achievable sensitivity for the CAD algorithm was 72% because 28% of the large polyps revealed at colonoscopy could not be located in retrospect on the CT colonographic images. The FROC curve illustrates the trade-off between sensitivity and false-positive detections. *TPF* = true-positive fraction per polyp. *FP/patient* = number of false-positive detections per patient.

and the CAD algorithm identified polyps missed by the other. The combination of the two yielded improved sensitivity for polyp detection. The reason for the improvement with CAD is not yet known; it may be because radiologists focus their attention primarily on analysis of the two-dimensional CT images while the CAD algorithm analyzes the three-dimensional data and may identify more complex abnormal shapes.

Ultimately, the combined sensitivity of the radiologist and CAD will need to be evaluated under conditions simulating the proposed clinical use of CAD—that is, CAD results would be given to the radiologist, who would determine them to be valid or false, and a final diagnostic decision would then be rendered. Such a study will also be necessary to determine the effect on specificity of the average of 11 false-positive detections per patient observed in this study and to determine how the use of CAD affects radiologists' image interpretation time.

Our CAD software processes CT colonographic examinations relatively quickly (4 minutes per patient). Even without any changes to our algorithm, improving computer hardware speeds will directly enhance performance. Thus, patient throughput is unlikely to be an issue for our CT colonography CAD software.

The current study was conducted with improved techniques. The degree of colonic distention achieved in this study was better than that achieved in our previous study, with a marked shift from partially distended to well-distended segments of colon (66% of segments were well distended in the present study vs 46% in the previous study, and 20% of segments were partially distended in the present study vs 40% in the previous study) (6). In addition, all but one CT colonographic examination in the present study was performed without section overlapping.

Despite these technical improvements, sensitivity for detection of large polyps was worse for the radiologists compared with that observed in other large clinical trials (1–3). The radiologists had a sensitivity of 82% (23 of 28) in our previous study (6) versus 48% (12 of 25) in the current study ($P = .02$, Fisher exact test). However, sensitivity was not significantly lower for the automated polyp detector. For example, the sensitivity of filter 7 was 68% (19 of 28) in our previous study and 48% (12 of 25) in the present study ($P = .17$, Fisher exact test).

There are a number of possible explanations for these differences in performance. The most likely explanation is that because the interpreting radiologists knew that the incidence of disease was lower in the current cohort of patients, they used more stringent criteria to maintain higher specificity. There are limited data in the literature regarding comparisons of the sensitivity of CT colonography between populations with different prevalences of disease (ie, asymptomatic versus symptomatic individuals). However, results of one large recent study indicated that sensitivity was not significantly different between asymptomatic (screening) and symptomatic patient groups (3).

Other possible explanations for the decreased performance of radiologists observed in this study are differences between the two studies in the distribution of types, shapes, sizes, and locations of polyps; the presence of more residual stool during examinations in the current study; and case-selection bias in the previous study (the current study was of a

consecutive series of patients). We have shown, however, that the sizes of the large polyps in the present study were not significantly different from those in the first study. The use of a wetter bowel preparation (ie, Colyte) instead of sodium phosphate (19) may also be a contributing factor, since in about 28% (20 of 70) of the examinations in which polyps were missed, the polyps may have been missed due to the presence of obscuring fluid or stool. However, the same bowel preparation was used in our earlier study. We also emphasize that every patient was imaged in both prone and supine positions, and both sets of images were interpreted by the radiologists and with CAD. These are very effective means of dealing with excessive fluid. It seems that these polyps were simply more difficult for the radiologists to detect.

The right-sided distribution of polyps could also have influenced detection. In both this study and our previous study, there was a trend toward a right-sided distribution of polyps. Sixteen of 25 large polyps in the current study and 18 of 28 large polyps in our previous study were proximal to the splenic flexure ($P = 1.0$, Fisher exact test). The right colon is larger in diameter (and therefore has more mucosal surface to be inspected), and the lesions in this portion of the colon tend to be flat. The confusion of polyps with normal structures like the ileocecal valve can also be problematic. Polypoid bulbous folds are normally present in the right colon and can mimic polyps, making detection of real polyps more difficult. Liquid stool with particulate debris is commonly seen in the right colon because the small bowel keeps pumping these contents through the ileocecal valve. Despite these factors, we found no significant differences in sensitivity for polyp detection on the basis of location (ie, proximal or distal colon) for either the radiologists or the CAD system.

We used 5-mm collimation with 2-mm overlap in this study to maintain consistency with the CT colonographic parameters used in our previous research (6). The multisection CT scanner permits scanning of the entire colon in a single breath hold with thinner collimation (ie, 3 mm or lower). We believe that the use of thinner collimation may improve the sensitivity of CAD and have initiated research to validate this hypothesis. However, although it seems reasonable that thinner collimation may improve sensitivity, there is currently no evidence that thinner sections result in improved detection of lesions. The use of thinner sections requires higher radiation doses if noise levels are to be kept constant. Therefore, solid evidence must be garnered for the use of thinner sections before this is accepted as the standard.

Note also that we chose to use filter 7 instead of filter 6, which we described in our previous report as "optimum" (6). Filter 6 proved to be insufficiently sensitive to improve on the results of the radiologists.

Most of the polyps (13 of 18) were found by the CAD system on supine CT colonographic images. The remaining five were detected only on prone images. These results support the addition of CT colonography with patients in the prone position to the standard protocol, as recommended by other researchers (2,20).

In preliminary studies of CAD systems, both Yoshida et al (21,22) and Paik et al (23) achieved higher sensitivity for polyp detection. Yoshida et al (21) reported 100% sensitivity for nine polyps—six between 0.5 and 1.0 cm and three larger than 1.0 cm—with 2.5 false-positive detections per patient with supine and prone scanning. Paik et al (23) reported 100% sensitivity for nine polyps larger than 0.85 cm, with 3.6 false-positive detections per patient with supine scanning. Compared with our study, in each of these studies the number of polyps was smaller, and the patient populations may have been markedly different and/or highly selected.

CAD is still in an early stage of development but has great potential. There are still many challenges ahead and areas that need improvement (7). Choices of features to extract and classifiers that combine and limit features are just being evaluated and developed. Measurements of CT numbers and shape are likely to be important factors in a working CAD system for CT colonography.

Early investigations suggest that labeling stool with ingested contrast agents may enable CT colonography of the colon without the use of bowel preparations, yielding a more "patient-friendly" examination (24). It is possible to eliminate labeled stool and fluid with image-processing techniques (25). The processed images could then be assessed with CAD algorithms designed for use with colons that have been cleansed with a bowel preparation.

Our database consists mainly of polyps rather than cancers. A survey of seven recent studies of CT colonography reveals that sensitivity for cancer detection with this modality has uniformly been 100% (4). Although sensitivity of CT colonography for cancer detection may be lower in colons not cleansed with a bowel preparation (for which no data are yet available), given the excellent performance of human observers for cancer detection, it is likely that the need for CAD is less for detection of such advanced lesions.

In summary, in a patient population having a lower prevalence of disease, we have shown how a computer-assisted polyp detection algorithm can improve the diagnostic sensitivity of CT colonography when used in conjunction with radiologists' interpretations.

Acknowledgments: We thank Andrew Dwyer, MD, for critical review of the manuscript. We thank Robert L. MacCarty, MD, and Tim J. Welch, MD, for CT colonographic image interpretation and John Garry, BS, for data management.

References

1. Fenlon HM, Nunes DP, Schroy PC III, Barish MA, Clarke PD, Ferrucci JT. A comparison of virtual and conventional colonoscopy for the detection of colorectal polyps. N Engl J Med 1999; 341:1496–1503. [Erratum: N Engl J Med 2000; 342: 524.]
2. Fletcher JG, Johnson CD, Welch TJ, et al. Optimization of CT colonography technique: prospective trial in 180 patients. Radiology 2000; 216:704–711.
3. Yee J, Akerkar GA, Hung RK, Steinauer-Gebauer AM, Wall SD, McQuaid KR. Colorectal neoplasia: performance characteristics of CT colonography for detection in 300 patients. Radiology 2001; 219: 685–692.
4. Summers RM, Hara AK, Luboldt W, Johnson CD. Computed tomographic and magnetic resonance colonography: summary of progress from 1995 to 2000. Curr Prob Diagn Radiol 2001; 30:141–168.
5. Summers RM, Beaulieu CF, Pusanik LM, et al. Automated polyp detector for CT colonography: feasibility study. Radiology 2000; 216:284–290.
6. Summers RM, Johnson CD, Pusanik LM, Malley JD, Youssef AM, Reed JE. Automated polyp detection at CT colonography: feasibility assessment in a human population. Radiology 2001; 219:51–59.
7. Summers RM. Challenges for computer-aided diagnosis for CT colonography. Abdom Radiol 2002; 27:268–274.
8. Hara AK, Johnson CD, Reed JE, et al. Reducing data size and radiation dose for CT colonography. AJR Am J Roentgenol 1997; 168:1181–1184.
9. Reed JE, Johnson CD. Automatic segmentation, tissue characterization, and rapid diagnosis enhancements to the computed tomographic colonography analysis workstation. J Digit Imaging 1997; 10:70–73.
10. Johnson CD, Toledano A, Herman B, Dachman AH, McFarland EG, Lu DS. CT colonography: performance evaluation in a multicenter setting (American College of Radiology imaging network study 6653) (abstr). Radiology 2001; 221(P):308.
11. Summers RM, Feng DH, Holland SM,

Sneller MC, Shelhamer JH. Virtual bronchoscopy: segmentation method for real-time display. Radiology 1996; 200:857–862.
12. Summers RM. Navigational aids for real-time virtual bronchoscopy. AJR Am J Roentgenol 1997; 168:1165–1170.
13. Summers RM. Image gallery: a tool for rapid endobronchial lesion detection and display using virtual bronchoscopy. J Digit Imaging 1998; 11:53–55.
14. Summers RM, Jerebko AK, Franaszek M, Malley JD. An integrated system for computer-aided diagnosis in CT colonography: work-in-progress. In: Lemke HU, Vannier MW, Inamura K, Farman AG, Doi K, eds. Computer assisted radiology and surgery. Berlin, Germany: Elsevier Science, 2001; 629–634.
15. Drebin RA, Carpenter L, Hanrahan P. Volume rendering. ACM Comput Graph 1988; 22:65–74.
16. Chakraborty DP. The FROC, AFROC and DROC variants of the ROC analysis. In: Beutel J, Kundel HL, Van Metter RL, eds. Handbook of medical imaging. Bellingham, Wash: SPIE Press, 2000; 771–796.
17. O'Brien M, Winawer SJ, Waye JD. Colorectal polyps. In: Winawer SJ, Kurtz RC, eds. Gastrointestinal cancer. New York, NY: Gower, 1992.
18. Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. Radiology 2001; 220:781–786.
19. Macari M, Lavelle M, Pedrosa I, et al. Effect of different bowel preparations on residual fluid at CT colonography. Radiology 2001; 218:274–277.
20. Chen SC, Lu DS, Hecht JR, Kadell BM. CT colonography: value of scanning in both the supine and prone positions. AJR Am J Roentgenol 1999; 172:595–599.
21. Yoshida H, Masutani Y, MacEneaney P, Dachman AH. Computer-aided detection of polyps in CT colonography based on geometric features. In: Chen CT, Clough AV, eds. Medical imaging 2001: physiology and function from multidimensional images. San Diego, Calif: SPIE, 2001; 4321:53–57.
22. Yoshida H, Masutani Y, MacEneaney PM, Doi K, Kim Y, Dachman AH. Detection of colonic polyps in CT colonography based on geometric features (abstr). Radiology 2000; 217(P):582.
23. Paik DS, Beaulieu CF, Jeffrey RB, Yee J, Steinauer-Gebauer AM, Napel S. Computer aided detection of polyps in CT colonography: method and free-response ROC evaluation of performance (abstr). Radiology 2000; 217(P):370.
24. Callstrom MR, Johnson CD, Fletcher JG, et al. CT colonography without cathartic preparation: feasibility study. Radiology 2001; 219:693–698.
25. Zalis ME, Hahn PF. Digital subtraction bowel cleansing in CT colonography. AJR Am J Roentgenol 2001; 176:646–648.