

Can radiographers be trained to triage CT colonography for extracolonic findings?

Thierry N. Boellaard · C. Yung Nio ·
Patrick M. M. Bossuyt · Shandra Bipat · Jaap Stoker

Received: 8 February 2012 / Revised: 1 May 2012 / Accepted: 9 May 2012 / Published online: 3 July 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract

Objectives Radiographers have been shown to be capable CT colonography observers. We evaluated whether radiographers can be trained to triage screening CT colonography for extracolonic findings.

Methods Eight radiographers participated in a structured training program. They subsequently evaluated extracolonic findings in 280 low-dose CT colonograms (cases). This dataset contained 66 cases with possibly important findings (E3) and 27 cases with probably important findings (E4) [classification based on the highest classified finding (C-RADS)]. The first 40 and last 40 CT colonograms were identical test cases. Immediate feedback was given after each reading, except for test cases. Radiographers triaged cases based on C-RADS classification and indicated the need for a radiologist read. We constructed learning curves for correct case triaging by calculating moving averages.

Results In the final test series, 84/120 (70 %) cases with E3 or E4 findings and 139/200 (70 %) without E3 or E4 findings were correctly triaged. Correct identification of cases with E3 findings improved with training from 46/88 (52 %) to 62/88 (70 %) ($P < 0.0001$) but not for E4 findings [both 22/32 (69 %) $P = 1.00$].

Conclusions Radiographers improve after training in correctly triaging extracolonic findings at CT colonography but do not reach a high enough accuracy to consider their structural involvement in screening.

Key Points

- Radiographers were trained to triage CT colonography for extracolonic findings.
- After training, radiographers improved sensitivity for likely unimportant findings.
- After training, radiographers did not improve sensitivity for possibly important findings.
- Radiographers should probably not be expected to identify all extracolonic findings.

Keywords Colonography, computed tomographic · Incidental findings · Learning curve · Radiology/education · Clinical competence

Introduction

Computed tomography (CT) colonography is currently adopted as one of the recommended screening tools for colorectal cancer and advanced adenomas by the American College of Radiology, the American Cancer Society, and the U.S. Multi-society Task Force on Colorectal Cancer [1]. Apart from yield and participation, costs have to be considered in the selection of a screening test. Recently doubts were raised about the cost-effectiveness of CT colonography compared with established screening techniques [2]. A modelling study estimated that CT colonography costs should not exceed 43 % of colonoscopy costs to be cost-effective [3]. One should therefore scrutinise all aspects of CT colonography screening contributing to the total cost, including reading strategies.

A potential approach to reduce costs is reading of CT colonography images by radiographers. Radiographers have

T. N. Boellaard (✉)
Department of Radiology, Academic Medical Center,
University of Amsterdam, Room: G1-235, Meibergdreef 9,
1105 AZ Amsterdam, The Netherlands
e-mail: t.n.boellaard@amc.uva.nl

C. Y. Nio · S. Bipat · J. Stoker
Department of Radiology, Academic Medical Center,
University of Amsterdam, Meibergdreef 9,
1105 AZ Amsterdam, The Netherlands

P. M. M. Bossuyt
Department of Clinical Epidemiology, Biostatistics,
and Bioinformatics, Academic Medical Center,
University of Amsterdam, Meibergdreef 9,
1105 AZ Amsterdam, The Netherlands

been demonstrated to be as accurate as radiologists in reading CT colonograms for intracolonic findings in a single read strategy [4, 5]. When such a strategy is implemented, images still require additional reading for examining extracolonic findings. In the asymptomatic screening population the incidence of potentially important extracolonic incidental findings is estimated at about 10 %, which results in about 2.5 % relevant new diagnoses [6]. With radiographers as readers for the intracolonic findings in screening, having an additional radiologist read for extracolonic findings in all CT colonograms would be less practical.

Radiographers may also be capable of evaluating extracolonic findings after sufficient training. Previous studies have demonstrated that radiographers can accurately read mammograms [7], plain x-rays [8], intravenous pyelograms [9] and double-contrast barium enemas [10]. Much lower accuracy results have been seen in a study in which non-radiologists evaluated paediatric brain CT but no dedicated training was given in that study [11].

Before radiographers can triage CT colonography examinations for extracolonic findings, they should achieve competence. We evaluated the accuracy and learning curve of formally trained radiographers in triaging extracolonic findings in CT colonography.

Methods

A research grant was received from the Nuts Ohra Foundation (Amsterdam, the Netherlands). The Nuts Ohra Foundation was not involved in designing and conducting this study, did not have access to the data, and was not involved in data analysis or preparation of this manuscript.

We trained consenting radiographers to triage CT colonograms for extracolonic findings using a training program comprising two parts: a basic training and 200 feedback training cases (see flow chart, Fig. 1). Before and after the feedback training cases, radiographers triaged the same 40 CT colonograms (initial and final test cases). In this study, the radiographers did not assess CT colonograms for intracolonic findings.

Basic training cases

For the initial training part of the programme (basic training), 100 clinical abdominal CT examinations were used: 40 CT examinations were enhanced with intravenous contrast medium, 54 were unenhanced and 6 were low-dose CT colonograms. One hundred CT examinations were assumed to be the minimum for the radiographers to become acquainted with the most important and/or frequent extracolonic findings (Table 1) [12–14]. All CT examinations had been performed with two CT systems: 4-slice (Mx8000; Philips Healthcare, Best, the Netherlands) and 64-slice (Brilliance, Philips Healthcare) CT

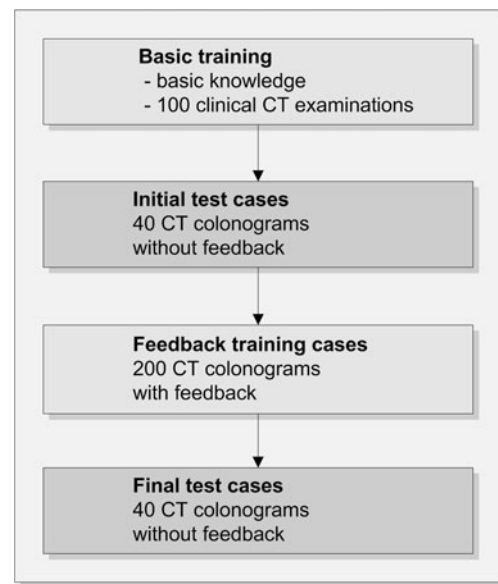


Fig. 1 Flow chart of the study

systems. A radiology research fellow (T.N.B.) searched the radiologists' reports from consecutive clinical CT examinations. Important and/or frequent diagnoses were tallied at least one time. For difficult lesions (e.g. bone metastasis, lymph nodes and adrenal lesions) more examples were selected to supplement the dataset. For lesions not found in consecutive clinical CT examinations, examples were obtained from subspecialist radiologists of the area required. All selected CT examinations that were considered representative by an abdominal radiologist (J.S.) were included and subsequently anonymised.

Feedback training and test cases

The local medical ethics committee waived informed consent for this specific study because the 200 feedback training cases and 40 test cases were derived from three existing CT colonography study databases containing both symptomatic and asymptomatic individuals [15–17]. The local medical ethics committee has approved these studies, in which all included study participants gave informed consent. These studies were performed with the same 4-slice and 64-slice CT systems, with which 65 and 175 unique CT colonograms respectively were assessed. All CT colonograms were obtained in both prone and supine position, using low-dose protocols and without intravenous contrast medium. Participants underwent an iodine contrast bowel preparation, in one study combined with barium.

From the three studies all CT colonograms with either an E3 or E4 finding according to the C-RADS classification were included in the dataset (Table 2) [15–17]. For each CT colonogram with an E3 or E4 finding, two consecutive enrolled CT colonograms without E3 or E4 findings were

Table 1 Extended version of C-RADS classification for radiographers. This version is based on the original classification as proposed by Zalis et al. [20]

| Organ | Finding | C-RADS |
|---------------|--|-----------------|
| Liver | Calcification | E2 |
| | Simple cyst | E2 |
| | Cirrhosis | E3 |
| | Enlarged | E3 |
| | Solid mass | E3 |
| | Solid mass in cirrhotic liver | E4 |
| | Multiple solid masses | E4 |
| Biliary | Steatosis | E2 |
| | Aerobilia | E2 |
| | Cholecystitis | E3 |
| | Calcified gallbladder wall | E2 |
| | Dilated biliary tract | E3 |
| | Gallbladder hydrops | E3 |
| | Stones/sludge | E2 |
| Pancreas | Tumour | E4 |
| | Atrophy | E2 |
| | Calcifications (>1) | E3 |
| | Cyst | E3 |
| | Dilated pancreatic duct | E3 |
| Spleen | Oedema | E3 |
| | Solid mass | E4 |
| | Accessory spleen | E1 ^a |
| | Calcification(s) | E2 |
| | Cyst | E2 |
| Adrenal | Enlarged | E3 |
| | Solid mass | E4 |
| | Hyperplasia | E2 |
| | Mass/adenoma 1–4 cm | E3 |
| | Mass/adenoma >4 cm | E4 |
| Urinary tract | Bladder stone | E2 |
| | Calcification parenchyma | E2 |
| | Complex cyst | E3 |
| | Dilatation urinary tract | E4 |
| | Fatty lesion (angiomyolipoma) | E3 |
| | Horseshoe kidney | E2 |
| | Mono kidney | E2 |
| | Polycystic kidney disease ^b | E4 |
| | Small kidney | E2 |
| | Small kidneys ^b | E4 |
| | Simple cyst | E2 |
| | Solid renal mass | E4 |
| | Stone <5 mm in urinary tract | E2 |
| | Stone >5 mm in urinary tract | E3 |
| Aneurysms | Aorta 3–5 cm | E3 |
| | Aorta >5 cm | E4 |
| | Iliac >1.5 cm | E3 |
| | Iliac >2.5 cm | E4 |

Table 1 (continued)

| Organ | Finding | C-RADS |
|---------------|---|-----------------|
| | Hepatic, splenic or renal artery | E3 or E4 |
| Lymph nodes | Enlarged | E4 |
| GI | Diverticulosis | E2 ^a |
| | Tumour (extracolonic) | E4 |
| Hernia | Abdominal wall | E2 |
| | Diaphragmatic | E2 |
| | Inguinal | E2 |
| Prostate | Umbilical | E2 |
| | Calcifications | E2 ^a |
| Ovaries | Hypertrophy | E2 |
| | Signs of malignancy | E4 |
| | Complex cyst | E3 |
| Uterus | Enlarged ovary | E4 |
| | Simple cyst <3 cm | E2 |
| | Simple cyst >3 cm | E3 |
| Ascites | Solid mass | E4 |
| | Enlarged | E2 |
| Pulmonary | Masses | E2 or E3 |
| | Small amount in fertile women (Douglas) | E2 |
| | Other | E3 |
| | Atelectasis | E2 or E3 |
| | Bullae/emphysema | E2 |
| | Calcified nodule | E2 |
| | Ground glass | E3 |
| | Ground glass in small amount posteriorly located | E2 |
| | Nodule <4 mm | E2 |
| | Nodule 4–9 mm | E3 |
| Nodule ≥10 mm | E4 | |
| Heart | Pleural thickening with or without calcifications | E2 |
| | Pleural effusion | E3 |
| | Tumour or multiple nodes | E4 |
| Skeletal | Pericardial effusion | E3 |
| | Coronary calcifications | E2 ^a |
| | Atypical cyst | E3 |
| | Cyst | E2 |
| | Degenerative changes | E2 ^a |
| | Haemangioma | E2 |
| | Possible metastasis | E4 |
| | Solid lesion | E4 |
| | Schmorl's node | E2 ^a |
| | Spondylolisthesis | E2 |

^a Not necessary to report^b In case no transplantation kidney is present

added from the same study. This resulted in a dataset of 234 CT colonograms, to which 6 consecutive cases without E3

Table 2 The original C-RADS classification as proposed by Zalis et al. and the number of CT colonograms classified E1–E4 in the initial and final test cases and the feedback training cases [20]

| C-RADS | Definition | Examples | Test cases (<i>n</i>) | Feedback cases (<i>n</i>) |
|--------|--|--|-------------------------|-----------------------------|
| E1 | Normal exam or normal variant | Retro-aortic left renal vein | 22 | 82 |
| E2 | Clinically unimportant finding | Simple cyst in the liver/kidney, gallbladder stones, vertebral haemangioma | 3 | 55 |
| E3 | Likely unimportant finding; incompletely characterised | Minimally complex or homogeneously hyperattenuating cyst | 11 | 44 |
| E4 | Potentially important finding | Kidney solid renal mass, lymphadenopathy, lung nodule ≥ 1 cm | 4 | 19 |
| Total | | | 40 | 200 |

or E4 findings were added to create a dataset of 240 cases. The 200 feedback training cases contained 44 E3 and 19 E4 cases according to the reference standard (Tables 2 and 3). The 200 feedback cases were preceded and followed by an additional set of 40 cases, which constituted the initial and final tests. The tests contained 11 E3 and 4 E4 cases (Tables 2 and 3). The CT colonograms in the two tests were identical. Cases for the final test were arranged in a randomised order using the ASELECT function in Excel 2003. Observers were not aware that the CT examinations for both tests were identical, nor that these served as a test. Both feedback training cases and test cases were anonymised, using a custom-made computer program that automatically removed all patient data.

Reference standard

All CT colonograms were prospectively read for extracolonic findings by an abdominal radiologist. This observer was not blinded to the intracolonic findings and had access to previous patient history. A second abdominal radiologist (C.Y.N.; previous experience approximately 100,000 abdominal CT examinations and >1,000 CT colonograms) retrospectively read all included anonymised CT colonograms to reduce the chance of visual errors and to ensure a correct C-RADS classification. After being informed of the possibly (E3) and probably (E4) important findings of the previous report, the second experienced abdominal radiologist gave his final judgement, which served as the reference standard in this study.

Observers

Thirteen radiographers were recruited using posters at our radiology department. After additional information, ten continued with the training program. Seven worked with CT systems during their daily work. Three radiographers had previous experience with reporting intracolonic findings,

but none of them had reported extracolonic findings before. For detailed information about the radiographers' experience see Table 4. The observers performed their training primarily in their own time.

Training program—basic training

All ten radiographers completed the basic training program. The program consisted of five components: study assignments, instruction manuals, presentation, hands-on training and watching expert reading. Radiographers were asked to study two correlative anatomy-imaging course books, including anatomical illustrations with corresponding axial CT images and to read two articles concerning the evaluation of abdominal lymph nodes and adrenal gland masses, as these are relatively frequent findings that require more insight [18, 19]. A course book about pathology in abdominal organs was given as reference. An instruction manual was provided to guide the observers through the CT examination step by step. For each organ the optimal window was indicated as well as what abnormalities to look for and what to measure. Observers also received an extended version of the C-RADS classification for extracolonic findings (Tables 1 and 2) [20].

The 16-h training program was given by an experienced abdominal radiologist and a research fellow (T.N.B.). Observers were familiar with the PACS system (Agfa IMPAX client, version 5.3) prior to the study, but received additional training on certain aspects, such as HU measurements. The C-RADS classification system was discussed followed by six normal anatomy CT examinations, of which three contained series with intravenous contrast medium, to see the anatomy in full detail [20]. Thereafter for every organ, the anatomy and all frequent and/or important extracolonic findings were discussed. Subsequently hands-on training was given using training cases with findings in that organ. Individual training was given for the most difficult 15 cases by an experienced abdominal radiologist (J.S.).

Table 3 Detailed distribution of actual extracolonic findings and data composition for the 280 cases

| | Initial and final test | 1st set | 2nd set | 3rd set | 4th set | 5th set |
|--------------------------|---|---|---|--|--|---|
| E3 findings (<i>n</i>) | 13 | 9 | 18 | 14 | 11 | 8 |
| E4 findings (<i>n</i>) | 5 | 5 | 4 | 6 | 2 | 3 |
| E3 findings | Aneurysm iliac artery (<i>n</i> =3); adrenal adenoma (<i>n</i> =4); lytic bone lesion of the ilium (<i>n</i> =1); ureteropelvic junction stenosis (<i>n</i> =1); renal irregular calcification (<i>n</i> =1); pulmonary nodule (<i>n</i> =1); aneurysm splenic artery (<i>n</i> =1); ovarian teratoma (<i>n</i> =1) | Adrenal adenoma (<i>n</i> =1); adrenal nodule (<i>n</i> =2); aortic aneurysm (<i>n</i> =2); renal stone (<i>n</i> =1); ovarian cyst (<i>n</i> =2); dilated urinary system (<i>n</i> =1) | Adrenal nodule (<i>n</i> =2); adrenal adenoma (<i>n</i> =2); enlarged spleen (<i>n</i> =3); abdominal collateral vessels (<i>n</i> =2); renal stone (<i>n</i> =2); pulmonary nodule (<i>n</i> =4); calcified renal lesion (<i>n</i> =1); extra renal pyelum (<i>n</i> =1); aneurysm splenic artery (<i>n</i> =1) | Diverticulitis (<i>n</i> =1); angioliipoma (<i>n</i> =1); renal stone (<i>n</i> =2); aortic aneurysm (<i>n</i> =1); adrenal nodule (<i>n</i> =1); ceecal soft tissue mass (<i>n</i> =1); hypodense liver lesion (<i>n</i> =1); pulmonary nodule (<i>n</i> =1); aneurysm iliac artery (<i>n</i> =1); adrenal adenoma (<i>n</i> =2); pulmonary ground glass (<i>n</i> =1); dilated pancreatic duct (<i>n</i> =1) | Aortic aneurysm (<i>n</i> =2); pancreatic calcification (<i>n</i> =1); adrenal adenoma (<i>n</i> =2); ovarian cyst (<i>n</i> =1); enlarged spleen (<i>n</i> =1); pericardial fluid (<i>n</i> =1); inhomogeneous pancreatic head (<i>n</i> =1); renal stone (<i>n</i> =2) | Adrenal nodule (<i>n</i> =4); para-aortal cyst (<i>n</i> =1); pulmonary nodule (<i>n</i> =1); aortic aneurysm (<i>n</i> =2) |
| E4 findings | Pulmonary nodule (<i>n</i> =1); aortic aneurysm (<i>n</i> =1); polycystic kidneys (<i>n</i> =1); solid renal mass (<i>n</i> =1); lymphadenopathy (<i>n</i> =1) | Mesenteric mass (<i>n</i> =1); solid renal mass (<i>n</i> =1); aneurysm iliac artery (<i>n</i> =1); hydronephrosis (<i>n</i> =1); pulmonary compaction (<i>n</i> =1) | Skeletal lesion (<i>n</i> =1); lymphadenopathy (<i>n</i> =2); solid renal mass (<i>n</i> =1) | Lymphadenopathy (<i>n</i> =2); adrenal mass (<i>n</i> =1); solid renal mass (<i>n</i> =1); pulmonary nodules (<i>n</i> =1); aortic aneurysm (<i>n</i> =1) | Aortic aneurysm (<i>n</i> =1); solid renal mass (<i>n</i> =1) | Aortic aneurysm (<i>n</i> =1); solid renal mass (<i>n</i> =1); pulmonary nodule (<i>n</i> =1) |

The number of E3 and E4 findings and the specific findings are provided for each dataset. Each dataset (initial and final test of 40 cases and five sets of feedback training cases) consisted of 40 cases

Table 4 Experience and daily work of the radiographers (R)

| Observer | Sex | Age at start of training | Years certified | Daily work | CT examinations performed | Number of CTCs assessed for intracolonic findings |
|----------|-----|--------------------------|-----------------|--------------|---------------------------|---|
| R1 | M | 30 | 4 | MRI, ED | 2,500 | 0 |
| R2 | F | 25 | 2 | Allround | 1,500 | 600 |
| R3 | F | 26 | 5.5 | CT | 16,000 | 750 |
| R4 | M | 52 | 36 | MRI, C, M, P | 0 | 0 |
| R5 | F | 28 | 1 | ED, C, M | 1,500 | 0 |
| R6 | M | 22 | 2 | CT, C | 4,000 | 0 |
| R7 | M | 56 | 39 | CT | 240,000 | 0 |
| R8 | M | 26 | 4 | CT | 12,500 | 200 |

ED Emergency department, C conventional X-ray, M mammography, P paediatric, CT computed tomography, CTC computed tomography colonography, MRI magnetic resonance imaging

Clinical background was presented where considered appropriate (e.g. clinical importance of a cystic pancreatic lesion compared with a simple renal cyst). Imaging findings were presented with emphasis on differentiating normal (E1) and frequent non-relevant extracolonic findings (E2) from other findings. We did not teach differential diagnoses for possible important extracolonic findings.

Training program—feedback

Eight of the ten radiographers reported all 280 CT colonograms (200 feedback cases and two times 40 test cases) for extracolonic findings in an electronic case record form. One could not continue for personal reasons, and one found it more time intensive than anticipated and stopped. Radiographers evaluated all examinations for C-RADS E2–E4 findings and performed size and HU measurements when appropriate. They expressed their level of confidence for every extracolonic finding on a ten-point scale and indicated whether or not they felt confident enough to read the case without a second read by a radiologist. A case was considered positively triaged by radiographers where one or more E3 findings ('E3 case') or E4 findings ('E4 case') were reported or where the need for read by a radiologist was indicated. Reading time per case was also documented.

For the 200 feedback cases, immediate feedback from the reference standard was disclosed for each case after findings had been submitted, as well as possible work-up and a final diagnosis for E3 and E4 lesions. For feedback cases observers were allowed to ask for further explanation by an abdominal radiologist; if this occurred this was documented. A third experienced abdominal radiologist (J.S.; previous experience approximately 40,000 abdominal CT examinations) evaluated the CT colonograms when a radiographer reported an E3 or E4 finding that did not correspond with the reference standard.

Statistical analysis

The findings of the observers were compared with the reference standard. For the initial and final test the sensitivity and specificity were calculated for all observers combined using the McNemar's test. An E3 or E4 case was defined as a case in which the highest classified lesion by the reference standard was an E3 or E4 finding respectively.

For individual observers and all observers combined, learning curves for case sensitivity and specificity were constructed using a moving average technique (window of 60 cases and steps of 20 cases). A learning effect for case sensitivity and specificity was studied using binary logistic regression with number of cases as the independent variable. Sensitivity and specificity pairs were plotted in receiver operating characteristic (ROC) space, with sensitivity for E3 and E4 combined as the true positive rate and the corresponding false positive rate.

A reading time analysis was performed using a moving average (window of 20 cases and steps of 10 cases). The average reading time for the initial and final 40 test cases was compared using a paired *t*-test. Reader confidence for the initial and final 40 test cases was compared with a Wilcoxon rank-sum test. All analyses were performed using SPSS 16.0.2.1. In statistical tests *P* values smaller than 0.05 were considered to indicate statistical significance.

Results

Exams

In the initial test, eight radiographers combined had a sensitivity of 52 % (46/88; 95 % CI: 42–63 %) for E3 cases and 69 % (22/32; 95 % CI: 52–86 %) for E4 cases. For the final test the average sensitivity was 70 % (62/88; 95 % CI: 61–80 %) for E3 cases and 69 % (22/32; 95 % CI: 52–86 %) for E4 cases, a significant increase of 18 % for E3 cases ($P < 0.0001$) and no

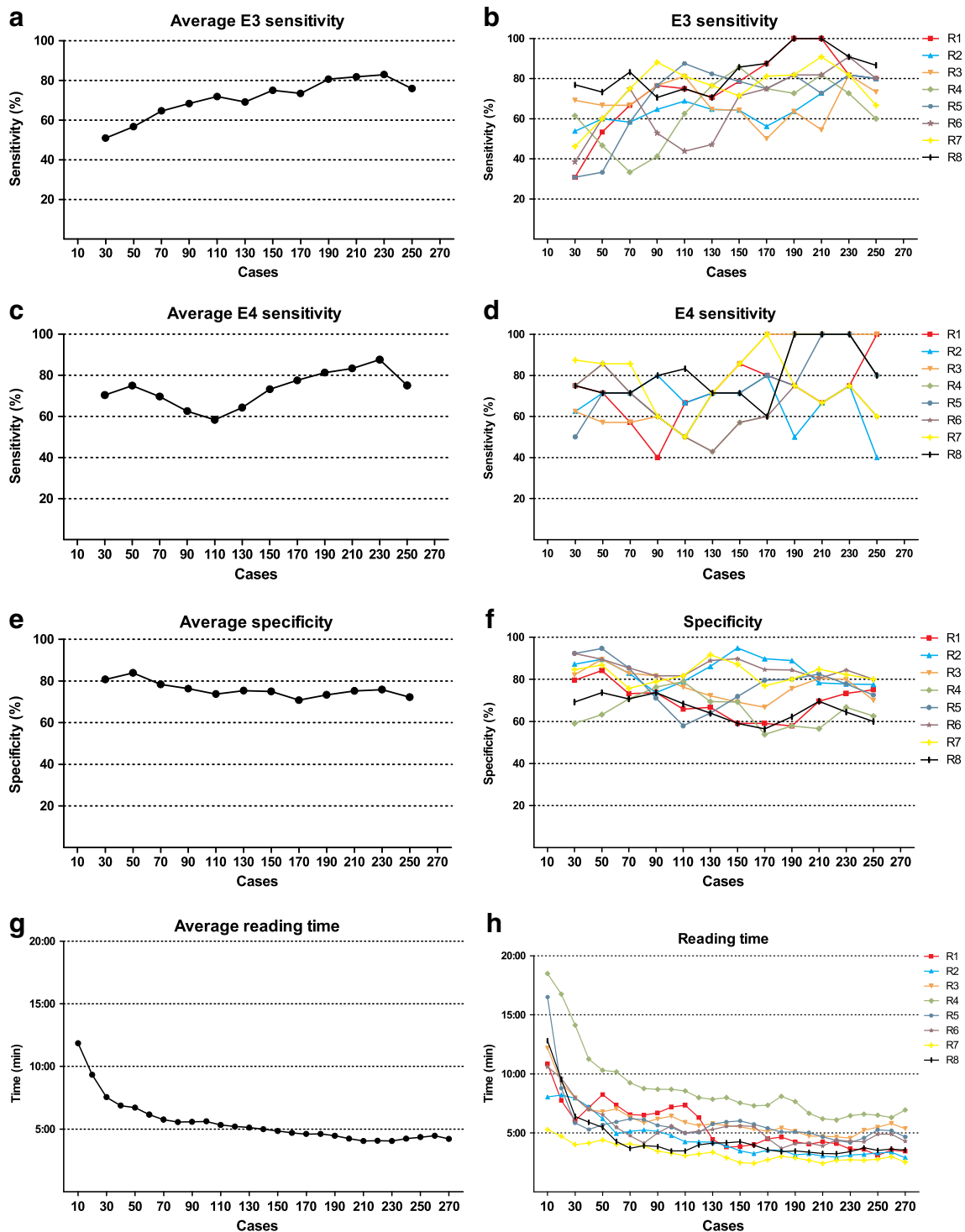


Fig. 2a–h Learning curves for radiographers. **a** Average sensitivity for E3 cases and **b** for radiographers individually. **c** Average sensitivity for E4 cases and **d** for radiographers individually. **e** Average specificity

and **f** for radiographers individually. **g** Average reading time and **h** for radiographers individually

difference for E4 cases ($P=1.00$). The combined specificity decreased from 83 % (165/200; 95 % CI: 77–88 %) in the initial test to 70 % (139/200; 95 % CI: 63–76 %) in the final test ($P<0.0001$).

The radiographers' average reading time per case decreased from 11:51 min to 4:13 min ($P<0.0001$). Median radiographers' confidence for all reported lesions increased from 8 out of 10 at the initial test to 9 out of 10 in the final test ($P<0.0001$).

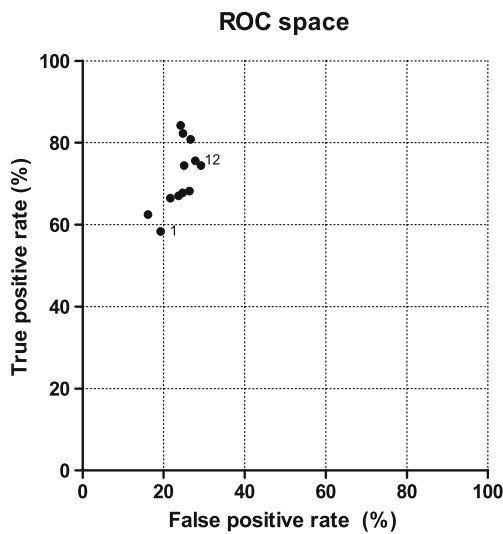


Fig. 3 Receiver operating characteristic (ROC) space for radiographers’ average proportion of true positives (sensitivity) for E3 and E4 combined versus the average percentage of false positive cases without E3 and E4 lesions. Each point in the ROC space indicates a time point in the learning curve. 1 indicates the first time point and 12 indicates the last time point

Learning curve

The case sensitivity learning curves for E3 and E4 for individual radiographers and eight radiographers combined are shown in Fig. 2a–d. The specificity learning curves are shown in Fig. 2e, f and the reading time curves in Fig. 2g, h. The ROC plot shows that with an increase in the proportion of true positives (sensitivity) the proportion of false positives also increased (Fig. 3).

In triaging E3 cases there was a significant positive learning effect for three radiographers; for E3 and E4 cases combined the learning effects were significant for two radiographers (Table 5). In correctly identifying E4 cases only, none of the radiographers achieved a significant learning effect. None of the radiographers showed a significant learning effect in specificity (Table 5).

In 15 cases, the additional findings by radiographers resulted in an altered case classification from E1 or E2 to E3. None of the cases with an altered classification were test cases. No follow-up was performed as a result of these findings.

The median number of days needed to report the 200 training cases was 57 days (range 28–115). The eight radiographers combined asked 116 questions to the abdominal radiologist about the feedback cases; these included 112 questions about radiographers’ findings that were not in the reference standard.

Discussion

This study shows that a dedicated formal training program can lead to a significant increase in the sensitivity of radiographers

Table 5 Binary logistic regression results of individual radiographers (R) for sensitivity (E3, E4 and combined) and specificity on a per-case basis. The regression coefficient B and the results of the Wald test are shown

| Radiographer | B | SE | P value |
|------------------------------|--------|-------|---------|
| E3 and E4 sensitivity | | | |
| R1 | 0.028 | 0.010 | 0.005** |
| R2 | 0.003 | 0.008 | 0.693 |
| R3 | 0.004 | 0.009 | 0.630 |
| R4 | 0.007 | 0.008 | 0.402 |
| R5 | 0.029 | 0.009 | 0.002** |
| R6 | 0.014 | 0.008 | 0.091 |
| R7 | 0.000 | 0.009 | 0.993 |
| R8 | 0.011 | 0.010 | 0.281 |
| E4 sensitivity | | | |
| R1 | 0.063 | 0.060 | 0.293 |
| R2 | -0.054 | 0.053 | 0.313 |
| R3 | 0.121 | 0.072 | 0.095 |
| R4 | 0.021 | 0.054 | 0.705 |
| R5 | 0.088 | 0.058 | 0.129 |
| R6 | -0.032 | 0.052 | 0.541 |
| R7 | -0.056 | 0.059 | 0.345 |
| R8 | 0.018 | 0.060 | 0.767 |
| E3 sensitivity | | | |
| R1 | 0.042 | 0.016 | 0.010** |
| R2 | 0.017 | 0.014 | 0.234 |
| R3 | -0.006 | 0.014 | 0.694 |
| R4 | 0.010 | 0.013 | 0.469 |
| R5 | 0.043 | 0.016 | 0.008** |
| R6 | 0.034 | 0.015 | 0.020* |
| R7 | 0.009 | 0.014 | 0.526 |
| R8 | 0.019 | 0.017 | 0.281 |
| Specificity | | | |
| R1 | -0.003 | 0.003 | 0.276 |
| R2 | -0.000 | 0.004 | 0.924 |
| R3 | -0.004 | 0.003 | 0.179 |
| R4 | -0.001 | 0.003 | 0.658 |
| R5 | -0.005 | 0.003 | 0.172 |
| R6 | -0.004 | 0.004 | 0.357 |
| R7 | -0.002 | 0.004 | 0.656 |
| R8 | -0.004 | 0.003 | 0.213 |

Positive values of B indicate a learning effect

*P < 0.05, **P < 0.01

in triaging E3 cases, without an increase in correctly triaging E4 cases. The overall triaging sensitivity and specificity at the end of the training were about 70 %, with substantial individual variation. There was a significant increase in reader confidence accompanied by a significant decrease in reading time. Specificity decreased after training, although this decrease was not significant for individual observers.

As shown previously, radiographers are capable of reading CT colonography for intracolonic findings [4, 5], but our results did not show high enough accuracy after training to facilitate an effective triage for extracolonic findings. The learning effect for triaging E3 findings at the end of training suggests that the maximum sensitivity may not yet have been reached with the number of training and feedback cases we offered. We were disappointed to learn that there was no learning effect for E4 cases and that specificity was low. The large number of different lesions, the differential diagnoses and the conspicuity of some lesions are plausible causes for the difficulty that radiographers had in reporting extracolonic findings. This is different from other learning curve studies concerning the evaluation of one organ with a limited number of differential diagnoses (e.g. intracolonic findings).

A recent study by Liedenbaum et al. showed that for reporting intracolonic findings, novice readers needed on average 164 CT colonograms with feedback to reach a desired level of competence [5]. In this study one of three radiographers reached sufficient competence within 164 CT colonography cases. As in our study, there were considerable differences in the individual performances of the radiographers. Some radiographers achieved very good sensitivity for the final test, but the specificity of these observers was below average. The increased reader confidence and reduced interpretation time observed in our study are encouraging findings, but of lesser importance than accuracy.

Apart from the possible lack of sufficient training, other important factors that may affect radiographer's capability in reading extracolonic findings are aptitude and level of professional education. Based on this study we cannot determine whether the latter two factors are principal factors that—despite a modified training program—will preclude radiographer reading of extracolonic findings.

We invested substantial effort in the training program. We ensured the training program was not just formal and reproducible, but also extensive. It included all frequent and important extracolonic findings. We attempted to make the training fit well with radiographers' knowledge. For an optimal learning effect, radiographers were blinded from feedback until submitting the report of a case. Only highly motivated radiographers participated, because free time had to be invested. We helped the radiographers with classifying lesions, by providing an extensive C-RADS classification. Furthermore, new findings of radiographers were evaluated by an independent experienced abdominal radiologist.

A number of limitations have to be acknowledged. The dataset proved to be too small, as sensitivity and specificity did not seem to reach a plateau phase. We estimated the number of cases needed to triage important findings, without data to support our assumptions. Our training program may also not have been optimal for this specific group of

observers. It was developed based on earlier experience by several groups who were successful in training non-radiologists to read imaging studies [5, 11, 21]. We combined both knowledge transfer and hands-on experience in a structured program with direct feedback. Additional findings by the eight radiographers combined in 15 cases altered the case classification from E1 or E2 into E3. All of these CT colonograms were feedback training cases; they did not influence the estimated sensitivity of the examination cases. The order of training cases was not randomised to make sure that lesions were well distributed over the dataset. The prevalence of extracolonic findings in this dataset was higher than in screening and results may be different in a screening population. Finally, we would like to mention that we performed multiple tests and some of the statistically significant results may have occurred by chance.

Our findings suggest that a radiographer-only strategy for screening CT colonography is not a viable option to further reduce the costs of screening CT colonography. Based on the triaging results shown in this study, the number of CT colonograms that would have to be read by a radiologist in a screening setting would still be around 40 %, while probably missing lesions. Even if performance had been comparable to that of a radiologist, medicolegal aspects of having CT colonography examinations read by radiographers and a radiographer-only reading strategy may have been a hurdle in many countries.

Acknowledgements This study was presented at the European Congress of Radiology 2012. Funding was provided by a NutsOhra Foundation grant. Furthermore we would like to acknowledge H.J.P.P. Eskes, K. Heijmen, D. Kauffman, K.J. Lobe, N.H.J. Lobe, C.C.M. Lute, M. Poulus and A. Tuit for their efforts.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- McFarland EG, Levin B, Lieberman DA et al (2008) Revised colorectal screening guidelines: joint effort of the American Cancer Society, U.S. Multisociety Task Force on Colorectal Cancer, and American College of Radiology. *Radiology* 248:717–720
- Lansdorp-Vogelaar I, Knudsen AB, Brenner H (2011) Cost-effectiveness of colorectal cancer screening. *Epidemiol Rev* 33:88–100
- Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Boer R, Wilschut J, Habbema JD (2009) At what costs will screening with CT colonography be competitive? A cost-effectiveness approach. *Int J Cancer* 124:1161–1168
- Jensch S, van Gelder RE, Florie J et al (2007) Performance of radiographers in the evaluation of CT colonographic images. *AJR Am J Roentgenol* 188:W249–W255
- Liedenbaum MH, Bipat S, Bossuyt PM et al (2011) Evaluation of a standardized CT colonography training program for novice readers. *Radiology* 258:477–487

6. Kim DH, Pickhardt PJ, Taylor AJ et al (2007) CT colonography versus colonoscopy for the detection of advanced neoplasia. *N Engl J Med* 357:1403–1412
7. van den Biggelaar FJ, Nelemans PJ, Flobbe K (2008) Performance of radiographers in mammogram interpretation: a systematic review. *Breast* 17:85–90
8. Brealey S, Scally A, Hahn S, Thomas N, Godfrey C, Coomarasamy A (2005) Accuracy of radiographer plain radiograph reporting in clinical practice: a meta-analysis. *Clin Radiol* 60:232–241.10
9. Bradley AJ, Rajashanker B, Atkinson SL, Kennedy JN, Purcell RS (2005) Accuracy of reporting of intravenous urograms: a comparison of radiographers with radiology specialist registrars. *Clin Radiol* 60:807–811
10. Law RL, Titcomb DR, Carter H, Longstaff AJ, Slack N, Dixon AR (2008) Evaluation of a radiographer-provided barium enema service. *Colorectal Dis* 10:394–396
11. Brandt A, Andronikou S, Wieselthaler N et al (2007) Accuracy of radiographer reporting of paediatric brain CT. *Pediatr Radiol* 37:291–296
12. Xiong T, Richardson M, Woodroffe R, Halligan S, Morton D, Lilford RJ (2005) Incidental lesions found on CT colonography: their nature and frequency. *Br J Radiol* 78:22–29
13. Flicker MS, Tsoukas AT, Hazra A, Dachman AH (2008) Economic impact of extracolonic findings at computed tomographic colonography. *J Comput Assist Tomogr* 32:497–503
14. Gluecker TM, Johnson CD, Wilson LA et al (2003) Extracolonic findings at CT colonography: evaluation of prevalence and cost in a screening population. *Gastroenterology* 124:911–916
15. Liedenbaum MH, van Rijn AF, de Vries AH et al (2009) Using CT colonography as a triage technique after a positive faecal occult blood test in colorectal cancer screening. *Gut* 58:1242–1249
16. Jensch S, de Vries AH, Peringa J et al (2008) CT colonography with limited bowel preparation: performance characteristics in an increased-risk population. *Radiology* 247:122–132
17. Stoop EM, de Haan MC, de Wijkerslooth TR et al (2012) Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial. *Lancet Oncol* 13:55–64
18. Einstein DM, Singer AA, Chilcote WA, Desai RK (1991) Abdominal lymphadenopathy: spectrum of CT findings. *Radiographics* 11:457–472
19. Johnson PT, Horton KM, Fishman EK (2009) Adrenal mass imaging with multidetector CT: pathologic conditions, pearls, and pitfalls. *Radiographics* 29:1333–1351
20. Zalis ME, Barish MA, Choi JR et al (2005) CT colonography reporting and data system: a consensus proposal. *Radiology* 236:3–9
21. Dachman AH, Kelly KB, Zintsmaster MP et al (2008) Formative evaluation of standardized training for CT colonographic image interpretation by novice readers. *Radiology* 249:167–177