CT colonography as surveillance technique for patients at increased risk for colorectal cancer

Jensch, S.

Citation for published version (APA):
Jensch, S. (2009). CT colonography as surveillance technique for patients at increased risk for colorectal cancer.

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Chapter 3

CT Colonography with Limited Bowel Preparation: Performance Characteristics in an Increased-Risk Population

Sebastiaan Jensch
Ayso H. de Vries
Jan Peringa
Shandra Bipat
Evelien Dekker
Lubbertus C. Baak
Joep F. Bartelsman
Anneke Heutinck
Alexander D. Montauban van Swijndregt
Jaap Stoker

Radiology 2008; 247:122-132
Abstract

**Purpose:** To prospectively evaluate the sensitivity and specificity of computed tomographic (CT) colonography with limited bowel preparation for the depiction of colonic polyps, by using colonoscopy as the reference standard.

**Methods:** Institutional review board approval and written informed consent were obtained. Patients at increased risk for colorectal cancer underwent CT colonography after fecal tagging, which consisted of 80 mL of barium sulfate and 180 mL of diatrizoate meglumine. Bisacodyl was added for stool softening. A radiologist and a research fellow evaluated all data independently by using a primary two-dimensional approach. Discrepant findings for lesions 6 mm or larger in diameter were solved with consensus. Segmental unblinding was performed. Per-patient sensitivity and specificity, per-polyp sensitivity, and number of false-positive findings were found (for lesions $\geq 6$ mm and $\geq 10$ mm in diameter). Per-patient sensitivities (blinded colonoscopy versus CT colonography) were tested for significance with McNemar statistics. Interobserver variability was analyzed per segment (prevalence-adjusted bias-adjusted $\kappa$ values $[\kappa_p]$).

**Results:** One hundred fourteen of 168 patients (105 men, 63 women; mean age, 56 years) had polyps, with 56 polyps 6 mm or larger and 17 polyps 10 mm or larger. Per-patient sensitivities were not significantly different for CT colonography (consensus reading) and colonoscopy ($P \geq 0.070$). Sensitivity of CT colonography for patients with lesions 6 mm or larger and 10 mm or larger was 76% and 82%, respectively, and specificity of CT colonography was 79% and 97%, respectively. Blinded colonoscopy depicted 91% (lesions $\geq 6$ mm) and 88% (lesions $\geq 10$ mm) of disease in patients. Per-polyp sensitivity for CT colonography was 70% (lesions $\geq 6$ mm) and 82% (lesions $\geq 10$ mm). Number of false-positive findings was 42 (lesions $\geq 6$ mm) and six (lesions $\geq 10$ mm). $\kappa_p$ Was 0.88 (lesions $\geq 6$ mm) and 0.96 (lesions $\geq 10$ mm).

**Conclusion:** CT colonography with limited bowel preparation has a sensitivity of 82% and specificity of 97% for patients with polyps 10 mm or larger.
Introduction

As with other colonic examinations, computed tomographic (CT) colonography requires a clean colon for optimal assessment of the bowel wall [1–5]. However, the necessary cathartic bowel preparation is often described by patients as the most burdensome part of colonic examinations [6–9]. This might negatively affect patients' willingness to participate in a screening program. In 2001, limited bowel preparation for CT colonography with iodine- or barium-based contrast material for fecal tagging was introduced [10]. With fecal tagging, any fecal material in the colon is labeled so that colorectal polyps or cancer can be distinguished from fecal material. Because no extensive bowel cleansing is necessary, this approach might increase patient compliance in a screening setting [11–13].

Feasibility studies with limited bowel preparation have revealed promising results [14–18]. In addition, results of one large diagnostic accuracy study by Iannaccone et al [19] showed a sensitivity of 90% (71 of 79) and specificity of 92% (114 of 124) for patients with polyps, regardless of polyp size. These results are similar or superior to those of accuracy studies in which a cathartic bowel preparation was used [1–5].

To our knowledge, no confirmatory study on limited bowel preparation has been published since that of Iannaccone et al [19]. Therefore, the aim of our study was to prospectively evaluate the sensitivity and specificity of CT colonography with limited bowel preparation for the depiction of colonic polyps, by using colonoscopy as the reference standard.

Materials and Methods

This study received grant support from the Dutch Cancer Society (Koningin Wilhelmina Fonds) (No. CKTO 2003-02). The authors retained control of all aspects of the study.

Patients

Patients with a personal or family history of colorectal polyps or cancer were invited to participate from January 10, 2004, until November 10, 2005. All patients were scheduled to undergo routine colonoscopy at the endoscopy department of the Academic Medical Center or the Onze Lieve Vrouwe Gasthuis. Exclusion criteria were age younger than 18 years, a personal history of inflammatory bowel disease or familial adenomatous polyposis, prior allergic reaction to an iodine-containing contrast
agent, known colorectal polyps that were not removed at prior endoscopy, or participation in a research project that involved ionizing radiation within 12 months preceding CT colonographic examination. The institutional review boards of both hospitals approved the study. All patients gave written informed consent. Information about radiation associated with the examination was given to all patients before written consent was obtained.

### Table 1. Bowel preparation for CT Colonography

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Instruction and Dosage*</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 days before CT colonography</td>
<td>Avoid fiber rich food 20 ml Barium Sulphate at dinner</td>
</tr>
<tr>
<td>1 day before CT colonography</td>
<td>Three doses of 20 mL barium sulphate and three doses of diatrizoate meglumine, once at breakfast, lunch and dinner Only liquids after dinner 20 mg bisacodyl (oral) 18 hours prior to CT colonography</td>
</tr>
<tr>
<td>day of examination</td>
<td>60 mL diatrizoate meglumine and 10 mg bisacodyl (oral) at breakfast 60 mL diatrizoate meglumine 1 hour before CT colonography</td>
</tr>
</tbody>
</table>

*Barium Sulphate (40% w/v, 30%w/w), Diatrizoate Meglumine (200mgI/ml).

**Diagnostic Procedures**

**Bowel preparation**

Patients were instructed not to eat high-fiber foods for 2 days before the examination. A combination of 80 mL of barium sulfate suspension (Tagitol V; E-Z-Em, Westbury, NY) and 180 mL of diatrizoate meglumine (200 milligrams of iodine per milliliter, prepared by hospital pharmacy, Academic Medical Center) was prescribed for fecal tagging (Table 1). Bisacodyl (30 mg, prepared by hospital pharmacy, Academic Medical Center) was given to soften fecal material for optimal contrast material and feces mixing and to avoid solid, sticky stool.

**CT colonographic image analysis**

CT colonographic data were evaluated for colorectal polyps and carcinoma independently by two observers who were blinded to clinical data. Observer 1 (J.P.) had 5 years of experience as an abdominal radiologist and had received training consisting of 50 cathartic untagged and 20 limited preparation tagged CT
Accuracy of limited prepped CT colonography

colonographic cases with colonoscopic verification. Observer 2 (A.H.d.V.) was a research fellow who had evaluated 100 cathartic untagged and 20 limited preparation tagged CT colonographic cases with colonoscopic verification as part of his fellowship in CT colonography. Evaluation of data was performed with a dedicated CT colonographic workstation (Viewforum; Philips Medical Systems) by using a primary two-dimensional evaluation technique, with instant onscreen correlation with multiplanar reformation images and three-dimensional endoluminal views for problem solving. Standard window settings were applied (window width, 1250 HU; window level, –50 HU), but observers were free to adjust these settings.

Figure 1. Flow-chart shows patient participation

- 468 pt at increased risk for colorectal cancer (316 [blinded]; 152 [blinded])
- 85 not requested
  - Reasons:
    - organizational, 25
    - unreachable, 60
- 383 requested
- 203 refused
- 180 participated
- CTC
- 174 underwent CTC
- 6 patients declined later on
  - Reasons:
    - second thoughts, 3
    - reported allergic reaction, 2
    - claustrophobia, 1
- 6 patients excluded for analysis
  - Reasons:
    - no segmental unblinding, 1
    - no colonoscopy, 1
    - inadequate bowel preparation, 4
- 168 patients for analysis
Chapter 3

Stool subtraction software was not used, and routine endoluminal fly-through was not performed. The observers classified CT colonographic polyps with regard to size, morphology (sessile, pedunculated, or flat [ie, height < 3 mm]), and location (cecum, ascending colon, transverse colon, descending colon, or sigmoid colon and rectum) by using the same classification as our gastroenterologists. Consensus reading, or a double-read strategy, was performed in case one observer found a lesion 6 mm or larger in diameter but the other did not or if both observers found the same lesion but disagreed on morphology, size, or location. Agreement forms were completed after each examination for polyps 6 mm or larger in diameter and included two-dimensional and three-dimensional images of the polyp at CT colonography, with the location marked in the three-dimensionally rendered colon. The research nurse (A.H.) used these agreement forms to provide feedback for segmental unblinding during colonoscopy. Interpretation times, with the exclusion of reporting time and evaluation of extracolonic findings, were recorded for each examination by both observers.

**CT colonographic stool-tagging analysis**

A questionnaire was filled out by observer 1 to subjectively assess the adequacy of stool tagging for every examination. A four-point scale was used in which a score of 1 indicated good tagging (homogeneous, diagnostic for all lesions); a score of 2, adequate tagging (fairly homogeneous and diagnostic for all lesions); a score of 3, poor tagging with untagged solid stool less than 6 mm (diagnostic for lesions ≥6 mm); and a score of 4, poor tagging with untagged solid stool 6 mm or larger (not diagnostic for lesions ≥6 mm).

<table>
<thead>
<tr>
<th>Table 2. Baseline patient characteristics (n=168)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male-to-female ratio</td>
</tr>
<tr>
<td>Age (y)</td>
</tr>
<tr>
<td>Hospital</td>
</tr>
<tr>
<td>Academic Medical Center / Onze Lieve Vrouwe Gasthuis Patients</td>
</tr>
<tr>
<td>Without polyps</td>
</tr>
<tr>
<td>With polyps (all sizes)</td>
</tr>
<tr>
<td>With a polyp ≥10mm</td>
</tr>
<tr>
<td>History of colorectal polyps or cancer</td>
</tr>
<tr>
<td>Family history of colorectal polyps or cancer</td>
</tr>
</tbody>
</table>

Note.-unless otherwise indicated, data are numbers of patients.
Colonoscopy

An experienced staff member (a gastroenterologist \( n = 15 \) or a gastrointestinal surgeon \( n = 4 \) with an average experience of 11 years; range, 1–26 years) or a gastroenterology fellow with direct supervision (including E.D., J.F.B., L.C.B.) performed the colonoscopic examinations with a standard colonoscope (CF-140L; Olympus, Tokyo, Japan). Patients received 2.0–10.0 mg midazolam (Dormicum; Roche, Basel, Switzerland) and 0.05–0.1 mg fentanyl (Fentanyl-Janssen; Janssen Pharmaceuticals, Beerse, Belgium) on request. The examination was recorded on videotape. Segmental location, morphology, and size of polyps were documented on a case-record form by the attending research nurse (A.H.). Polyp size was measured with open biopsy forceps (8 mm in length). Segmental unblinding was performed only for lesions 6 mm or larger that both observers agreed on (independently or through consensus). Histological findings were obtained at colonoscopy, except in those cases in which it was technically not possible or when material was lost in the colonic lumen during the procedure.

**Determination of lesion status**

Colonoscopy after unblinding of the CT colonographic findings served as a reference standard. For CT colonography, a polyp was considered true-positive if \( a \) the segment or adjacent segment corresponded with the reference standard segment and \( b \) the polyp size as estimated by the endoscopist corresponded with size as measured at CT colonography, considering a margin of error of 50%. Polyps 10 mm or larger at colonoscopy that were not identified by either observer were reevaluated with knowledge of the colonoscopic findings by an independent radiologist with experience in the evaluation of more than 1000 CT colonographic studies. This reevaluation was performed to differentiate between technical errors (polyps obscured by technical failure [e.g., inhomogeneous tagging]), perception errors (polyps visible in retrospect), and occult lesions (polyps not visible in retrospect). This classification was also used by MacCarty et al [20]. Furthermore, a research fellow in CT colonography (S.J.) who was not involved in evaluation of the CT colonographic assessed the nature of false-positive findings 10 mm or larger (eg, untagged stool or haustral folds).

**Patient experience and preference**

To evaluate patient experience, we asked patients to fill out a questionnaire on the day of the CT colonographic examination. Patients were asked if they had
experienced diarrhea as a side effect and graded it on a five-point scale (none, mild, moderate, severe, extreme, with 0 indicating none and 5 indicating extreme). Furthermore, the overall burden of the complete preparation was evaluated on the same scale. Patient preference was assessed 5 weeks after colonoscopy. This, like adverse reactions to tests, tends to temper with time [7,8], and the attitude at that time point will better reflect the attitude toward future screening. Patients were asked if they had experienced the bowel preparation for CT colonography or that for colonoscopy as most burdensome and if they would prefer CT colonography or colonoscopy in the future as the colonic examination of choice.

**Statistical Analysis**

**Outcome parameters**

Observers were instructed before the start of the study that only polyps at CT colonography that were matched to true polyps (i.e., adenomatous, hyperplastic, or hamartomatous polyps) on the basis of the histological report or, if histological findings were not acquired, on the basis of the endoscopic report would be considered true-positive findings. All other lesions at CT colonography that were matched to polypoid lesions at colonoscopy that have no malignant potential (e.g., lipomas or pseudopolyps) would be classified as false-positive findings. For CT colonography, a patient was considered to have a true-positive finding if CT colonography depicted at

---

**Figure 2. True positive patients with a lesion ≥10mm and ≥6mm**

Figure 2a shows a villous polyp of 20mm (arrow) in the sigmoid colon with adherent contrast on the surface of the lesion. A second smaller polypoid lesion (arrowhead) was visible that contained contrast in the centre of the lesion and was correctly classified as tagged stool. Figure 2b shows a 6mm polyp (arrow) in the descending colon that is situated within a contrast tagged fluid level. Both lesions were correctly identified by both observers.
least one polyp seen at colonoscopy, on the basis of the location and size criteria described previously. A patient was considered to have a false-negative finding if CT colonography depicted no polyps or only those of a lower size category in comparison to those depicted with the reference standard.

Sensitivity, specificity, and positive and negative predictive values were calculated for CT colonography on a per-patient basis and were stratified according to polyp size (all sizes, $\geq 6$ mm, $\geq 7$ mm, $\geq 8$ mm, $\geq 9$ mm, and $\geq 10$ mm, as well for the size range 6–9 mm). Per-patient sensitivity was calculated for blinded colonoscopy. For calculation of the per-polyp sensitivity for CT colonography and blinded colonoscopy, generalized estimating equations (GEE) were used to revise the data clustering and dependency (21) because some patients had more than one polyp. For CT colonography, the number of false-positive findings was calculated on a per-lesion basis. For the per-polyp sensitivity and false-positive findings, data were also stratified according to polyp size (all sizes, $\geq 6$ mm, $\geq 7$ mm, $\geq 8$ mm, $\geq 9$ mm, and $\geq 10$ mm, as well for the size range 6–9 mm).

### Table 3. Histology and morphology of polyps at colonoscopy

<table>
<thead>
<tr>
<th></th>
<th>Polyps 6-9mm (n=39)</th>
<th>Polyps $\geq$10mm (n=17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No histology finding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histology finding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyperplastic</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Adenomatous*</td>
<td>27</td>
<td>14</td>
</tr>
<tr>
<td>tubulovillous</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>tubular</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>serrated</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>non-specified</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Hamartomatous</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Morphology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sessile</td>
<td>26</td>
<td>9</td>
</tr>
<tr>
<td>Pedunculated</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Flat</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Note- Data are number of polyps.
*No specification of the degree of dysplasia is provided by the pathologist in our institution.
Diagnostic value
We used the McNemar test to compare per-patient sensitivity values between blinded conventional colonoscopy, consensus reading, and observer 1 and observer 2 (colonoscopy vs consensus reading, colonoscopy vs observer 1, colonoscopy vs observer 2, consensus reading vs observer 1, consensus reading vs observer 2, and observer 1 vs observer 2). We also used the McNemar test to compare per-patient specificity values between consensus reading and both observers (consensus reading vs observer 1, consensus reading vs observer 2, and observer 1 vs observer 2). With the GEE model, we took into account the comparison of per-polyp sensitivities between blinded colonoscopy, consensus reading, and both observers (colonoscopy vs consensus reading, colonoscopy vs observer 1, colonoscopy vs observer 2, consensus reading vs observer 1, consensus reading vs observer 2, and observer 1 vs observer 2).

Interobserver variability
Per-segment agreement and prevalence-adjusted bias-adjusted \( \kappa \) values (\( \kappa_p \) values) were calculated (22,23) for lesions 6 mm or larger and lesions 10 mm or larger. Observers were considered to agree if both recorded corresponding lesion(s) in the same segment or if both recorded no findings. Values were interpreted as those for the regular Cohen \( \kappa \) statistic: \( \kappa_p \leq 0.20 \) indicated poor agreement; \( \kappa_p = 0.21–0.40 \), fair agreement; \( \kappa_p = 0.41–0.60 \), moderate agreement; \( \kappa_p = 0.61–0.80 \), good agreement; and \( \kappa_p = 0.81–1.00 \), very good agreement.

Interpretation time
Mean interpretation times and 95% confidence intervals are provided. Differences in evaluation time between the observers were tested for significance by using the paired Student \( t \) test.

GEE analyses were performed with software (SPSS 15 for Windows; SPSS, Chicago, Ill). All other analyses (per-patient sensitivity, per-patient specificity, and interpretation time) were performed with software (SPSS for Windows, version 12.0.2; SPSS).

\( P \) values less than .05 were considered to indicate statistically significant differences.
Results

Patients

Of 468 eligible patients who were scheduled to undergo colonoscopy during the inclusion period, 180 participated in the study. Twelve (7%) of 180 patients were excluded for different reasons, which left 168 patients for analysis (Fig 1, Table 2).

<table>
<thead>
<tr>
<th>Table 4. Per-Patient Sensitivity and Specificity according to Polyp Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
</tr>
<tr>
<td>Colonoscopy</td>
</tr>
<tr>
<td>Consensus</td>
</tr>
<tr>
<td>Observer 1</td>
</tr>
<tr>
<td>Observer 2</td>
</tr>
<tr>
<td><strong>Specificity</strong></td>
</tr>
<tr>
<td>Consensus</td>
</tr>
<tr>
<td>Observer 1</td>
</tr>
<tr>
<td>Observer 2</td>
</tr>
</tbody>
</table>

Note - Data are number of patients and sensitivity with in parenthesis 95% confidence intervals.

a colonoscopy detected significant more patients with lesions of all sizes than observer 1 and 2 (p<0.001, p<0.001).
b colonoscopy detected significant more patients with lesions 6-9mm compared to observer 2 (p=0.012).
c colonoscopy detected significant more patients with lesions ≥6mm than observer 2 (p=0.008).

Colonoscopy

Blinded and unblinded colonoscopic results revealed that 45 (27%) of 168 patients had a polyp 6 mm or larger and 17 (10%) patients had a polyp 10 mm or larger. Table 3 shows results of histological and morphological characterization, subdivided for lesions 6–9 mm and 10 mm or larger. No colorectal carcinomas were found.
Blinded colonoscopy revealed a lesion 6 mm or larger in 41 (91%) of 45 patients and a polyp 10 mm or larger in 15 (88%) of 17 patients (Table 4).

CT colonography
The mean scanner-room examination time for patients was 21 minutes (range, 13–48 minutes). Buscopan was administered in 144 patients, glucagon was administered in 21, and three patients received no spasmolytic agent. Eighty-eight patients were scanned with 50 mAs, and 80 patients were scanned with 70 mAs. The average amount of carbon dioxide insufflated was 3.9 L (range, 2.0–8.0 L). No complications occurred.

Interpretation time
Observer 1 needed significantly more time to evaluate a complete study than observer 2: a mean interpretation time of 18 minutes (95% confidence interval: 17, 18 minutes) versus a mean interpretation time of 13 minutes (95% confidence interval: 12, 14 minutes) (P < .001).

Per-patient analysis
No patient had more than one polyp 10 mm or larger. The consensus reading resulted in 34 (76%) of 45 patients with a lesion 6 mm or larger, 20 (71%) of 28 patients with a 6–9-mm lesion, and 14 (82%) of 17 patients with a lesion 10 mm or larger (Table 4, Fig 2). Specificity was 79% (97 of 123) for the identification of
No significant differences in detection rates were observed between the consensus reading (CT colonography) and colonoscopy (all P values, ≥ .070). Colonoscopy did help identify significantly more patients with a lesion of any size than did observers 1 (P < .001) and 2 (P < .001) and significantly more patients with a lesion 6–9 mm (P = .012) or 6 mm or larger (P = .008) than did observer 2. No other significant differences in detection rates were found between colonoscopy and separate observers (all P values, ≥ .057). Furthermore, no significant differences in sensitivities or specificities were observed between consensus reading and observers (sensitivity: all P values, ≥ .063; specificity: all P values, ≥ .092) or between separate observers (sensitivity: all P values, ≥ .267; specificity: all P values, ≥ .115). The consensus reading resulted in 34 (76%) of 45 patients with a polyp 6 mm or larger, in comparison to 29 (64%) of 45 patients for observer 2 (P = .267); no increase was
observed compared with observer 1. With regard to larger lesions, the consensus reading resulted in the identification of 14 (82%) of 17 patients with a polyp 10 mm or larger, in comparison to 13 (76%) of 17 patients for observer 1 (P > .99) and 12 (71%) of 17 patients for observer 2 (P = .50). Table 5 shows negative and positive predictive values.

**Per-polyp analysis**

With regard to the three missed lesions 10 mm or larger (two flat polyps—one in the transverse and one in the ascending colon—and one sessile polyp in the ascending colon), no perception or technical errors were observed for the consensus reading (Table 6, Fig 3). Two (33%) of the six false-positive findings for lesions 10 mm or larger were matched to polypoid lesions at colonoscopy: an inverted appendix stump (Fig 4) and a mucosal bleb (cecum); one finding proved to be a lipomatous ileocecal valve, and three findings were attributed to poorly tagged stool. In general, detection rates for colonoscopy were better for lesions 6 mm or larger and 6–9 mm lesions (Table 6).

**Figure 4. False-positive finding ≥10mm at CTC**

Figure displays a false-positive finding (arrow) at CTC (axial image right and 3D image in the middle) that had a true polypoid morphology at endoscopy (left) but proved to be an inverted appendix stump.

For the other size thresholds, no significant differences in per-polyp sensitivity were found between CT colonography (consensus or observer 1 or observer 2) and colonoscopy (all P values, ≥.101), except for observer 2, whose detection rates were lower for lesions 7 mm or larger (P = .016). All data on per-polyp sensitivity values, with significant corresponding P values for all comparisons (colonoscopy vs consensus reading, colonoscopy vs observer 1, colonoscopy vs observer 2, consensus reading vs...
Accuracy of limited prepped CT colonography

Observer 1, consensus reading vs observer 2, and observer 1 vs observer 2), are provided in Table 6.

Table 6. Polyp Sensitivity and Number of False-Positive Findings according to Size

<table>
<thead>
<tr>
<th>Sensitivity (%)</th>
<th>All sizes</th>
<th>6-9mm</th>
<th>≥6mm</th>
<th>≥7mm</th>
<th>≥8mm</th>
<th>≥9mm</th>
<th>≥10mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colonoscopy</td>
<td>384/410</td>
<td>37/39</td>
<td>52/56</td>
<td>137/40</td>
<td>24/26</td>
<td>15/17</td>
<td>15/17</td>
</tr>
<tr>
<td></td>
<td>94 (90-96)</td>
<td>95 (81-99)</td>
<td>93 (92-97)</td>
<td>93 (92-97)</td>
<td>88 (74-98)</td>
<td>88 (63-97)</td>
<td>88 (63-97)</td>
</tr>
<tr>
<td>Consensus</td>
<td>n/a</td>
<td>25/39</td>
<td>39/56</td>
<td>70 (56-80)</td>
<td>83 (68-91)</td>
<td>14/17</td>
<td>14/17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>64 (49-77)</td>
<td>90 (82-97)</td>
<td>97 (82-97)</td>
<td>85 (66-94)</td>
<td>82 (57-94)</td>
<td>82 (57-94)</td>
</tr>
<tr>
<td>Observer 1</td>
<td>134/410</td>
<td>27/39</td>
<td>40/56</td>
<td>32/40</td>
<td>21/26</td>
<td>13/17</td>
<td>13/17</td>
</tr>
<tr>
<td></td>
<td>33 (26-40)</td>
<td>69 (52-82)</td>
<td>71 (57-83)</td>
<td>80 (65-90)</td>
<td>76 (62-92)</td>
<td>76 (51-91)</td>
<td>76 (51-91)</td>
</tr>
<tr>
<td>Observer 2</td>
<td>108/410</td>
<td>19/39</td>
<td>31/56</td>
<td>27/40</td>
<td>19/26</td>
<td>12/17</td>
<td>12/17</td>
</tr>
<tr>
<td></td>
<td>26 (21-33)</td>
<td>49 (36-62)</td>
<td>55 (43-67)</td>
<td>68 (52-80)</td>
<td>73 (53-87)</td>
<td>71 (46-87)</td>
<td>71 (46-87)</td>
</tr>
<tr>
<td>FP findings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consensus</td>
<td>n/a</td>
<td>42</td>
<td>17</td>
<td>13</td>
<td>10</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Observer 1</td>
<td>184</td>
<td>44</td>
<td>19</td>
<td>14</td>
<td>11</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Observer 2</td>
<td>208</td>
<td>39</td>
<td>18</td>
<td>15</td>
<td>13</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

Note - Data are number of patients and sensitivity with in parenthesis 95% confidence intervals.

a colonoscopy detected significant more lesions of all sizes than observer 1 and 2 (p<0.001, p<0.001).
b observer 1 detected significant more lesions of all sizes than observer 2 (p=0.008).
c colonoscopy detected significant more lesions 6-9mm than consensus, observer 1 and 2 (p=0.006, p=0.007, p<0.001).
d consensus and observer 1 detected significant more lesion 6-9mm than observer 2 (p=0.009, p=0.024).
e colonoscopy detected significant more lesions ≥6mm than consensus, observer 1 and 2 (p=0.008, p=0.008, p<0.001).
f consensus and observer 1 detected significant more lesions ≥6mm than observer 2 (p=0.003, p=0.024).
g colonoscopy and consensus detected significant more lesions ≥7mm than observer 2 (p=0.016, p=0.010).

Interobserver agreement

Observers 1 and 2 recorded lesion(s) 10 mm or larger in 19 and 23 segments, respectively: Concordant findings were reported in 11 segments. No findings 10 mm or larger were recorded in 966 segments by both observers.
Chapter 3

This resulted in a per-segment agreement of 98% (977 of 1000 segments) and a corresponding very good \( \kappa_p \) value of 0.96. For lesions 6 mm or larger, the per-segment agreement was 94% (940 of 1000 segments), with an associated \( \kappa_p \) value of 0.88.

**Image quality**

Four (2%) of 172 evaluated studies were considered non-diagnostic because of poor tagging and were therefore excluded from analysis (Fig 1). The radiologist evaluated stool tagging as good in 160 (95%) of the 168 included studies and as adequate in seven (4%) of 168 studies. In one (1%) of 168 studies, image quality was rated as poor, with untagged solid stool less than 6 mm. Figure 5 shows different tagging examples.

| Table 7. Burden of diarrhea and overall burden of the bowel preparation (n=168) |
|------------------|------|------|------|------|------|
|                  | None | Mild | Moderate | Severe | Extreme |
| Burden of diarrhea* | 17   | 48   | 44     | 23     | 20     |
| Overall burden#    | 54   | 69   | 28     | 12     | 3      |

*11 patients experienced no diarrhea, 5 patients did not fill out the questionnaire with regard to the burden of diarrhea.

*2 patients did not fill out the questionnaire with regard to the overall burden of the bowel preparation.

**Patient experience and preference**

Diarrhea was present in 152 (93%) of 163 patients who filled out the questionnaire with regard to diarrhea (Table 7). From among the 165 patients who returned the questionnaire after 5 weeks, 144 (87%) rated the bowel preparation for colonoscopy as more burdensome than the bowel preparation for CT colonography. With regard to patient preference, 114 (70%) of 164 patients preferred CT colonography to colonoscopy, 13 (8%) were indifferent, and 36 (22%) favored colonoscopy as a colonic examination.
Discussion

Our study results had a sensitivity for the consensus reading of 76% (34 of 45) and 82% (14 of 17) for patients with a colorectal polyp 6 mm or larger and those with a polyp 10 mm or larger, respectively. Detection rates were higher for colonoscopy, for other patients with a polyp 6 mm or larger (41 [91%] of 45) and patients with a lesion 10 mm or larger (15 [88%] of 17), than those at CT colonography, but this

Figure 5. Stool tagging evaluation

Figure shows different levels of homogeneity of stool tagging. 5a was evaluated as good stool tagging; a 15mm adenomatous polyp is visible in the sigmoid colon (arrow); 5b shows adequate tagging; 5c is an example of poor tagging in the sigmoid but the examination was considered diagnostic for lesions ≥6mm; 5d shows poor tagging and this patient was excluded because our reviewers rated this examination as non-diagnostic.
difference was not statistically significant. CT colonography correctly depicted 97 (79%) of 123 patients without polyps 6 mm or larger and 146 (97%) of 151 patients without polyps 10 mm or larger. For these size thresholds, negative predictive values were 90% (97 of 108) and 98% (146 of 149), and positive predictive values were 57% (34 of 60) and 74% (14 of 19) for polyps 6 mm or larger and polyps 10 mm or larger, respectively. Interobserver agreement per segment was 94% (940 of 1000) and 98% (977 of 1000) for lesions 6 mm or larger and lesions 10 mm or larger, respectively.

Our ability to identify patients with lesions 6 mm or larger and those with lesions 10 mm or larger was not as good as in the study by Iannaccone et al [19], in which sensitivity ranged from 90% (44 of 49) for patients with a polyp 6 mm or larger to 100% (24 of 24) for patients with a polyp 10 mm or larger. These differences in accuracy could be explained by interpretation problems, but could also possibly be at least partly explained by visualization problems. An argument in favor of an interpretation problem is that the observers in our study were less experienced (70 and 120 cases) than those in the other study (100, 200, and 300 cases, respectively, for the three observers). Although we believe that a higher level of experience might have improved the detection of medium-size (6–9-mm) polyps, this is not an adequate explanation for the difference with regard to the detection rates of lesions 10 mm or larger. Namely, no lesions 10 mm or larger that were missed at the consensus reading were visible in retrospect. Therefore, the problem with the identification of polyps 10 mm or larger seems to be a visualization rather than an interpretation problem.

There are two major differences between both studies that can possibly explain these visualization errors. First, we used a different bowel preparation with a lower dose of diatrizoate meglumine but added barium and bisacodyl. In theory, inhomogeneous tagging might result in the masking of true lesions. However, stool tagging was considered good in 160 (95%) of 168 of the studies, including the cases with missed lesions, and only four studies were excluded because of poor stool tagging. Therefore, we do not believe that our tagging regimen was the cause of the visualization errors. Second, we investigated only patients at increased risk for colorectal cancer, while the other study had a mixed population, of which 106 (52%) of 203 were high-risk patients. Recently, MacCarty et al [20] suggested that patients at increased risk who regularly undergo colonoscopy tend to have lesions that are hard to detect with CT colonography because easy-to-see lesions are detected and removed at previous colonoscopic examinations, while less conspicuous lesions
Accuracy of limited prepped CT colonography

remain in situ. Van Gelder et al [5], who studied a patient group comparable to ours, found one-third (14 of 48) of polyps 10 mm or larger to be flat, and 71% (10 of 14) of these flat polyps were missed at CT colonography. In our study, four (24%) of the 17 polyps that were 10 mm or larger were flat, and two (50%) of these were missed. Our results underscore the findings of MacCarty et al; this indicates that CT colonography with fecal tagging might therefore perform better in screening populations (e.g., patients at average risk who have not previously undergone screening with colonoscopy). The relatively high number of 42 false-positive findings by using 6 mm as a size threshold was somewhat disappointing because fecal tagging is supposed to be associated with a lower rate of false-positive findings because stool is not mistaken for polyps [15,17,24]. The actual polyp size threshold for colonoscopy referral in a screening setting is still debated, but many institutions consider a size threshold of 6 mm clinically relevant.

An explanation for our high number of false-positive findings in this category is that although stool tagging was homogeneous, the average volume-rendering effect at a section thickness of 3.2 mm might have caused small (6- or 7-mm) pieces of stool to look like small polyps. This explains the sharp decline in the number of false-positive findings at a size threshold of 8 mm or larger, which can be deduced from the fact that only 13 false-positive findings 8 mm or larger were observed, and six of these were 10 mm or larger. Furthermore, the observers were relatively inexperienced, which, in our opinion, undoubtedly contributed to the number of false-positive findings. As mentioned before, we did not use stool subtraction software, which might have helped lower the number of false-positive findings. Despite the high number of medium-sized false-positive findings, specificity was 79% (97 of 123) for patients with a lesion 6 mm or larger and increased to 94% (136 of 145) at a size threshold of 8 mm. Furthermore, negative predictive values were high (lesions ≥6 mm, 90% [97 of 108]; lesions ≥10 mm, 98% [146 of 149]). We applied a double-reading strategy (consensus reading for discrepant lesions) for lesions 6 mm or larger. Although this is a time-consuming and costly strategy, it has been suggested to improve sensitivity [4]. Although a higher detection rate for the consensus reading compared with that for the separate observers was observed, no significant improvement was found by using double reading (P ≥.063). Alternatives to a double reading by radiologists are computed-aided detection [25,26] or the deployment of radiographers [27,28].

The radiologist needed significantly (P < .001) more time to review the cases in comparison to the radiology fellow (18 vs 13 minutes). An explanation is that the research fellow had more experience in CT colonography prior to the study than the
radiologist (120 vs 70 cases). However, the longer review time might also reflect a more careful evaluation of the data by the radiologist. This can be substantiated by the fact that higher sensitivities were calculated for the radiologist than for the research fellow. Some limitations of our study should be considered. A cohort of 168 patients at increased risk for colorectal carcinoma was evaluated in our study. Our results can therefore not be extrapolated to a screening setting of patients at average risk for developing colorectal cancer. Two observers were used in our study, and although interobserver variability was high, the use of two observers might limit generalizibility of our data. Relatively short training of 70 cases for the radiologist and 120 cases for the radiology fellow was applied, and this is probably not enough to achieve a high level of competence with a primary two-dimensional evaluation method [29]. We used a section thickness of 3.2 mm with our four-section CT scanner to enable scanning in one breath hold, and thin collimation is associated with better accuracy in the clean colon [30–32]. Presumably, thin collimation is also important in the limitedly prepared colon because inhomogeneously tagged stool can be better differentiated from polyps, which subsequently lowers false-positive rates and improves detection of particularly flat polyps. Finally, a disadvantage of using a limited bowel preparation for CT colonography in comparison to using a cathartic bowel preparation is that same-day colonoscopy for polyp removal cannot be performed.

In conclusion, we report that CT colonography with limited bowel preparation in a population at increased risk for colorectal cancer had a sensitivity and specificity of 82% (14 of 17) and 97% (146 of 151), respectively, for patients with polyps 10 mm or larger. These results are in line with those of performance studies with extensive bowel preparation [4,5]. Because a cathartic bowel preparation could potentially be a deterrent to screening, we believe that limited bowel preparation can be of value for the implementation of CT colonography as a screening technique for colorectal polyps and cancer.

Acknowledgments

The authors thank Jasper Florie, MD, Karin Horsthuis, MD, and Henk W. Venema, PhD, for their critical review of the manuscript. We thank C. Yung Nio, MD, for his critical review of the CT colonographic cases with missed lesions and Koos A. H. Zwinderman, PhD, for his valuable contribution with regard to the statistical analyses (GEE) of our data.
Reference list