CT colonography in faecal occult blood test positives
Liedenbaum, M.H.

Citation for published version (APA):
Liedenbaum, M. H. (2010). CT colonography in faecal occult blood test positives
Evaluation of a standardized CT colonography training program in novice readers

Marjolein H. Liedenbaum
Shandra Bipat
Patrick M.M. Bossuyt
Roy S. Dwarkasing
Margriet C. de Haan
Roel J. Jansen
Dominique Kauffman
Christiaan van der Leij
Manou S. de Lijster
Cindy C. Lute
Marije P. van der Paardt
Maarten Thomeer
IJsbrand A. Zijlstra
Jaap Stoker

Submitted
Abstract

**Purpose:** To evaluate a standardized CT colonography (CTC) training program in novice readers.

**Methods:** Six physicians (one radiologist, three residents and two radiology research-fellows) and three technicians started a CTC training program. A series of 200 CTC’s with colonoscopic verification was selected from an existing research database, with 100 normal cases and 100 cases with at least one polyp ≥6 mm. CTC reading was done individually with immediate feedback of the colonoscopy outcome after each CTC. The increase in per-polyp sensitivity was calculated per 50 CTC’s, for lesions ≥6mm. Using logistic regression analyses, the number of CTCs to reach 90% sensitivity for lesions ≥6mm was estimated. Reading times were registered.

**Results:** The average per-polyp sensitivity for lesions ≥6mm was 77% in the first set of 50 CTCs, 77% in the second set (p=0.99 vs. first set), 80% in the third set (p=0.26 vs. second set) and 91% in the fourth set (p<0.001 vs. the third set). The estimated number of CTCs to be evaluated to reach a 90% per polyp sensitivity for lesions ≥6mm was 164. Six of the nine readers reached this level of competence within 175 CTCs. Reading times decreased significantly from the first 50 CTC’s to the second 50 CTCs in 6 readers.

**Conclusions:** Novice CT colonography readers can obtain a sensitivity equal to that of experienced readers after practicing on average 164 CTCs. For most readers 175 CTCs are sufficient for adequate training, while some readers will need additional training after this.
INTRODUCTION

The evaluation of computed tomography colonography (CT colonography) examinations is preferably performed by experienced readers. It has been shown that carcinoma and polyp detection improve with practice.\textsuperscript{1-7} So far, no consensus exists about the level of experience and how the desired level can be acquired best. Several training programs have been developed, relying on different methods for training. These include reading of CT colonographies with colonoscopy feedback, peer reviewed reading of CT colonography, and a course with presentations of cases and instructions on the use of CT colonography software. The number of CT colonography examinations that has been recommended for adequate training varies from 50 to more than 100.\textsuperscript{2,5,7}

Training in interpreting radiological images is necessary to reduce errors in detection of abnormalities. Three types of errors are recognized: errors of search (the radiologists’ gaze completely misses the abnormality), errors of detection (the eyes of the radiologist pass over the abnormality but not long enough to be recognized) and errors of decision (the abnormality is not correctly identified).\textsuperscript{8} The first two types can be avoided when a reader gains experience in examining the specific exam. Errors of decision can be decreased when a set of examples of true and false positives is provided to a radiologist for training.

One study was using a teaching file of 40 CT colonographies in training radiologists and technologists. The investigators found that the trainees performed worse than experienced radiologists.\textsuperscript{5} Another study assessed the training effect in medical students and technologists using a teaching file of 50 cases. After this training the participants in the training program interpreted 50 cases, which resulted in a sensitivity and specificity similar to that of experienced radiologists.\textsuperscript{9} In yet another study novice readers were presented 100 CT colonography learning cases. The authors of this study concluded that computer aided detection can improve sensitivity in the first 20 cases.\textsuperscript{4}

So far, no study has tried to estimate the number of CT colonography examinations to be evaluated by physicians and technicians to reach a sufficient level of competence. There are some colonoscopy studies that have evaluated a learning curve in gastroenterology residents.\textsuperscript{10,11} One study found that for a sufficient competence in screening and diagnostic colonoscopy, an experience of more than 150 procedures is required.\textsuperscript{10} During the training period the cecal intubation time decreased and the success rate for completion of the procedure improved significantly, but the polyp detection rate remained equal. The aim of our study presented was primarily to determine how many CT colonography training examinations have to be evaluated by novice readers to obtain an adequate level of competence in polyp detection.
METHODS

Philips Medical Systems (Best, the Netherlands) provided CT colonography software and four workstations for this study. The authors had full control of all data and information submitted for publication.

Patient selection
Cases were selected from an existing database of 302 fecal occult blood test (FOBT) positive individuals between 50 and 75 years old, who had participated in a study reported in detail elsewhere. The study had been approved by the local Medical Ethics Committee. Informed consent was obtained from all participants, including informed consent for additional studies with the CT colonography examinations. All FOBT positive participants had received a bowel preparation of meglumine-ioxithalamate (Telebrix Gastro 300 mg I/ml; Guerbet, Cedex, France) starting two days (in total 7*50 ml of Telebrix Gastro) or one day (in total 4*50 mL of Telebrix Gastro) before CT colonography examination. CT colonography had been performed with automatic insufflation of CO₂ (Bracco, PROTOCO2L insufflator, New York, USA) and scans were made in prone and supine positions. A low dose scan protocol was used on a 64-slice CT scanner (Brilliance, Philips Medical Systems, Best, the Netherlands) with 120 kV, pitch 1.2, slice thickness 0.9 mm, rotation time 0.4s, 40 ref mAs and dose modulation or on another 64-slice scanner (SOMATOM Sensation, Siemens Medical Solutions, Erlangen, Germany) with 120 kV, 1.4, 0.5s and 32 reference mAs with dose-modulation. All CT colonography examinations had been read prospectively by two of seven experienced observers, who had read between 125 and 700 CT colonography interpretations with colonoscopic verification. Observers had marked all polyps and indicated the morphology, size and location. All CT colonographies of insufficient quality due to insufficient distension or inhomogeneous tagging of the feces judged by the experienced readers were excluded.

All participants had undergone a subsequent colonoscopy with segmental unblinding within 2 weeks after CT colonography. A radiology researcher [ML] had matched all CT colonography polyps with the found colonoscopy polyps based on criteria for size, location and morphology matching.

Case selection
A radiology CT colonography research fellow [ML] selected 200 CT colonography training cases from the series of 302. Cases were arranged in groups of 50 cases with equal lesion prevalence. In each group of 50 cases there were 25 normal examinations, i.e. without lesions ≥6 mm (i.e. polyps and carcinomas) at colonoscopy, and 25 examinations with lesions ≥6 mm. The 25 examinations with lesions ≥6 mm included approximately 15 examinations with one or more polyps of ≥10 mm and 10 examinations with one or more lesions between 6 and 9 mm. Table 1 summarizes the actual number of lesions per group of 50 cases, as well as the sensitivity of the double read of the two experienced readers in the previous study.
Table 1: Distribution of lesions among the 200 cases

<table>
<thead>
<tr>
<th>Number of lesions</th>
<th>1st 50 cases</th>
<th>2nd 50 cases</th>
<th>3rd 50 cases</th>
<th>4th 50 cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without lesions ≥6mm</td>
<td>25</td>
<td>26</td>
<td>25</td>
<td>23</td>
</tr>
<tr>
<td>Lesions ≥10mm</td>
<td>15</td>
<td>16</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>Lesions only between 6-9mm</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Number of lesions ≥10mm</td>
<td>9 pedunculated 7 sessile 2 flat</td>
<td>14 pedunculated 6 sessile</td>
<td>12 pedunculated 5 sessile 3 flat</td>
<td>11 pedunculated 4 sessile</td>
</tr>
<tr>
<td>Number of polyps 6-9mm</td>
<td>3 pedunculated 10 sessile 2 flat</td>
<td>6 pedunculated 10 sessile 5 flat</td>
<td>3 pedunculated 19 sessile 2 flat</td>
<td>4 pedunculated 13 sessile</td>
</tr>
<tr>
<td>Carcinomas</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Sensitivity experienced readers (double read)</td>
<td>*Lesions ≥10mm 95% 95% 86% 94%</td>
<td>*Lesions ≥6mm 83% 88% 93% 89%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Lesions include polyps (adenomatous and hyperplastic) and carcinomas.

CT colonography training program

Nine novice readers started and completed the CT colonography training program (see Table 2 for levels of radiological experience). They were one radiologist, three radiology residents, two radiology research-fellows and three technicians. None of the readers had any previous experience with CT colonography reading.

Table 2: Experience of participating readers

<table>
<thead>
<tr>
<th>Reader</th>
<th>Experience in Radiology Department</th>
<th>FP training?²</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>12 years</td>
<td>No</td>
</tr>
<tr>
<td>R2</td>
<td>3 years</td>
<td>Yes</td>
</tr>
<tr>
<td>R3</td>
<td>4 years</td>
<td>Yes</td>
</tr>
<tr>
<td>R4</td>
<td>1 year</td>
<td>No</td>
</tr>
<tr>
<td>R5</td>
<td>½ year</td>
<td>No</td>
</tr>
<tr>
<td>R6</td>
<td>0 years</td>
<td>Yes</td>
</tr>
<tr>
<td>T1</td>
<td>3 years</td>
<td>Yes</td>
</tr>
<tr>
<td>T2</td>
<td>8 years</td>
<td>No</td>
</tr>
<tr>
<td>T3</td>
<td>33 years</td>
<td>Yes</td>
</tr>
</tbody>
</table>

¹Radiologist, radiology residents and research fellows; ²Training in polyps and pitfalls

The training program started with reading two articles on pitfalls in CT colonography.¹⁵,¹⁶ Then a basic course was given, with presentations on anatomy and pathology of the colon, CT colonography bowel preparation, performance of the examination, CT colonography viewing methods and instructions on how to use the viewing software (see description of the course in Fig. 1). After this, all participants received an individual training of four CT
colonography cases by an experienced radiology CT colonography research fellow [M.L.], mainly for instructions in using the software.

Half of the readers - 2 technicians and 3 physicians - received an additional hour of training in pitfalls and polyps. The other group - 1 technician and 3 physicians - did not receive the additional training. The reader groups were matched on age and gender. The additional training consisted of 40 images of polyps, stool, lipomas, carcinomas, folds, diverticulas and veins. The correct lesion type had to be selected, using multiple choice questions. After each question, the correct answer was given immediately.

Then 200 CT colonography cases were read individually on a workstation with specialized software (Philips View Forum, Version 6.2.2, Best, the Netherlands). A primary 3D read with 2D problem solving was done by all readers using an enhanced 3D viewing method (unfolded cube) after electronic cleansing. At the time of study the cleansing algorithm was not yet commercially available (View Forum, Philips Medical Systems, Best, the Netherlands). All readers performed a secondary 2D reading for evaluation of collapsed segments and uncleansed areas (if present). Readers recorded the level of certainty (0, 25, 50, 75 or 100%), diameter and location (cecum, ascending, transverse, descending, sigmoid colon or rectum) of all lesions found. The polyp diameter was measured in the 2D MPR setting using electronic calipers. All readers recorded the time used for examining the primary 3D and 2D read per position (supine and prone). After each case the reader was invited to verify the reading results by checking the colonoscopy matched lesions, which were also available on the workstation after finalizing the reading of that case. A radiology CT colonography research fellow [ML] was available for answering questions on false negative or false positive polyps. The last 25 cases were presented as an exam, so no feedback was available for the readers for these cases.

**Analysis of lesion detection**

Only lesions registered with a certainty level greater than or equal to 50% were taken into account for analysis. All lesions found by the novice readers that matched the previously found CT colonography lesions verified at colonoscopy were considered as true positives. Lesions reported by the novice readers that did not match the lesions verified at colonoscopy were considered as false positives. When a novice reader did not report a lesion in a case, and no lesion was found at colonoscopy, that was defined as a true negative. When the readers did not report a lesion, and lesions had been found at colonoscopy which had not been seen by any of the expert readers on CT colonography, these were also considered true negatives, since we wanted to compare novice readers to experienced readers. Most of these lesions were technical false negatives, i.e. retrospectively not visible (see reference 18). Lesions that were perceptive errors, i.e. retrospectively visible, in the original study were considered as true positive when found by the novice readers. A per-polyp analysis was performed to calculate sensitivity for lesions ≥10 mm and ≥6 mm. A per-patient analysis was performed for calculation of the specificity for lesions ≥6 mm.
We a priori determined a required level of sufficient polyp detection. This level was set at a sensitivity of 95% for lesions ≥10 mm, 90% for lesions ≥6 mm and a per-patient specificity of 80% for lesions ≥6 mm. We chose these levels because they corresponded to the results of the individual experienced observers from our former study, when excluding the technical false negatives (i.e. retrospectively not visible) for calculation of the sensitivity.

Statistical analysis
Previous studies showed that a learning curve of 50 CT colonographies is not sufficient and that there can be an improvement in performance even after 100 CT colonographies. We estimated that 200 CT colonography examinations were necessary to evaluate whether novice readers could achieve the desired sensitivity level after the learning series.

We calculated the sensitivity and specificity for each set of 50 cases, for each reader separately, and for all readers combined. Changes in average sensitivity and specificity between consecutive sets of 50 cases were evaluated for statistical significance.
using the Chi-square test statistic. We also compared the average sensitivity and specificity per set of 50 cases of the group of readers who received the additional training to the group of readers who did not receive this training using the Chi-square test.

For each reader a learning curve was estimated using the outcomes of per-polyp sensitivity for lesions ≥10 mm and lesions ≥6 mm (including colorectal cancers, adenomas and hyperplastic polyps) and per-patient specificity for lesions ≥6 mm. This curve was calculated using logistic regression analysis. We used restricted cubic splines to select the most appropriate functional relation between number of cases evaluated and sensitivity and specificity. We tested for a learning effect by comparing a null model (intercept only) with a model that included one or more coefficients for the number of cases evaluated, using the generalized likelihood ratio test statistic. For each reader separately and for the average of all readers the number of cases necessary to reach the desired level of competence for sensitivity and specificity was estimated based on the fitted logistic regression model.

Differences in reading time were calculated by comparing mean reading times in consecutive sets of 50 cases and comparison of outcomes were done using the student-T-test. Statistical analyses were performed using SPSS version 15.0.1 for Windows (SPSS). For all analysis, a p-value of <0.05 indicated a significant difference between groups.

RESULTS

Learning curve CT colonography

Per polyp sensitivity

The average sensitivity for lesions ≥6 mm in the first set of 50 cases was 77% (95% CI: 72 to 82). In the second, third and fourth set of 50 cases the average sensitivity was 77% (95% CI: 72 to 81), 80% (95% CI: 76 to 84) and 91% (95% CI: 87 to 94) respectively. A significant difference in sensitivity was only found between the third and fourth set of 50 cases (p<0.001) but not between the first and the second (p=0.99), and between the second and the third set (p=0.26). In Table 3 the sensitivity for all readers is presented.

When we calculated the average sensitivity for lesions ≥6 mm per 50 cases for the group of readers who had received the additional pitfalls training we found sensitivities of 77% (95% CI: 71 to 84), 75% (95% CI: 69 to 81), 79% (95% CI: 73 to 84) and 90% (95% CI: 85 to 95) respectively. For the readers who did not receive the additional training, sensitivities were 76% (95% CI: 68 to 83), 78% (95% CI: 72 to 85), 82% (95% CI: 76 to 88) and 91% (95% CI: 87 to 96) for the four sets of 50 cases respectively.

When comparing the sensitivities per sets of 50 cases of the readers who had received the additional training to the readers who did not receive training we found p-values of p=0.77, p=0.51, p=0.41 and p=0.33 respectively.
### Table 3 Per polyp sensitivity

<table>
<thead>
<tr>
<th>Reader</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lesions ≥6mm</td>
<td>Lesions ≥10mm</td>
</tr>
<tr>
<td></td>
<td>Cases</td>
<td>Lesions ≥6mm</td>
</tr>
<tr>
<td>R1</td>
<td>80 76 84 90</td>
<td>85 100 100 100</td>
</tr>
<tr>
<td></td>
<td>(69-100) (84-100) (82-100) (79-100)</td>
<td>(90-100) (90-100) (81-100)</td>
</tr>
<tr>
<td>R2</td>
<td>80 79 77 100</td>
<td>90 100 100 100</td>
</tr>
<tr>
<td></td>
<td>(77-100) (84-100) (82-100) (79-100)</td>
<td>(90-100) (90-100) (81-100)</td>
</tr>
<tr>
<td>R3</td>
<td>77 74 86 91</td>
<td>80 95 100 100</td>
</tr>
<tr>
<td></td>
<td>(62-98) (86-100) (82-100) (79-100)</td>
<td>(73-99) (83-100) (75-100) (88-100)</td>
</tr>
<tr>
<td>R4</td>
<td>73 89 86 97</td>
<td>85 100 100 100</td>
</tr>
<tr>
<td></td>
<td>(69-100) (84-100) (82-100) (79-100)</td>
<td>(78-200) (90-100) (75-100) (88-100)</td>
</tr>
<tr>
<td>R5</td>
<td>77 82 84 100</td>
<td>85 95 100 100</td>
</tr>
<tr>
<td></td>
<td>(69-100) (86-100) (92-100) (79-100)</td>
<td>(83-100) (83-100) (70-96) (75-100)</td>
</tr>
<tr>
<td>R6</td>
<td>83 79 70 84</td>
<td>90 95 84 93</td>
</tr>
<tr>
<td></td>
<td>(77-100) (86-100) (68-100) (82-100)</td>
<td>(78-100) (83-100) (75-100) (59-93)</td>
</tr>
<tr>
<td>T1</td>
<td>67 66 77 84</td>
<td>75 86 89 93</td>
</tr>
<tr>
<td></td>
<td>(56-94) (71-100) (76-100) (82-100)</td>
<td>(90-100) (78-100) (81-100) (70-98)</td>
</tr>
<tr>
<td>T2</td>
<td>73 66 74 78</td>
<td>75 90 95 75</td>
</tr>
<tr>
<td></td>
<td>(56-94) (78-100) (85-100) (54-96)</td>
<td>(59-91) (78-100) (70-98) (70-98)</td>
</tr>
<tr>
<td>T3</td>
<td>80 79 84 91</td>
<td>85 100 100 100</td>
</tr>
<tr>
<td></td>
<td>(69-100) (84-100) (92-100) (79-100)</td>
<td>(68-96) (83-100) (59-93) (59-93)</td>
</tr>
<tr>
<td>Average</td>
<td>77 77 80 91</td>
<td>83 96 96 96</td>
</tr>
<tr>
<td></td>
<td>(78-89) (93-99) (94-99) (93-99)</td>
<td>(81-89) (84-92) (82-91) (81-90)</td>
</tr>
</tbody>
</table>

Percentages indicate sensitivity. Numbers between brackets are 95% Confidence Intervals.

The sensitivity for lesions ≥10 mm was 83% (95% CI: 78 to 89), 96% (95% CI: 93 to 99), 96% (95% CI: 94 to 99) and 96% (95% CI: 93 to 99) for the four sets of 50 cases (Table 3). A significant difference was found between the first and second set of 50 cases (p<0.001), but not between later sets (p=0.72 and p=0.76, respectively).

Estimated individual sensitivity learning curves for polyps ≥6 mm are presented in Figure 2. When observing these curves it is visible that almost all readers first had an increase in sensitivity, then a slight decrease or plateau-phase after 80-100 cases whereafter sensitivity increased again.

**Per patient specificity**

The average per patient specificity for lesions ≥6 mm was 85% (95% CI: 81 to 89) for the first set of 50 cases, 88% (95% CI: 84 to 92), 87% (95% CI: 82 to 91) and 86% (95% CI: 81 to 90) for the second, third and fourth set of 50 cases, respectively. See table 3 for individual results of the learning curve specificity for lesions ≥6 mm. No significant differences between subsequent case groups were found. For lesions ≥10 mm the average per patient specificity was 94% (95% CI: 91 to 97), 95% (95% CI: 92 to 97), 96% (95% CI: 94 to 98) and 97% (95% CI: 94 to 99) for the four respective sets of 50 cases. No significant differences between subsequent case groups were found.
Fig. 2 Per polyp analysis, lesions ≥6mm
When we calculated the specificity for lesions $\geq 6$ mm per 50 cases for readers who had received the additional pitfalls training we found specificities of 90% (95% CI: 85 to 95), 93% (95% CI: 89 to 97), 86% (95% CI: 80 to 92) and 85% (95% CI: 79 to 91) respectively. For the readers who did not receive the additional training, specificities were 79% (95% CI: 71 to 86), 81% (95% CI: 74 to 88), 87% (95% CI: 80 to 94) and 87% (95% CI: 80 to 94) for the four sets of 50 cases respectively. When comparing the specificities per sets of 50 cases of the readers who had received the additional training to the readers who did not receive training we found p-values of $p=0.01$, $p=0.05$, $p=0.89$ and $p=0.64$ respectively.

**Number of cases to reach the desired level of competence**

We estimated that, on average, readers would have to read 164 CT-colonography cases to reach a per-polyp sensitivity of 90% for polyps $\geq 6$ mm. In Table 4 the estimates for all individual readers are presented. Three readers (R6, T1 and T2) did not reach the desired level of per-polyp sensitivity within 200 cases. The other 6 readers had reached this level after 131 to 168 cases. When analyzing the per-polyp sensitivity for lesions $\geq 10$ mm we
found that on average 67 cases had to be examined to reach the desired level. Two readers (R1 and T2) did not reach this level. For the per-patient specificity of lesions \( \geq 6 \text{mm} \) we found that all readers, except reader R1, did not improve in specificity during the training program. The likelihood ratio test statistic did not read significance. All readers had a specificity of 80% or more, except reader R1 (see Table 3).

Table 4 Number of cases per reader in order to reach the desired level of competence

<table>
<thead>
<tr>
<th>Reader</th>
<th>Sensitivity lesions ( \geq 6 \text{mm} )¹</th>
<th>Sensitivity lesions ( \geq 10 \text{mm} )²</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>168</td>
<td>57</td>
</tr>
<tr>
<td>R2</td>
<td>146</td>
<td>53</td>
</tr>
<tr>
<td>R3</td>
<td>156</td>
<td>76</td>
</tr>
<tr>
<td>R4</td>
<td>131</td>
<td>57</td>
</tr>
<tr>
<td>R5</td>
<td>131</td>
<td>74</td>
</tr>
<tr>
<td>R6</td>
<td>&gt;200</td>
<td>&gt;200</td>
</tr>
<tr>
<td>T1</td>
<td>&gt;200</td>
<td>195</td>
</tr>
<tr>
<td>T2</td>
<td>&gt;200</td>
<td>&gt;200</td>
</tr>
<tr>
<td>T3</td>
<td>163</td>
<td>61</td>
</tr>
<tr>
<td>Average</td>
<td>164</td>
<td>67</td>
</tr>
</tbody>
</table>

¹Desired level of competence for the sensitivity for lesions \( \geq 6 \text{mm} \) is 90%. ²Desired level of competence for sensitivity for lesions \( \geq 10 \text{mm} \) is 95%. ³Desired level of competence for specificity for lesions \( \geq 6 \text{mm} \) is 80%.

Reading times
Reader R1 did not register all reading times, therefore only the reading times from the other eight observers were analyzed. In fig. 3 all reading times per observer per 50 cases are given. The mean reading time for all observers was 12'58" (SD 4'52"). For the first 50 cases mean reading times for all observers were 16'26" (SD 5'55"), 12'09" (SD 4'04") for the second 50 cases (\( p<0.001 \) when compared to the first 50 cases), 11'30" (SD 4'06") for the third 50 cases (\( p=0.43 \) when compared to the second 50) and 11'48" (SD 3'24") for the last 50 cases (\( p=0.70 \) when compared to the third 50). Reading times decreased significantly from the first 50 CTC’s to the second 50 CTC’s in 6 readers.

DISCUSSION

The results of this study show that novice readers can reach a sensitivity equal to that of an experienced reader after practicing 175 CT colonography training cases with colonoscopy feedback. We estimated that an average number of 164 cases would be needed to reach a desired level of sensitivity in detecting lesions \( \geq 6 \text{mm} \). A few readers did not reach the desired level of sensitivity and specificity after 200 cases and probably would need more training.
It has been generally accepted that for adequate reading of CT colonography a dedicated training program with a sufficient number of training cases has to be followed.\textsuperscript{6} CT colonography is nowadays recognized as a possible screening technique and it can replace the barium enema examination for detection of both polyps and cancer.\textsuperscript{10-21} Consequently, this relatively new examination has become widespread in use and many new novice readers have to start reading CT colonography examinations. From earlier studies it became clear that more than 50 to 100 CT colonography cases were necessary to obtain a sufficient accuracy for lesion detection.\textsuperscript{5,7,9} In our study on average 164 CT colonography training cases were needed to gain a sufficient sensitivity. From a practical point of view, one could round up this number and use a CT colonography training program with 175 colonoscopy verified cases which most likely will suffice in the majority of observers. Three readers, of which two technologists and one radiology research fellow, did not reach this level, and probably need additional training. Because our study was limited to 200 cases we do not know the number of cases they should have to read additionally. It might also be possible that one or more of these readers will never reach an adequate level of competence.

Fig. 3 Reading times

![Reading times graph](Image)

Average reading times presented per reader. One reader (R1) did not record all reading times, therefore this information is missing.

Although we analyzed only a small number of readers, it became clear that different readers in a radiology practice are able to learn CT colonography: radiologists, radiology residents, radiology research fellows and technologists. Earlier studies have also shown that radiology residents and technicians can reach sensitivities similar to that of as radiologists.\textsuperscript{1,22} From our study it became clear that smaller lesions (6 to 9 mm) are more difficult to detect. A significant increase in sensitivity in detecting lesions $\geq$6 mm was only
seen between 150 to 200 cases. For larger lesions, i.e. lesions ≥10 mm, we found an earlier increase, namely between 50 and 100 cases. This illustrates that smaller lesions are less conspicuous and readers need more perceptual learning to reduce the errors of search and detection.\textsuperscript{3,23} Other than errors of search and detection, errors of decision exist. These can be avoided when more knowledge on e.g. anatomy and pathology is obtained. For this reason we started our training program with a series of lectures. About half of the readers also received an additional training program on pitfalls in CT colonography. Some studies that evaluated CT colonography learning used a training program to teach CT colonography pitfalls and true positive lesions.\textsuperscript{2,4,24} In our study we did not find an effect on sensitivity of readers who received this additional training, but we did find an effect on specificity for lesions ≥6mm. The group of readers who did not receive the additional training had a significantly lower specificity in the first and second set of 50 cases than the group that received the additional training. This was however mainly due to one reader (R1) in the first group that had a very low specificity. It is however difficult to draw solid conclusions on this, because we only analyzed a small number of readers in our study.

When using CT colonography it is important to limit the number of false positive cases, because this will induce many unnecessary colonoscopies. We found that only one reader (R1) did not reach the desired level of specificity. All other readers stayed on a nearly constant level of specificity larger than 80% and no increase in learning was observed. This indicates that most readers will not report many false positives when they start reading CT colonography. Reader R1 however might have needed more training after 200 CT colonographies to reach a desired specificity level. A few other studies evaluated specificity in CT colonography training programs. In two studies the specificity increased, but this was observed in a training set of only 50 or 60 CT colonographies in total.\textsuperscript{2,3} In our study a training set of only 50 CT colonographies was not sufficient to gain an optimal sensitivity, thus an increase of specificity within the first series of 50 cases will not be relevant for the desired number of training cases.

Reading times decreased with learning, especially after the first 50 cases. These results are consistent with results of earlier studies. In the study of Hock et al. with 100 CT colonography cases reading times decreased from 15.67 minutes in the first training session (without CAD) to 13.42 minutes in the last session.\textsuperscript{4} In the study of Burling et al. experienced radiologists also reported significantly faster than novice radiologists and technicians.\textsuperscript{25}

Although we performed an extensive study on learning curves in CT colonography using a large training subset of 175 training cases and 25 exam cases, some aspects of training in CT colonography still remain unclear. We used a set of cases with a 50% disease prevalence, optimal for calculating both sensitivity and specificity. An enriched training set with a higher number of lesions might result in a steeper learning curve for sensitivity. The training set in the study of Dachman et al., for example, consisted of 83% abnormal CT colonography cases.\textsuperscript{7} A disadvantage of this method is that specificity is possibly less well trained, which could result in a higher number of false positives and consequently a lower positive predictive value in a daily practice, where lesion prevalence is much lower. Furthermore the use of CAD could influence the learning curve in novice.
This study has some limitations. A first limitation is that only one of the physicians was an abdominal radiologist, the others were residents and research fellows in training. One previous study showed that non radiologists can perform CT colonography equally good as radiologists. In our study one technologist and also most of the radiology fellows and residents performed equally or even slightly better than the radiologist. Because we only had a limited number of readers, differences between different reader groups could not be meaningfully calculated and tested for significance. A second limitation is that no feedback was provided in the last 25 cases, because these cases were considered as an exam in our department. Some of the readers (R1, R6 and T2) performed differently in those last cases, with a slight decrease in sensitivity and/or specificity (not significant). For example more fecal remnants and other polyp-like structures were marked as true lesion than in the previous cases. Thirdly we calculated the sensitivity considering only the true positive lesions of the experienced observers from our earlier study. Therefore the sensitivity percentages reported here cannot be compared to sensitivities reported in other studies. Fourthly, we only evaluated the learning curve using a primary 3D reading paradigm and not a 2D paradigm. Because previous research had shown that novice readers perform better when using primary 3D reading we have chosen to use only this reading paradigm. Fifthly, we did not analyze detection of carcinomas separately, because the number of carcinomas in our dataset was only small (see Table 1). We found it most important that true lesions were detected (either adenomas, carcinomas or hyperplastic polyps). Another limitation is that we evaluated a learning curve on patients that had received an iodine tagging bowel preparation, while using a cleansing algorithm. Therefore results might not be generalized to preparations with cathartic or barium preparations.

**Conclusions**

To conclude we found that CT colonography reading can be adequately performed after dedicated training by inexperienced radiologists, radiology fellows and radiology technicians. An average number of 164 CT colonographies with colonoscopic verification was needed to reach a sensitivity that equals that of an experienced reader. Six of nine readers reached the level of sufficient competence within 175 training cases. A few readers do not reach this level after 200 cases and might need more training or might even not be able to reach this level at all.

**Acknowledgements**

We thank Wouter van Elzelingen of the University of Amsterdam for his contribution to this study. We thank Philips Medical Systems for providing us their workstations for this project.
References